Munich Center for Mathematical Philosophy (MCMP)
Faculty of Philosophy, Philosophy of Science and Religious Studies
Ludwig-Maximilians-Universität München(LMU - Munich)

What Was True Yesterday Might Not Be True Today

Ana Flávia de Faria Cholodovskis

Munich - 2025

Supervised by **Prof. DDr. Hannes Leitgeb**

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Fakultät für Philosophie, Wissenschaftstheorie und Religionswissenschaft

What Was True Yesterday Might Not Be True Today

Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Fakultät für Philosophie, Wissenschaftstheorie und Religionswissenschaft
der Ludwig-Maximilians-Universität München

vorgelegt von
Ana Flávia de Faria Cholodovskis
aus Santo Antônio do Monte (Minas Gerais), Brasilien

München, 2025

Erstgutachter: Prof. DDr. Hannes Leitgeb
Zweitgutachter: Prof. Dr. Peter Verdée
Tag der mündlichen Prüfung (Disputation): 07.07.2025

Contents

Abstract

Belief change in context-sensitive environments poses significant challenges, particularly due to the role of indexicals in shaping reasoning and arguments. Indexicals such as “I,” “here,” and “now” introduce unique challenges to traditional models, as their meanings are deeply tied to the agent’s perspective. These challenges reveal a fundamental tension in belief revision: the need to account for the world’s objective structure and the agent’s subjective position within it. A central issue in this discussion is the definition of context and its influence on asserting truths. Context-sensitive reasoning often exposes gaps in existing frameworks, where the interplay between external facts and internal perspectives remains underexplored. To address this, the thesis introduces a Two-Dimensional approach that considers both the agent’s knowledge of the world and their situated perspective, providing a structured lens for analyzing belief change dynamics. The “Alone in the Wilderness” scenario presented in the thesis serves as a guiding example throughout this work, illustrating the limitations of traditional models and the necessity of context-sensitive frameworks. This scenario captures the profound interplay of belief, knowledge, and perspective, demonstrating how an agent’s understanding of their situation evolves when isolated from external cues. Through such cases, the thesis showcases the applicability of Two-Dimensional frameworks while critically examining their limitations, ultimately proposing new directions for addressing the challenges of belief revision in context-sensitive domains.

Imagine waking up alone in the wilderness, lost and confused. In your hand is a letter explaining that you must survive until rescue arrives. You have no idea when or even if they will come, where you are, or what day it is. Yet, you can think and articulate sentences such as, “Today is windy,” or, “I didn’t eat yesterday,” even without knowing the exact date. You don’t need to know which day it is on the calendar to determine if it’s windy or if you’ve gone without food for a day. Moreover, you can effortlessly shift from using “today” to “yesterday” to describe an event from the prior day without confusion. This scenario highlights a fascinating aspect of human cognition: our ability to use indexicals expressions like “today,” “yesterday,” “I,” “here,” and “now” that are inherently tied to specific contexts, even when the precise details of the context remain unknown. Such expressions are context-sensitive, enabling individuals to describe their beliefs and experiences without needing exhaustive knowledge about their surroundings or even themselves. In our wilderness example, a person could write sentences such as, "I am here," or, "It is cold now," without knowing their exact location or personal details.

Indexical expressions play a central role in how we describe and assert beliefs, particularly those about ourselves. This thesis explores the nature of indexical reasoning and its implications for Belief Revision. In the wilderness, every decision you make relies on reasoning under uncertainty: Should you ration your food? Should you move or stay put? These decisions depend on the context you construct, using available evidence and linguistic expressions to frame your thoughts. Indexicals are crucial in this process, as they allow you to structure beliefs and make decisions even with incomplete information. By examining the wilderness scenario, this thesis investigates the interplay between context-sensitive reasoning and Belief Revision. How can beliefs be updated when new information arises? For instance, discovering footprints nearby could significantly alter your beliefs about rescue or potential danger. This work connects such questions to formal systems of reasoning, providing tools to model and analyze the dynamics of belief and knowledge in context-dependent situations. Beyond decision-making scenarios, the process of revising beliefs with indexical information raises deeper questions about how defeasible reasoning interacts with context-sensitivity. To address these issues, it is crucial to identify the underlying problem and its theoretical foundations.

This thesis explores the role of context-sensitive reasoning in Belief Revision, emphasizing the need to integrate Two-Dimensional Semantics into frameworks for dynamic belief change. Traditional models of Belief Revision typically assume that the context in which beliefs are held remains fixed. However, as demonstrated by authors such as David Lewis and Robert Stalnaker, beliefs are often shaped not only by the external world but also by the agent's perspective and context. This dual dependency reveals the limitations of one-dimensional approaches and highlights the necessity of a more refined framework. Building on insights from Two-Dimensional Semantics, this thesis distinguishes between two crucial components of belief: those that pertain to the external world and those grounded in the agent's context. By analyzing the interplay between these two types of contexts, the thesis establishes the conditions under which a Two-Dimensional framework becomes indispensable for accurately modeling belief change as sometimes Two-Dimensional models are necessary for Belief Revision and in others, simpler one-dimensional approaches suffice. Through this analysis, the thesis bridges the gap between classical Belief Revision theory and the complexities of everday reasoning, offering a more comprehensive understanding of how context-sensitive beliefs evolve in dynamic environments.

By presenting a novel framework for context-sensitive Belief Revision, grounded in the integration of Two-Dimensional Semantics and the theory of indexicals and self-locating beliefs, this work aims to address the issue of how one can rationally revise their beliefs in the presence of indexical expressions. The work begins with an exploration of the theoretical foundations, introducing key concepts from Belief Revision and modal logic. Using illustrative examples, the introductory chapter sets the stage by demonstrating how context influences an agent's ability to revise beliefs, particularly when faced with uncertainty about their position in time, space, or knowledge. This section provides the necessary background for understanding the challenges addressed in the subsequent chapters and offers a comprehensive literature review.

The Part I of the thesis delves deeper into the mechanics of belief dynamics, and the formal proposals to deal with defeasible reasoning. Through a formalization of these intensions, this section builds a foundation for understanding how context-sensitive reasoning operates and why it is indispensable for addressing complex scenarios. This is essential to understand the construction and limitations of a Two-Dimensional framework for Belief Revision. This approach is exemplified through real-life-inspired scenarios, such as John the Ranger, who must navigate uncertainty about his location in the wilderness, and Jane the Detective, who adapts her beliefs as new evidence reshapes her understanding of events. These case studies serve as practical illustrations of how the framework operates, demonstrating its flexibility and accuracy in capturing dynamic revisions.

In the Part II of the thesis delves into the intricate phenomena of indexicality, laying the foundation for understanding how beliefs involving indexicals are formed, maintained, and revised. One chapter focus on the unique challenges posed by indexical expressions, which depend on the speaker's context to convey meaning. This chapter develops a systematic approach to formalizing indexicals necessary if we want to incorporate Indexicals into the framework of Belief Revision, emphasizing their dependence on both linguistic and epistemic contexts. Following this, we discuss in detail what a Context is and how contexts can shape and constrain an agent's beliefs and reasoning. It introduces the new concepts of *K-contexts* and *S-contexts* to distinguish between knowledge about the world and self-locating perspectives, offering a structured way to model how contexts influence belief dynamics. Together, these chapters establish the theoretical and formal tools necessary for addressing the complexities of context-sensitive reasoning, setting the stage for the practical applications and advanced modeling discussed in later parts of the thesis.

Part III of the thesis builds upon the theoretical groundwork laid earlier, advancing approaches to context-sensitivenes from the probabilistic side. It

refines traditional models by integrating probabilistic reasoning and addressing limitations in handling conflicting or uncertain information. This section emphasizes the importance of flexibility in belief dynamics, particularly in accommodating context-sensitive elements. The thesis also incorporates probabilistic reasoning to enhance the modeling of Belief Revision, particularly in scenarios involving uncertainty and conflicting information. By integrating probabilities into the revision process, the framework allows for more nuanced adjustments to belief strength, reflecting the varying degrees of confidence agents may hold. This probabilistic perspective is essential for addressing real-world reasoning challenges, where evidence is often incomplete or context-sensitive. It further strengthens the framework's applicability to dynamic environments and decision-making processes. A key application explored in the thesis is the Sleeping Beauty problem, which illustrates the complex interplay between indexicals, context, and probabilistic reasoning. This scenario demonstrates how an agent's self-locating beliefs—such as those tied to temporal contexts—can influence their epistemic updates. The analysis reveals the practical utility of the Two-Dimensional framework in resolving ambiguities arising from shifting contexts, providing a concrete example of how the proposed model captures the nuances of context-sensitive reasoning in intricate cases.

A significant focus of Part IV is the process of belief revision involving indexical expressions. Using previously introduced scenarios, it demonstrates how agents reconcile changes in their self-locating perspectives with their knowledge about the world. This section formalizes the interplay between different dimensions of belief, providing practical insights into resolving inconsistencies in context-sensitive reasoning. Finally, in Part IV we critically evaluate whether a Two-Dimensional framework is necessary for addressing complex cases involving indexicals. It balances the theoretical advantages of this approach against potential limitations, reflecting on its broader implications for both philosophical inquiry and practical applications. By offering a robust and adaptable framework for modeling context-sensitive belief dynamics, this section concludes the thesis with a comprehensive outlook on its contributions and directions for future research. In the final chapters, the thesis develops a detailed account of how beliefs involving indexicals and context-sensitive information can be revised systematically. It introduces a Two-Dimensional framework to capture the nuances of these revisions. The final chapters apply this framework to various illustrative cases, offering practical insights into the mechanics of belief adjustment as well as critically evaluating the necessity of Two-Dimensional Semantics in addressing such revisions, balancing the theoretical benefits against practical limitations. Together, these chapters provide a comprehensive roadmap for understanding and implementing context-sensitive Belief Revision while reflecting on the broader implications of adopting a Two-Dimensional approach. The thesis concludes by synthesizing its contributions, offering both a theoretical framework and practical tools for modeling belief dynamics. It emphasizes the significance of integrating context and agent-centered perspectives into Belief Revision, contributing to ongoing discussions in epistemology, logic, and beyond. By bridging foundational theory with applied reasoning, this work opens new avenues for understanding the complexity of everyday reasoning and sets the stage for future advancements in this vibrant area of research.

In summary, this thesis explores the dynamics of Belief Revision in context-sensitive environments, emphasizing the role of indexicals such as “I,” “here,” and “now.” These expressions, deeply tied to the agent’s perspective, reveal challenges in reconciling external facts with internal viewpoints. By integrating Two-Dimensional Semantics, the thesis develops a framework distinguishing between knowledge about the world and self-locating perspectives, addressing the interplay of context, belief, and reasoning. Illustrative scenarios, including probabilistic reasoning and the Sleeping Beauty problem, demonstrate the framework’s flexibility in handling uncertainties and shifting contexts. This work advances the understanding of belief change, bridging classical theories with the complexities of everyday reasoning and offering insights into fields such as epistemology, cognitive science, and artificial intelligence.

Zusammenfassung

Die Arbeit zeigt, dass die Zwei-Dimensionale Semantik vor allem dann erforderlich ist, wenn die Revision von Überzeugungen komplexe, kontextabhängige Elemente berücksichtigt, wie sie bei Indexikalität und selbstlokalisierenden Überzeugungen vorkommen. In weniger komplexen Fällen, in denen der Kontext eine untergeordnete Rolle spielt, reicht ein eindimensionaler Ansatz aus, der eine effizientere und einfachere Lösung bietet. Schließlich wird die Dissertation die verschiedenen Möglichkeiten zur Implementierung von 2D-Modellen in der Überzeugungsrevision und deren Grenzen aufzeigen. Es wird auch die Frage behandelt, wann eine Vereinfachung auf eindimensionale Modelle möglich und rational ist.

Zu Beginn der Dissertation wird die theoretische Grundlage der Belief Revision und der Zwei-Dimensionalen Semantik gelegt. Im klassischen Ansatz der Belief Revision werden Überzeugungen basierend auf neuen Informationen aktualisiert, wobei der Kontext als konstant angenommen wird. Dieser Ansatz geht davon aus, dass die Welt und das Wissen des Agenten unverändert sind, abgesehen von den neuen Informationen, die eingeführt werden. Die Dissertation zeigt jedoch auf, dass viele Überzeugungen kontextabhängig sind und von der Perspektive des Agenten abhängen. Diese Perspektive wird oft durch Indexikalität und selbstlokalisierende Überzeugungen geprägt, bei denen nicht nur das Wissen über die Welt relevant ist, sondern auch der Zeitpunkt, der Ort und der Agent selbst. Die herkömmlichen Modelle der Belief Revision stoßen hier an ihre Grenzen, da sie die dynamische Natur von Kontexten und die damit verbundene Unsicherheit nicht ausreichend abbilden. Zwei-Dimensionale Semantik wird hier als Modell vorgestellt, das zwischen der tatsächlichen Welt und der Perspektive des Agenten unterscheidet, wobei sowohl die Welt als auch der Kontext des Agenten in die Beliefs einfließen.

Im ersten Kapitel werden die Grundlagen der Zwei-Dimensionalen Semantik erklärt, die ursprünglich von Kaplan [46], Chalmers [24] und Stalnaker [100] entwickelt wurde. Diese Theorien untersuchen die Unterscheidung zwischen der tatsächlichen Welt und möglichen Welten aus der Perspektive eines Individuums. Der Hauptbeitrag des ersten Kapitels ist es, die Notwendigkeit der Unterscheidung zwischen K-Intensionen (weltbezogene Überzeugungen) und

S-Intensionen (kontextbezogene Überzeugungen) zu verdeutlichen. K-Intensionen beschreiben Überzeugungen, die ausschließlich auf der Welt basieren, wie sie aus einer objektiven Perspektive betrachtet wird, während S-Intensionen den Kontext des Agenten, insbesondere die Position des Agenten in der Welt (Zeit, Ort, und individuelle Perspektive), in Betracht ziehen. Diese Unterscheidung ist von zentraler Bedeutung, da sie es ermöglicht, kontextabhängige Überzeugungen korrekt zu modellieren und die Komplexität der realen Welt, in der der Agent lebt, adäquat zu berücksichtigen. In vielen Fällen reicht es nicht aus, nur die Welt zu betrachten, sondern auch die Position des Agenten in der Welt, die den Kontext für die jeweiligen Überzeugungen bildet.

Das erste Kapitel führt daher die Grundlagen der Zwei-Dimensionalen Semantik in die Belief Revision ein und diskutiert die Relevanz von Kontext und Perspektive für die Aktualisierung von Überzeugungen. Es wird gezeigt, dass die Einführung von S-Intensionen und K-Intensionen eine präzisere Modellierung von Belief Revision ermöglicht, insbesondere in Szenarien, in denen der Agent unsicher ist, nicht nur über die Welt, sondern auch über seine Position innerhalb der Welt. Dies ist besonders relevant für die Behandlung von Indexikalität und selbstlokalisierenden Überzeugungen, die in vielen alltäglichen und philosophischen Kontexten vorkommen. Weiterhin wird ein methodischer Rahmen für die Integration von Zwei-Dimensionaler Semantik in die Belief Revision vorgestellt, der aufzeigt, wie Überzeugungen effizient aktualisiert werden können, wenn sowohl die Welt als auch der Kontext berücksichtigt werden müssen.

Im Verlauf der Arbeit wird dann detailliert untersucht, wie dieser erweiterte Rahmen in praktischen Beispielen wie der „Alone in the Wilderness“ -Erzählung oder dem „Sleeping Beauty“ -Paradoxon angewendet werden kann. Anhand dieser Beispiele wird gezeigt, dass die Unterscheidung zwischen K- und S-Intensionen entscheidend dafür ist, wie wir Überzeugungen aktualisieren und neue Informationen einbeziehen, die in einem bestimmten Kontext Sinn machen. Diese Anwendung wird sowohl theoretisch als auch praktisch weiter vertieft, um das Potenzial der Zwei-Dimensionalen Semantik für die Belief Revision zu demonstrieren.

Ein zentrales Beispiel, das in dieser Dissertation behandelt wird, ist das Beispiel von John, dem Ranger, der sich allein in der Wildnis befindet. In diesem Szenario geht es nicht nur darum, welche Welt John für wahr hält, sondern auch, welche Welt in Bezug auf seine spezifische Situation (seinen „zentralen Kontext“) relevant ist. John ist sich nicht sicher, wo er sich befindet, was zu einer erheblichen Unsicherheit bezüglich seiner Überzeugungen führt. Während er die Welt als „möglich“ oder „wahrscheinlich“ betrachtet, verändert sich seine Überzeugung, je nachdem, welche Informationen er in Bezug auf seine Position und Zeit erhält. Diese Unsicherheit in Johns Wahrnehmung wird durch die Komplexität von kontextabhängigen Überzeugungen und deren Überarbeitung noch verstärkt, was das Beispiel zu einem idealen Fall für die Anwendung der Zwei-Dimensionalen Semantik macht.

In Johns Fall treten zwei Arten von Unsicherheit auf. Einerseits ist er unsicher über die genaue Welt, in der er sich befindet (unsicher über den Ort). Andererseits ist er unsicher über den Kontext (unsicher über die aktuelle Zeit und seine eigene Position). Diese doppelte Unsicherheit macht das Beispiel ideal, um die Unterscheidung zwischen K- und S-Intensionen zu demonstrieren. Johns Unsicherheit über den Ort ist eine klassische Fall von Unsicherheit über die Welt, die durch die herkömmliche AGM-Überarbeitungslogik behandelt werden kann. Seine Unsicherheit über den Kontext erfordert jedoch eine Zwei-Dimensionale Betrachtung, da sowohl die Welt als auch seine Position innerhalb dieser Welt berücksichtigt werden müssen. Diese Unterscheidung ist besonders wichtig, da Johns Überzeugungen nicht nur durch objektive Welten, sondern auch durch seinen spezifischen Standort und Zeitpunkt beeinflusst werden.

Johns Unsicherheit kann durch die Unterscheidung zwischen K-Intensionen und S-Intensionen präzisiert werden. Die K-Intensionen beziehen sich auf Johns Überzeugungen über die Welt, unabhängig von seinem Standort und seiner Position. In diesem Fall würde er eine Überzeugung darüber haben, dass er sich möglicherweise in einem bestimmten geografischen Gebiet befindet, aber er könnte nicht sicher sein, welches Gebiet es genau ist. Diese Überzeugung lässt sich durch die herkömmliche, eindimensionale AGM-Logik überarbeiten, da sie sich nur auf den Inhalt der Welt bezieht.

Die S-Intensionen hingegen beziehen sich auf Johns Unsicherheit über seine Position innerhalb der Welt – also die Frage, wo er sich genau befindet und zu welchem Zeitpunkt. Diese Art von Unsicherheit ist kontextabhängig, da sie von der Perspektive und dem Wissen von John abhängt. Um die Unsicherheit zu überarbeiten, muss man sowohl die Welt als auch den Kontext berücksichtigen, was eine Zwei-Dimensionale Modellierung erfordert. Wenn John z.B. durch die Beobachtung eines markanten geografischen Merkmals (z.B. einen Berg oder einen Fluss) seine Position zu einer bestimmten Zeit und an einem bestimmten Ort präzisiert, verändert sich seine Überzeugung über den „Ort“ und seine Zeit innerhalb der Welt. Diese kontextabhängige Unsicherheit kann nicht allein mit einem eindimensionalen Belief Revision-Ansatz erfasst werden.

Um die Überzeugungsrevision in Johns Fall adäquat zu modellieren, ist es notwendig, beide Dimensionen der Zwei-Dimensionalen Semantik anzuwenden. Dies bedeutet, dass die Überzeugungen über die Welt sowohl auf den aktuellen Wissensstand (die K-Intension) als auch auf den aktuellen Kontext (die S-Intension) zurückgeführt werden müssen. Ein einfaches Beispiel hierfür ist die Unterscheidung zwischen der Überzeugung, dass „heute ein schöner Tag ist“ (S-Intension) und der Überzeugung, dass „es morgen regnen wird“ (K-Intension), wobei letzteres nicht von der persönlichen Position oder Zeit des Agenten abhängt.

Wenn John z.B. glaubt, dass er in einem bestimmten geografischen Gebiet ist, aber nicht sicher ist, ob er sich in der Nähe eines bestimmten Berges oder Flusses aufhält, ist seine Unsicherheit über die Welt eine K-Intension. Wenn er jedoch glaubt, dass er an einem bestimmten Punkt der Zeit in dieser Gegend ist, wird diese Unsicherheit durch die S-Intension erfasst. Ein vollständiges Modell seiner Überzeugungen erfordert daher eine Kombination dieser beiden Dimensionen, die die Unsicherheit sowohl über die Welt als auch über den Kontext berücksichtigen.

Diese Unterscheidung ist von zentraler Bedeutung, da sie aufzeigt, dass die Überzeugungsrevision in Fällen von kontextabhängigen Überzeugungen komplexer wird, als es die klassische AGM-Logik annehmen würde. Es reicht nicht aus, nur die Welt zu betrachten, sondern es muss auch die Position des Agenten innerhalb dieser Welt berücksichtigt werden. Das Beispiel „Allein in der Wildnis“ veranschaulicht, wie die Zwei-Dimensionale Semantik ein präziseres und nuancierteres Modell für die Überarbeitung von Überzeugungen bietet, insbesondere in Situationen, in denen sowohl der Kontext als auch die Welt berücksichtigt werden müssen. Dies ist von großer Bedeutung für die Entwicklung einer robusteren Theorie der Belief Revision, die auch die Rolle von Selbstlokalisierung und Kontextsensitivität in die Überlegungen einbezieht.

Die Dissertation verwendet auch andere Beispiele, um das Konzept der Zwei-Dimensionalen Semantik weiter zu illustrieren. Ein weiteres Beispiel ist das „Sleeping Beauty“ -Paradox, das mit Belief Revision und Unsicherheit arbeitet. In diesem Paradoxon wird die Überlegung aufgeworfen, wie man die Überzeugungen einer Person über den Zeitpunkt ihrer Wachsamkeit aktualisieren sollte, wenn neue Informationen vorliegen. Das Paradox zeigt auf, wie Überzeugungen über die Welt und der Kontext, in dem diese Überzeugungen bestehen, differenziert betrachtet werden müssen.

Im Fall der „Sleeping Beauty“ -Paradoxie geht es darum, wie Beauty ihre Überzeugungen aktualisieren sollte, wenn sie durch ein Experiment (das häufig in der Philosophie verwendet wird) ihren Zustand ändert. Das Experiment besagt, dass sie sich an einem Tag wiedererkennt, aber ihre Erinnerung an die vorangegangenen Tage gelöscht wurde. Beauty befindet sich daher in einer Situation der Unsicherheit über den Zeitpunkt ihrer Wachsamkeit. Sie weiß nicht, ob sie sich am ersten oder am zweiten Tag des Experiments befindet, und ihre Überzeugung über den Zeitpunkt der Wachsamkeit muss überarbeitet werden, wenn neue Informationen über den Verlauf des Experiments hinzukommen.

Das Paradoxon zeigt, dass Überzeugungen über die Welt (z.B. der Glaube, dass Beauty entweder am ersten oder am zweiten Tag wach ist) und der Kontext, in dem diese Überzeugungen bestehen (z.B. die Unterscheidung zwischen dem ersten und zweiten Tag des Experiments), differenziert betrachtet werden müssen. Im Rahmen der Zwei-Dimensionalen Semantik hilft diese Unterscheidung, die Überarbeitung von Überzeugungen zu strukturieren, indem sowohl die Welt als auch der Kontext (d.h. der Zeitpunkt der Wachsamkeit) berücksichtigt werden müssen. Der Unterschied zwischen den K- und S-Intensionen wird hier besonders deutlich: Beauty hat eine K-Intension über die Welt (z.B. der Tag des Experiments), aber ihre S-Intension hängt von ihrem aktuellen Zustand und dem Kontext ab.

Ein weiteres Beispiel, das zur Veranschaulichung verwendet wird, ist der „Messy Shopper“ . In diesem Fall wird ein Käufer gezeigt, der durch ein chaotisches Geschäft navigiert und unsicher ist, welche Entscheidungen er aufgrund der Vielzahl von Optionen treffen soll. Der Käufer befindet sich in einer unsicheren Situation, da er nicht nur unsicher darüber ist, welches Produkt er kaufen soll, sondern auch unsicher über die Art und Weise, wie er die Optionen interpretieren soll. Diese Unsicherheit betrifft sowohl die Welt als auch den Kontext. Der Käufer weiß nicht nur nicht, welches Produkt das beste für ihn ist (Unsicherheit über die Welt), sondern er ist sich auch unsicher darüber, wie er den Kontext (die verschiedenen Kategorien von Produkten oder die besten Angebote) richtig einordnen soll.

Die Zwei-Dimensionale Perspektive hilft, diese Unsicherheit sowohl in Bezug auf die Welt als auch den Kontext zu modellieren und zu überarbeiten. Die K-Intension des Käufers bezieht sich auf seine Überzeugungen über die Welt der Optionen, die er kaufen kann. Die S-Intension bezieht sich auf den spezifischen Kontext, in dem der Käufer sich befindet, und wie dieser Kontext seine Entscheidungen beeinflusst. Die Zwei-Dimensionale Semantik ermöglicht es, diese beiden Unsicherheiten miteinander zu verbinden und den Entscheidungsprozess des Käufers realistischer zu modellieren, indem sowohl die objektive Welt (die Produkte) als auch der subjektive Kontext (der Kaufprozess) berücksichtigt werden.

Beide Beispiele – das „Sleeping Beauty“ -Paradox und der „Messy Shopper“ – verdeutlichen die Bedeutung von Kontextsensitivität und der Notwendigkeit, Überzeugungen in Bezug auf beide Dimensionen zu überarbeiten: die Welt und den Kontext. Die Zwei-Dimensionale Semantik bietet einen strukturierten Rahmen, um zu verstehen, wie Überzeugungen sowohl über die Welt als auch über den Kontext hinweg verändert werden können. Dies ist besonders relevant für Belief Revision, da diese Beispiele zeigen, wie die Unsicherheit über den Kontext genauso wichtig ist wie die Unsicherheit über die Welt. In vielen realen Situationen, wie sie in den Beispielen dargestellt sind, ist die Korrektur von Überzeugungen nicht nur eine Frage des Wissens über die Welt, sondern auch des Wissens über die eigene Position innerhalb dieser Welt.

Die Beispiele verdeutlichen auch die Flexibilität und die Herausforderungen der Zwei-Dimensionalen Semantik: Während sie eine präzisere Modellierung von Überzeugungen und ihrer Revision ermöglicht, bringt sie auch Komplexität mit sich, die bei der Modellierung von Unsicherheit und Kontextabhängigkeit berücksichtigt werden muss.

Am Ende der Dissertation wird die Notwendigkeit einer Zwei-Dimensionalen Perspektive auf die Überarbeitung von Überzeugungen zusammengefasst. Die Unterscheidung zwischen K- und S-Intensionen ist entscheidend, um kontextabhängige Unsicherheiten angemessen zu adressieren und eine präzisere, realistischere Methode zur Überarbeitung von Überzeugungen zu entwickeln. Die Beispiele von John, der „Sleeping Beauty" und dem „Messy Shopper" zeigen, dass die Zwei-Dimensionale Semantik einen umfassenderen Rahmen bietet, um mit den verschiedenen Arten von Unsicherheit und Kontextsensitivität in der Überarbeitung von Überzeugungen umzugehen.

Ein zentrales Ergebnis dieser Arbeit ist, dass die Zwei-Dimensionale Semantik in bestimmten Fällen unerlässlich ist, insbesondere dann, wenn kontextabhängige Überzeugungen eine Rolle spielen. In diesen Fällen ist es nicht ausreichend, nur die Welt zu betrachten, sondern auch die Position des Agenten in der Welt, die seine Überzeugungen beeinflusst. In den untersuchten Beispielen wird deutlich, wie die Zwei-Dimensionale Semantik hilft, sowohl die Unterscheidung zwischen der Welt und dem Kontext als auch die Wechselwirkungen zwischen den beiden Dimensionen zu berücksichtigen. So wird die Überarbeitung von Überzeugungen in komplexen, dynamischen Szenarien realistischer modelliert, was die Genauigkeit und Nützlichkeit von Belief Revision Modellen erheblich steigert.

Gleichzeitig wurde jedoch auch festgestellt, dass in vielen Fällen eine einfachere, eindimensionale Revision ausreicht, solange der Kontext keine zentrale Rolle spielt oder die Unsicherheit nur in Bezug auf die Welt besteht. Für diese weniger komplexen Szenarien bieten die klassischen Belief Revision Modelle weiterhin eine effektive Lösung. Das bedeutet, dass die Wahl zwischen Zwei-Dimensionaler und eindimensionaler Semantik nicht nur von der Komplexität des jeweiligen Problems abhängt, sondern auch von der Bedeutung des Kontextes für die Überzeugungen des Agenten. Diese Erkenntnis trägt dazu bei, die Anwendbarkeit der Zwei-Dimensionalen Semantik zu erweitern, ohne ihre Komplexität unnötig zu steigern, wenn sie nicht erforderlich ist.

Die Erkenntnisse dieser Arbeit bieten einen wertvollen Beitrag zur Diskussion über die Dynamik von Überzeugungen und ihren Veränderungen und eröffnen neue Perspektiven für die Anwendung von Zwei-Dimensionaler Semantik in verschiedenen Disziplinen, einschließlich der Kognitionswissenschaft, der Künstlichen Intelligenz und der Philosophie. Insbesondere könnte diese Arbeit dazu beitragen, die Modellierung von Entscheidungsprozessen in Bereichen wie maschinelles Lernen und autonome Systeme zu verbessern, wo Kontextsensitivität und Unsicherheit eine wesentliche Rolle spielen. Auch in der philosophischen Forschung zu Wissen und Wahrheit, insbesondere in der Epistemologie, könnte der entwickelte Rahmen neue Impulse geben.

Die nächsten Schritte in dieser Forschung könnten darin bestehen, die Zwei-Dimensionale Semantik weiter zu verfeinern und auf komplexere Situationen anzuwenden, wie z. B. in der interaktiven Entscheidungsfindung oder in komplexen maschinellen Lernmodellen. Insbesondere die Modellierung von Kontexten in dynamischen, interaktiven Umgebungen könnte von Bedeutung sein, da hier häufig multiple, sich ändernde Perspektiven und unsichere Informationen verarbeitet werden müssen. Der theoretische Rahmen, der in dieser Dissertation entwickelt wurde, bietet eine solide Grundlage für zukünftige Arbeiten in diesen Bereichen und eröffnet neue Forschungsfragen, die eine weitere Vertiefung und Erweiterung der Zwei-Dimensionalen Semantik erfordern. Zukünftige Studien könnten untersuchen, wie die Zwei-Dimensionale Semantik auf neue Problemstellungen angewendet werden kann, beispielsweise bei der Automatisierung von Entscheidungsprozessen, der Optimierung von KI-Modellen oder in der Analyse von Interaktionen zwischen Agenten in komplexen Systemen.

Acknowledgments

The first time I realized the need to combine philosophy with analytical tools was when I encountered the Hypothetico-Deductive Model, the foundation of the scientific method. I wish I had discovered it sooner. Over time, my philosophical interests evolved. With so many questions racing through my mind, I came to understand that I may never find answers to all of them, but I could focus on one at a time. And perhaps, through dedication and hard work, I might find answers to just a few. Throughout this journey, I have been fortunate to receive support from so many incredible people. How could I ever fully express my gratitude for the daily exchange of words, the shared smiles, or the distant but heartfelt encouragement? I hope this thesis honors the collective efforts that have helped me reach this point.

I would also like to thank my advisor, Prof. DDr. Hannes Leitgeb, for his unwavering enthusiasm and for having the remarkable ability to explain what I was missing or misunderstanding. His guidance has been a genuine source of motivation, continually pushing me to learn more. Being part of the MCMP has been an extraordinary experience, and I am deeply grateful for the opportunity to be here.

To my friends and colleagues, thank you for the countless discussions and revisions: Abílio Rodrigues, Peter Verdée, Evelyn Erickson, Bruno Jacinto, Matteo De Benedetto, Sena Bozdag, Pedro Lemos, Dante Pinto, Sebastian Hengst, Sarah Hiller, Benedict Eastaugh, Norbert Gratzl, Alessandra Marra, Marianna Antonutti, Pedro Afonso Branco . . . and so many others! Your valuable insights have been instrumental in shaping this work. I also want to extend my heartfelt thanks to Ursula Danninger, who was always there to assist whenever I needed help.

To my core support system: Rodrigo and Alice, my daughter, thank you. I dedicate this work to you both, as well as to Nico, my son, who was born during the course of this project. To all my family, in particular to my mother, Heloisa, and to my father, Henrique, thank you: this work wouldn't be possible without your encouragement.

I couldn't have done it without all of you! Thank you so much!

Chapter 1

Introduction

Imagine waking up alone in a vast, untamed wilderness. The surroundings are unfamiliar dense forests, distant mountains, and unmarked trails. You have no map, no clock, and no companions. Clutched in your hand is a single letter. It reads: *"You must survive alone until rescue arrives."* But the letter offers no certainty no promise of when help will come or even if it will come at all. You don't know what day it is, where you are, or how long you've been stranded. The only thing you can rely on is your ability to observe, reason, and make decisions.

In this scenario, survival depends on your capacity to orient yourself and adapt to your surroundings. You begin to reason and use language to make sense of your environment. You might utter sentences like, *"Today is windy,"* or *"I didn't eat yesterday,"* even though you have no idea what day today or yesterday corresponds to on a calendar. Over time, you start naming key locations to help you navigate: *Lily's Fish Pond*, where you catch fish; *Ox-Foxy Hill*, a safe place to hunt; and *Blueberry-Musk Camp*, where you gather berries. These names are deeply personal they derive their meaning only from your context and serve as tools to organize your experience. However, as the days stretch on, you begin to lose track of time. You know that the rescue team will search for you for exactly 100 days. After that, on Day 101, you must abandon hope. Yet, without a reliable calendar, how can you determine whether today is Day 99 or Day 103? In the absence of concrete temporal markers, your reasoning relies on subjective perceptions and context-sensitive expressions.

This wilderness narrative vividly illustrates the philosophical problem at the heart of this thesis: the interplay between context-sensitive expressions and rational reasoning. As Kaplan explained, *"The referent of a pure indexical depends on the context, and the referent of a demonstrative depends on the associated demonstration"* [46, p. 491]. These expressions such as *"I," "here,"* and *"now"* are not merely linguistic conveniences; they are fundamental tools for making sense of the world in uncertain and dynamic environments. In this scenario, the utterance *"Today is windy"* depends entirely on the context in which it is expressed. The meaning of *"today"* shifts with each passing day, yet the statement remains coherent within its specific context. Indexicals like *"I"* and *"here"* function similarly, allowing you to express beliefs and navigate your surroundings despite the absence of fixed references. As Perry noted, *"It is not only that language is sensitive to the user and his context; the thoughts we have, that our utterances express, are similarly sensitive"* [73, p. 3].

These expressions illustrate the complex relationship between language, context, and belief. Unlike proper names or fixed propositions, Indexicals depend on parameters like the speaker's identity, location, and temporal position. This reliance on context raises critical questions about how beliefs involving Indexicals are formed, interpreted, and revised. For example, how can others understand your statement *"Here is Lily's Fish Pond"* when they lack the contextual information that gives the name its meaning? Such examples highlight the inherent fragility of context-dependent communication and reasoning. This fragility extends beyond individual misunderstandings it fundamentally shapes how we acquire, structure, and adjust our beliefs. Misinterpretations or incomplete contextual information can lead to flawed reasoning processes, and these flaws are often magnified when reasoning collaboratively or across different contexts. Addressing these challenges requires us to examine how context-sensitive expressions function within reasoning systems and to explore the mechanisms by which individuals can rationally revise beliefs when faced with context-related ambiguities. This brings us to the broader questions guiding this thesis: how do context-sensitive expressions influence our reasoning and Belief Revision processes, and

how can one rationally revise their beliefs under indexical information?

When reasoning about our beliefs, we encounter challenges that stem not only from the complexity of our reasoning processes but also from the intricate nature of language. Human reasoning inherently involves dynamic changes in our set of beliefs as we process new information, resolve contradictions, and communicate our thoughts with others. A particularly challenging aspect of this process arises from language's reliance on context-sensitive expressions terms like "today," "I," "here," and many more. These expressions are deeply embedded in the fabric of our reasoning, yet their meaning depends heavily on the context in which they are used.

For reasoning to be successful, we must understand the contextual features these expressions rely on, such as the time of utterance, the individuals involved, and the specific location being referenced. However, our access to these contextual elements is not always precise. Misunderstandings, ambiguity, or uncertainty about the relevant contextual features can lead us astray, resulting in reasoning that produces false conclusions. These errors are not merely theoretical concerns; they are practical problems we face in everyday reasoning and communication, particularly when we need to revise our beliefs in light of conflicting or new information. The central question, then, is not merely about the nature of context-sensitivity as a linguistic phenomenon. Instead, it is about how individuals can navigate these challenges rationally. Specifically, how is it possible for someone to revise their beliefs and reasoning processes when they encounter issues related to context-sensitivity? What does it mean to change one's mind in a rational manner when the underlying problem stems from uncertainties or ambiguities in contextual understanding?

To investigate these questions thoroughly, it is crucial to identify and analyze the key concepts that underlie this problem. Understanding the interplay between language, context, and reasoning requires a clear conceptual framework that can account for how context-sensitive expressions influence Belief Revision and reasoning processes. It also demands tools capable of capturing the dynamic and fluid nature of these processes. This leads us to the role of formal systems as tools for contextual reasoning, the focus of the next section. Formal systems offer a structured way to represent, analyze, and even simulate reasoning processes involving context-sensitive expressions. They allow us to model how contextual features interact with beliefs and how Belief Revisions can be carried out systematically. By bridging the gap between intuitive reasoning and formal representation, we can better understand the mechanisms that enable rational Belief Revision in the face of context-sensitivity issues. In the following section, we will explore how formal systems provide the necessary tools for tackling the challenges posed by context-dependent reasoning.

The challenges posed by context-sensitive expressions, such as *"I," "here,"* and *"now,"* call for a rigorous framework to analyze how beliefs are formed, expressed, and revised. Logic, as the systematic study of reasoning, offers powerful tools to model and evaluate these processes. By formalizing the relationships between propositions, individuals, and contexts, logic provides the precision needed to address complex phenomena like indexicality. For instance, understanding how the statement *"Today is windy"* can convey different meanings depending on the speaker and time requires tools that go beyond everyday reasoning. These tools not only clarify abstract philosophical questions but also allow us to test hypotheses about reasoning in structured, formal environments. This chapter introduces key formal systems in logic, demonstrating their strengths and limitations in addressing context-sensitive reasoning and Belief Revision.

1.1 Formal Systems as Tools for Reasoning

Logical systems provide formal frameworks to analyze and reflect upon the nature of reasoning and epistemic questions. By defining precise rules and structures, they enable the exploration of how beliefs are formed, justified, and revised in light of new information. Formal logics also help identify valid reasoning patterns, avoid fallacies, and model the epistemic states of agents, making them indispensable tools for addressing fundamental philosophical and practical problems. Thus, we now turn our attention to the foundations of the formal systems that will be presented along this work. Let us start with Classical Logics as they provide the foundation for most traditional logical systems. Classical Logics are formal systems that adhere strictly to the three classical principles that ensure consistency and provide the framework for reasoning in classical logic. Intuitively, we can define those principles as follows:

Principle of Non-Contradiction (PNC): A given proposition cannot be both true and false at the same time.

Principle of Explosion (Ex Falso Quodlibet): From a contradiction any proposition can be inferred.

Principle of the Excluded Middle (PEM): Either a proposition is true, or its negation is true.

These principles are fundamental to classical formal systems such as Propositional Logic and First-Order Logic and will be presented formally in the next sections. Systems that reject or modify these principles, such as paraconsistent or intuitionistic logics, provide alternative frameworks for reasoning while challenging the classical assumptions.

1.1.1 Propositional Logic

Building on the classical principles, Propositional Logic, also known as Sentential Logic, serves as a foundational system in formal reasoning. It provides a systematic framework for analyzing declarative sentences, known as **propositions**, based on their truth values and logical relationships. Each proposition is assigned a truth value: either true (T) or false (F) [11, p. 3]. Propositional Logic captures these classical principles through the use of logical connectives, which will be formally introduced and explored in detail. Logical connectives such as conjunction ($\wedge$), disjunction ($\vee$), negation ($\neg$), implication ($\rightarrow$), and biconditional ($\leftrightarrow$) allow for the combination of propositions to form more complex statements, providing a precise method for analyzing logical relationships and reasoning patterns.

Propositional Logic, also known as Sentential Logic, is a foundational system in formal reasoning. It provides a systematic framework for analyzing declarative sentences, known as **propositions**, based on their truth values and logical relationships. Each proposition is assigned a truth value: either true (T) or false (F) [11, p. 3]. Propositional Logic will consider logical connectives such as:

- $\wedge$ (*and*),
- $\vee$ (*or*),
- $\neg$ (*not*),
- $\rightarrow$ (*implies*),
- $\leftrightarrow$ (*if and only if*),

that allow the construction of compound statements whose truth values depend systematically on the truth values of their components [20, p. 12].

Propositional Logic is widely used for analyzing arguments, detecting contradictions, and ensuring logical consistency. Its principles underpin more advanced logical systems and serve as a benchmark for evaluating the soundness and completeness of reasoning frameworks.

Syntax of Propositional Logic The syntax of Propositional Logic defines the formal rules for constructing valid expressions, called **well-formed formulas (WFFs)**. WFFs are defined recursively as follows:

1. **Base Case:** Every propositional variable $(P, Q, R, \ldots)$ is a WFF.
2. **Recursive Rules:** If ϕ and ψ are WFFs, then the following are also WFFs:
   - $\neg\phi$,
   - $\phi \wedge \psi$,
   - $\phi \vee \psi$,
   - $\phi \rightarrow \psi$,
   - $\phi \leftrightarrow \psi$.
3. **Closure Rule:** Nothing else is a WFF.

This definition ensures that every WFF is syntactically valid. For example, the formula $P \wedge (Q \vee \neg R)$ is a WFF, while $PQ\wedge$ is not.

Semantics of Propositional Logic The semantics of Propositional Logic assign truth values to WFFs. Propositional variables are interpreted as either true (T) or false (F), and the truth values of compound formulas are determined systematically using truth tables. For example, the semantics of the conjunction ($\wedge$) are defined as follows:

Truth Tables for Main Logical Connectives						
P	**Q**	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

Table 1.1: Truth tables for main logical connectives: Negation ($\neg$), Conjunction ($\wedge$), Disjunction ($\vee$), Implication ($\rightarrow$), and Biconditional ($\leftrightarrow$).

A formula in Propositional Logic is called a **tautology** if it is true under every possible valuation of its propositional variables. For example, the formula $P \vee \neg P$ is always true regardless of the truth value of P. A **contradiction**, on the other hand, is a formula that is false under every valuation, such as $P \wedge \neg P$. Lastly, a formula is classified as a **contingency** if it is neither a tautology nor a contradiction, meaning its truth value depends on the valuation of its variables. An example of a contingency is $P \wedge Q$, which is true in some cases and false in others depending on the truth values of P and Q. These classifications help in understanding the logical properties and behavior of formulas within a given logical system.

These semantic rules provide a robust framework for evaluating logical validity, detecting inconsistencies, and ensuring coherence in reasoning. Given two propositional variables P and Q, we can now define formally the three Classical Principles presented before as follows:

Principle of Non-Contradiction (PNC): $\vdash \neg(P \wedge \neg P)$. This ensures that contradictions are not allowed in Classical Logics, providing a basis for consistent reasoning. Semantically, $\vdash \neg(P \wedge \neg P)$ means that $\neg(P \wedge \neg P)$ is logically true. Deductively, it means that $\neg(P \wedge \neg P)$ is provable without assumptions.

Principle of Explosion (Ex Falso Quodlibet): $P \wedge \neg P \vdash Q$. This highlights the importance of avoiding contradictions, as their presence would render the system trivial by allowing any statement to be derived. Semantically, $P \wedge \neg P \vdash Q$ means that $P \wedge \neg P$ is not true in any interpretation. Deductively, it means that the arbitrary formula Q is derivable from $P \wedge \neg P$.

Principle of the Excluded Middle (PEM): This principle asserts that for any proposition, either the proposition is true, or its negation is true. Formally, $\vdash P \vee \neg P$. Semantically, $\vdash P \vee \neg P$ means that $P \vee \neg P$ is logically true, and hence in every interpretation P is true or $\neg P$ is true. Deductively, it means that $P \vee \neg P$ is provable without assumptions.

Moreover, Propositional Logic is governed by several fundamental principles that ensure its utility in formal reasoning:

- **Consistency:** A set of propositions is consistent if no contradiction can be derived from it. For example, the set $\{P, \neg P\}$ is inconsistent because both P and $\neg P$ cannot be true simultaneously.
- **Monotonicity:** If a conclusion can be derived from a set of propositions, it remains derivable when additional propositions are added to the set. For example, if $P \rightarrow Q$ and P imply Q, adding R to the set does not affect the validity of the inference.
- **Soundness:** A logical system is sound if every formula derivable in the system is logically true under its semantics.
- **Completeness:** A logical system is complete if every formula that is logically true under its semantics can be derived within the system.

These principles provide a foundation for assessing the reliability and expressiveness of Propositional Logic. For instance, the combination of soundness and completeness ensures that Propositional Logic is both reliable and sufficient for analyzing truth-functional relationships. Consider the statement:

"If it is windy, then I will stay near the camp."

This can be represented in Propositional Logic as:

$$P \rightarrow Q,$$

where P represents *"It is windy"* and Q represents *"I will stay near the camp."* The truth value of $P \rightarrow Q$ depends on the truth values of P and Q, as shown in the truth table for implication. Additionally, Propositional Logic can detect inconsistencies in reasoning. For example, the set $\{P, \neg P\}$ is inconsistent because it violates the Principle of Non-Contradiction.

While Propositional Logic is powerful for analyzing truth-functional relationships, it abstracts away from semantic content and context. For example, the sentence *"It is windy today."* cannot be fully analyzed without considering the temporal context given by the expression *"today"*) or the speaker's perspective as Propositional Logic needs more Expressive Power. These limitations motivate the need for more expressive frameworks, such as First-Order Logic and Modal Logic, which can account for relationships between objects and dynamic contexts.

1.1.2 First-Order Logic: Extending the Expressive Power

First-Order Logic (FOL) extends Propositional Logic by introducing variables, predicates, and quantifiers, allowing for the formalization of statements about objects, their properties, and relationships. This extension significantly enhances the expressive power of formal systems, enabling reasoning about more complex scenarios beyond the capabilities of Propositional Logic [11, p. 19]. FOL is particularly suited for analyzing structured domains where objects and relationships play a central role.

Syntax of First-Order Logic The syntax of FOL consists of the following components:

1. **Variables** $(x, y, z, \ldots)$: Represent objects in the domain.
2. **Constants** $(a, b, c, \ldots)$: Represent specific objects in the domain.
3. **Predicates** ($P(x), Q(x, y), \ldots$): Represent properties of objects or relationships between objects. For example, $F(x)$ could represent *" x is a fish."*
4. **Logical Connectives** $(\wedge, \vee, \neg, \rightarrow, \leftrightarrow)$: Combine formulas.
5. **Quantifiers:**
   - $\forall$ (*for all*): Denotes universal quantification.
   - $\exists$ (*there exists*): Denotes existential quantification.

The rules for constructing well-formed formulas (WFFs) in FOL are defined recursively:

1. **Base Case:** If P is an n-place predicate and $t_1, t_2, \ldots, t_n$ are terms (constants or variables), then $P(t_1, t_2, \ldots, t_n)$ is a WFF.
2. **Recursive Rules:** If ϕ and ψ are WFFs, then the following are also WFFs:
   - $\neg\phi$,
   - $\phi \wedge \psi$,
   - $\phi \vee \psi$,
   - $\phi \rightarrow \psi$,
   - $\phi \leftrightarrow \psi$,
   - $\forall x \phi$,

- $\exists x \phi$.

3. **Closure Rule:** Nothing else is a WFF.

For example:

$$\forall x(F(x) \rightarrow E(x)),$$

is a WFF representing the statement, *"For all x, if x is a fish, then x is edible."*

Semantics of First-Order Logic The semantics of First-Order Logic (FOL) provide a formal framework for interpreting the symbols used in logical expressions. The meaning of terms, formulas, and predicates is determined with respect to a **domain of discourse** and an **interpretation function**, which collectively define the truth conditions of logical statements. The key components of the semantics are as follows:

- **Domain of Discourse ($\mathcal{D}$):** The non-empty set of all objects under consideration. This set serves as the universe in which variables and terms take their values.
- **Interpretation Function ($\mathcal{I}$):** An interpretation function assigns meaning to the non-logical symbols of the language, including constants, predicates, and function symbols. Formally, an interpretation consists of:
  - An assignment of each constant symbol c to a specific element $\mathcal{I}(c) \in \mathcal{D}$. For example, if the domain consists of marine animals, a constant a might refer to a specific fish.
  - A function $\mathcal{I}(P)$ that assigns an n-ary predicate symbol P to a subset of $\mathcal{D}^n$, representing the set of all n-tuples for which the predicate holds. For instance, if $F(x)$ represents "x is a fish," then $\mathcal{I}(F) \subseteq \mathcal{D}$ would be the subset of objects in the domain that are fish.
  - An assignment of function symbols to actual functions on $\mathcal{D}$. If f is a function symbol of arity n, then $\mathcal{I}(f)$ is a function from $\mathcal{D}^n$ to $\mathcal{D}$. For example, if $f(x)$ represents "the mother of x," then $\mathcal{I}(f)$ maps each element in $\mathcal{D}$ to its corresponding mother.
- **Variable Assignment (σ):** A function that maps each variable to an element of the domain. Since variables do not have fixed interpretations, a variable assignment function $\sigma : V \rightarrow \mathcal{D}$ assigns each variable a value from the domain.
- **Truth Conditions for Atomic Formulas:** Given an interpretation $\mathcal{I}$ and a variable assignment σ, the truth value of an atomic formula is determined as follows:

– If t is a term (either a variable x or a constant c), then its denotation is:

$$\llbracket t \rrbracket^{\mathcal{I},\sigma} = \begin{cases} \sigma(x) & \text{if } t \text{ is a variable } x, \\ \mathcal{I}(c) & \text{if } t \text{ is a constant } c. \end{cases}$$

– If $P(t_1, \ldots, t_n)$ is an atomic formula where P is an n-ary predicate and $t_1, \ldots, t_n$ are terms, then:

$$\mathcal{I}, \sigma \models P(t_1, \ldots, t_n) \text{ if and only if } (\llbracket t_1 \rrbracket^{\mathcal{I},\sigma}, \ldots, \llbracket t_n \rrbracket^{\mathcal{I},\sigma}) \in \mathcal{I}(P).$$

- **Truth Conditions for Logical Connectives:** The truth values of compound formulas are defined recursively based on their structure:
  – $\mathcal{I}, \sigma \models \neg\varphi$ if and only if $\mathcal{I}, \sigma \not\models \varphi$.
  – $\mathcal{I}, \sigma \models (\varphi \vee \psi)$ if and only if $\mathcal{I}, \sigma \models \varphi$ or $\mathcal{I}, \sigma \models \psi$.
  – $\mathcal{I}, \sigma \models (\varphi \wedge \psi)$ if and only if $\mathcal{I}, \sigma \models \varphi$ and $\mathcal{I}, \sigma \models \psi$.
  – $\mathcal{I}, \sigma \models (\varphi \rightarrow \psi)$ if and only if $\mathcal{I}, \sigma \not\models \varphi$ or $\mathcal{I}, \sigma \models \psi$.
- **Truth Conditions for Quantifiers:**
  – $\mathcal{I}, \sigma \models \forall x\, \varphi(x)$ if and only if for every $d \in \mathcal{D}$, we have $\mathcal{I}, \sigma[x \mapsto d] \models \varphi(x)$, where $\sigma[x \mapsto d]$ is the modified variable assignment that maps x to d.
  – $\mathcal{I}, \sigma \models \exists x\, \varphi(x)$ if and only if there exists some $d \in \mathcal{D}$ such that $\mathcal{I}, \sigma[x \mapsto d] \models \varphi(x)$.
- **Satisfaction and Validity:** A formula φ is satisfied in an interpretation $\mathcal{I}$ with a variable assignment σ, denoted as $\mathcal{I}, \sigma \models \varphi$, if it evaluates to true under those conditions. A formula is **valid** (denoted $\models \varphi$) if it is true under every interpretation.

Thus, the semantics of First-Order Logic establish a rigorous connection between the syntactic expressions of the language and their meanings in a formalized model, allowing for precise reasoning about objects and their relationships.

First-Order Logic (FOL) is governed by key logical principles that ensure its consistency, reliability, and applicability in formal reasoning. Moreover, FOL inherits several principles from Propositional Logic while introducing new considerations:

- **Consistency:** A set of formulas is *consistent* if it does not lead to a contradiction. That is, there exists at least one interpretation where all formulas in the set hold true. Conversely, a set is *inconsistent* if it entails a contradiction. For example, consider the set:

$$\{\forall x(F(x) \to E(x)), \neg E(a)\}.$$

This set becomes inconsistent if $F(a)$ is true, since it would force $E(a)$ to be true while simultaneously asserting $\neg E(a)$.

- **Monotonicity:** The inference rules of FOL exhibit *monotonicity*, meaning that if a conclusion follows from a given set of premises, it remains valid even when additional premises are introduced. Formally, if:

  $$\{F(x)\} \models E(x),$$

  then adding another formula to the premise set does not invalidate the conclusion:

  $$\{F(x), G(x)\} \models E(x).$$

  This property ensures that new knowledge does not invalidate previous inferences in classical FOL.

- **Soundness:** A logical system is *sound* if every formula derivable within the system is semantically true in all models. In other words, if φ is provable in FOL, then φ is also true under every interpretation:

  $$\vdash \varphi \quad \Rightarrow \quad \models \varphi.$$

  This guarantees that no false conclusions can be derived from valid premises.

- **Completeness:** FOL is *complete* if every formula that is true in all models is also derivable within the system. That is, if φ is true under every interpretation, then it has a formal proof within FOL:

  $$\models \varphi \quad \Rightarrow \quad \vdash \varphi.$$

  The completeness theorem, established by Gödel, ensures that FOL is powerful enough to capture all truths expressible in its formalism.

These properties collectively ensure the robustness of First-Order Logic, making it a reliable framework for formal reasoning and automated deduction. Consider the following example:

"Every place where I find food also has water."

This can be formally represented in First-Order Logic as:

$$\forall x(F(x) \rightarrow W(x)),$$

where:

- $F(x)$ represents *"x provides food"*, and
- $W(x)$ represents *"x provides water"*.

Such formalizations enable precise reasoning about structured domains, ensuring that logical relationships between concepts are explicitly defined and rigorously analyzed. Despite its expressive power, FOL assumes fixed interpretations for predicates and constants, making it inadequate for handling context-sensitive expressions such as *"I find water here now."* The sentences involving Indexicals such as *"I," "here,"* and *"now"* (whose meanings depend on the context of utterance) cannot be fully capture by the expressive power of FOL. Addressing these challenges requires extending FOL with additional frameworks, such as Modal Logic and Two-Dimensional Semantics [20, p. 35]. Thus, the limitations of FOL in handling context-sensitivity lead us to explore Modal Logic, which introduces the concepts of possible worlds and Accessibility Relations. Modal Logic provides a more dynamic framework for reasoning about necessity, possibility, and context-sensitive phenomena.

1.1.3 Modal Logic: Extending Classical Frameworks

Modal Logic extends both Propositional Logic and First-Order Logic by introducing modal operators that allow reasoning beyond truth values in a single, static domain. While Propositional Logic and FOL are concerned whether a statement is true or false within a fixed interpretation, Modal Logic expands this framework to incorporate notions of Necessity, Possibility, and more modal operators. These extensions make Modal Logic particularly valuable for analyzing dynamic contexts, including reasoning processes that involve context-sensitive expressions and epistemic variability [41, p. 1].

By formalizing concepts such as Possibility ($\Diamond$) and Necessity ($\Box$), Modal Logic provides a powerful framework for representing truth as something that may depend on different possible worlds rather than a single, fixed reality. This perspective enables a more nuanced understanding of indexicality and context-dependency, as the truth of statements may shift depending on the world, time, or agent in which they are evaluated. As such, Modal Logic plays a foundational role in formal semantics, philosophy of language, and theories of knowledge, allowing for precise formalization of concepts like belief, knowledge, obligation, and temporal change.

Syntax of Modal Logic The syntax of Modal Logic builds on Propositional Logic, introducing additional modal operators that allow for reasoning about necessity and possibility. The fundamental components of Modal Logic include:

1. **Propositional Variables** $(P, Q, R, \dots)$: Represent atomic propositions that can take the truth values true or false in different possible worlds.
2. **Logical Connectives** $(\wedge, \vee, \neg, \rightarrow, \leftrightarrow)$: Standard logical operators used to combine propositional formulas.
3. **Modal Operators:**
   - $\Box$ (*necessarily*): Indicates that a proposition holds in all accessible possible worlds.
   - $\Diamond$ (*possibly*): Indicates that a proposition holds in at least one accessible possible world.

The formation of well-formed formulas (WFFs) in Modal Logic follows specific syntactic rules:

1. **Base Case:** Every propositional variable $(P, Q, \dots)$ is a WFF.
2. **Recursive Rules:** If ϕ and ψ are WFFs, then the following are also WFFs:
   - $\neg\phi$ (negation),
   - $\phi \wedge \psi$ (conjunction),
   - $\phi \vee \psi$ (disjunction),
   - $\phi \rightarrow \psi$ (implication),
   - $\phi \leftrightarrow \psi$ (biconditional),
   - $\Box\phi$ (necessity),
   - $\Diamond\phi$ (possibility).
3. **Closure Rule:** Nothing else is a WFF.

For example, the formula:

$$\Box P \rightarrow \Diamond Q$$

expresses the statement: *"If P is necessarily true, then Q is possibly true."* This showcases how modal operators interact within logical expressions.

Semantics of Modal Logic The semantics of Modal Logic are defined using the notion of *possible worlds*, where each world represents a distinct way the world might be. The interpretation of modal formulas depends on a formal structure known as a **Kripke frame** or **Kripke model**. The key components are:

- **Set of Possible Worlds** (W): A non-empty set of worlds, denoted as $W = \{w_1, w_2, w_3, \dots\}$. These worlds represent different scenarios in which formulas may hold.
- **Accessibility Relation** (R): A binary relation $R \subseteq W \times W$ that determines which worlds are accessible from others. If $w_i R w_j$, it means that world w_j is accessible from w_i. The Accessibility Relation can exhibit the following properties:
  - **Reflexivity:** $\forall w \in W, wRw$ Every world is accessible from itself. This means that if something is necessary, it must be true in the current world. Reflexivity characterizes system T.
  - **Symmetry:** $\forall w, v \in W,$ if wRv then vRw If a world v is accessible from w, then w is also accessible from v. This implies that if something is possible from one perspective, then its converse is also possible. Symmetry is a key feature of system B.
  - **Transitivity:** $\forall w, v, u \in W,$ if wRv and vRu, then wRu If a world v is accessible from w, and another world u is accessible from v, then u is accessible from w. This property ensures that if something is necessarily necessary, then it is necessary. It characterizes system $S4$.
  - **Euclidean Property:** $\forall w, v, u \in W,$ if wRv and wRu, then vRu If two worlds v and u are accessible from the same world w, then they are accessible to each other. This property ensures that if something is possibly true, then it is possibly necessary. It defines system $S5$.
  - **Seriality:** $\forall w \in W, \exists v \in W$ such that wRv Every world has at least one accessible world. This property ensures that $\Diamond\phi$ is always satisfied for at least one world. Seriality is a characteristic feature of deontic Modal Logic (D).
- **Interpretation Function** ($\mathcal{V}$): Assigns truth values to propositional variables in each world. Formally, $\mathcal{V} : W \times \{P, Q, R, \dots\} \to \{\text{true}, \text{false}\}$, meaning that each atomic proposition is evaluated in different possible worlds.

Truth Conditions for Modal Operators Given a Kripke model $\mathcal{M} = (W, R, \mathcal{V})$, the satisfaction of modal formulas at a world w (denoted as $\mathcal{M}, w \models \phi$) is defined recursively:

- $\mathcal{M}, w \models P$ if and only if P is true in w, according to the interpretation function $\mathcal{V}$.
- $\mathcal{M}, w \models \neg\phi$ if and only if $\mathcal{M}, w \not\models \phi$.
- $\mathcal{M}, w \models \phi \wedge \psi$ if and only if $\mathcal{M}, w \models \phi$ and $\mathcal{M}, w \models \psi$.
- $\mathcal{M}, w \models \phi \vee \psi$ if and only if $\mathcal{M}, w \models \phi$ or $\mathcal{M}, w \models \psi$.
- $\mathcal{M}, w \models \phi \rightarrow \psi$ if and only if $\mathcal{M}, w \not\models \phi$ or $\mathcal{M}, w \models \psi$.
- $\mathcal{M}, w \models \Box\phi$ if and only if for all worlds w' such that wRw', we have $\mathcal{M}, w' \models \phi$.
- $\mathcal{M}, w \models \Diamond\phi$ if and only if there exists some world w' such that wRw' and $\mathcal{M}, w' \models \phi$.

Accessibility Relations and Modal Logic Systems The Accessibility Relation (R) plays a central role in defining the semantics of Modal Logic. It determines which possible worlds are considered accessible from a given world, shaping the interpretation of necessity ($\Box$) and possibility ($\Diamond$). Formally, R is a binary relation on the set of possible worlds, $R \subseteq W \times W$, where $w_i R w_j$ means that world w_j is accessible from world w_i. Different constraints on R give rise to various modal systems. The different modal systems arise from imposing particular constraints on the Accessibility Relation:

- **System** K: No restrictions on R. This is the most basic Modal Logic system, named after Saul Kripke.
- **System** T: Reflexive R. Ensures that if something is necessarily true, then it is true in the actual world.
- **System** $S4$: Reflexive and transitive R. Models knowledge and belief systems where accessibility follows transitive closure.
- **System** $S5$: An equivalence relation (reflexive, symmetric, and transitive). Assumes that all possible worlds are mutually accessible.
- **System** B: Reflexive and symmetric R. Captures Modal Logics where possibility and necessity mirror each other.

- **System** D: Serial R. Guarantees that some possible world is always accessible, useful for deontic logic.

By adjusting the properties of the Accessibility Relation, different Modal Logics can model necessity and possibility in various ways, making Modal Logic a flexible tool for analyzing knowledge, obligation, time, and other modal concepts.

Consider a Kripke model with three possible worlds: $W = \{w_1, w_2, w_3\}$, where w_1 has access to w_2 and w_3 ($w_1 R w_2$, $w_1 R w_3$), and let P be true in w_2 but false in w_1 and w_3. The formula $\Diamond P$ expressing *"It is possible that P is true,"* holds in w_1 because there exists an accessible world (w_2) where P is true. However, the formula $\Box P$ expressing *"P is necessarily true,"* does not hold in w_1 because P is false in w_3, which is also accessible from w_1.

Modal Logic provides a formal framework for reasoning about necessity and possibility across different scenarios, making it particularly useful in philosophy, computer science, and linguistics. Its formal syntax and semantics, defined through possible worlds and Accessibility Relations, enable precise modeling of knowledge, time, obligation, and other modal concepts. Moreover, Modal Logic adheres to principles that govern reasoning across possible worlds: Consistency, Monotonicity, Soundness, Completeness are to be understood analogously as in Propositional or First-Order Logic

These principles ensure that Modal Logic maintains logical rigor while providing the flexibility needed to analyze necessity and possibility. However, to formally capture these modal notions, we require a structured framework that allows truth values to be evaluated beyond a single, fixed reality. In that sense, the Possible Worlds Semantics provide a rigorous framework for understanding modality, counterfactual reasoning, and context-dependent truth. In Logic, it formalizes necessity and possibility, enabling precise treatments of modal, epistemic, and temporal logics. It also underpins the semantics of conditionals and Belief Revision, crucial for Artificial Intelligence and Formal Epistemology.

1.1.4 Possible Worlds: A Framework for Modal Reasoning

The concept of possible worlds provides a systematic approach to reasoning about necessity, possibility, and alternative scenarios. Rather than evaluating statements in isolation, Modal Logic considers how truth values may vary across different worlds, allowing for a formal representation of context-sensitive reasoning. Each possible world represents a complete and self-contained way the world could be, encompassing all facts, entities, and their relationships. This framework is essential for understanding a range of philosophical and logical issues, including context, Indexicals, and Belief Revision [58, p. 1].A possible world is a hypothetical scenario or state of affairs in which every proposition has a definite truth-value, either true or false. The actual world is simply one among an infinite set of possible worlds, often denoted as w_0. Other worlds $(w_1, w_2, \ldots)$ represent alternative ways reality might have been. By structuring Modal Logic around this idea, we gain a powerful tool for analyzing statements about what must be true, what could be true, and how truth varies across different circumstances.

Formally, a possible world can be understood as a model in which propositional variables $(P, Q, R, \ldots)$ are assigned truth values; and their relationships and properties are interpreted relative to a given possible world. For example, consider a proposition P: *"It is windy"*. In one world (w_1), the proposition P might be true, while in another world (w_2), P might be false. Therefore, the world and its accessibility relations must be considered when determining the truth-value of a modal statement. Suppose the sentence $Q \rightarrow R$ meaning *"If I gather food, then I survive"*. Intuitively, the formula $\Box(Q \rightarrow R)$ is true in a world w if in all accessible worlds w', gathering food ensures survival. Recall that Possible Worlds Semantics is governed by several key principles, including Consistency and Monotonicity. These principles ensure that reasoning about necessity and possibility remains coherent and logically consistent. Possible worlds are a versatile tool for analyzing indexical expressions and belief dynamics. For example, we could use Possible Worlds Semantics to deal with statements like *"I am here now"*, as they can be evaluated relative to a primary world of utterance and secondary worlds of evaluation. Or even think in terms of an agent's beliefs about the actual world and alternative possibilities can be formalized using Accessibility Relations. Consider the Wilderness example:

$$\Diamond(\text{\textit{“I find food”}}),$$

represents the possibility of finding food in at least one accessible world, while:

$$\Box(\text{\textit{“If I find food, I survive”}}),$$

indicates that survival is guaranteed in all accessible worlds if food is found. While Possible Worlds Semantics provides a powerful framework for reasoning about necessity and possibility, it does not inherently capture the nuances of context-sensitive expressions, such as temporal or personal perspectives. Extensions such as Two-Dimensional Semantics incorporate additional parameters, enabling the analysis of statements where meaning depends on both context and evaluation [46, p. 89]. This is crucial when we talk about belief change dynamics and context-sensitiveness.

Possible Worlds Semantics lays the groundwork for understanding more sophisticated frameworks like AGM Belief Revision and Logic of Demonstratives.

Indeed, Modal Logic and Possible Worlds Semantics seem *prima facie* particularly suited for reasoning about beliefs and contexts. For instance, the statement $\Box$ (*"If I gather food, then I will survive."*), can be formalized as $\Box(F \rightarrow S)$ where F represents *"I gather food"* and S represents *"I survive."* This formula indicates that in all accessible worlds, survival is guaranteed if food is gathered. Moreover, Modal Logic allows us to reason about the possibility of uncertain outcomes. For example:

$$\Diamond(\text{"I find water."}),$$

can represent the agent's belief that there exists at least one possible world where water is found.

Despite its expressive power, standard Modal Logic does not account for the nuances of Indexicals and context-sensitive expressions. These expressions depend on the identity of the speaker, the location, and the time of utterance parameters not inherently captured by Modal Logic's possible world semantics [20, p. 45]. The limitations of Modal Logic in addressing context-sensitive phenomena evoke the development of further extensions. The next section explores how these refinements integrate context and dynamism into logical systems, bridging the gap between formal reasoning and everyday applications.

1.1.5 Belief Revision: Core Concepts and Principles

Belief Revision is the process by which rational agents update their belief sets when confronted with new evidence or information. Rooted in epistemology and formal logic, this framework provides tools for maintaining consistency and coherence in an agent's beliefs, even as those beliefs evolve over time. The seminal AGM framework, developed by Alchourrón, Gärdenfors, and Makinson [2], offers a formal foundation for understanding Belief Revision. Let us take a look at some of the core concepts and principles surrounding AGM Standard Belief Revision.

A belief set is a collection of propositions that an agent accepts as true at a given time. This set must satisfy two key properties:

- **Consistency:** The belief set does not contain any contradictions.
- **Closure:** The belief set is closed under logical consequence; if P and $P \rightarrow Q$ are in the set, then Q must also be included.

Now, consider that a **set** is typically denoted by an uppercase letter, such as A, B, or C, and its elements are enclosed within curly brackets as follows: $A = \{a, b, c\}$ [9]. The fundamental relation in set theory is *membership*, denoted by the symbol $\in$. If an element x belongs to a set A, we write $x \in A$. Otherwise, we denote non-membership as $x \notin A$. Common operations on sets include:

- **Union**: $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$.
- **Intersection**: $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.
- **Difference**: $A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$.
- **Subset**: $A \subseteq B$ iff $\forall x(x \in A \rightarrow x \in B)$.

One of the primary difficulties with naïve set theory arises from unrestricted comprehension, the assumption that any well-defined property determines a set. This leads to paradoxes such as **Russell's paradox**, which considers the set: $R = \{x \mid x \notin x\}$. If $R \in R$, then by definition $R \notin R$, a contradiction. Such paradoxes motivated the development of axiomatic set theories, such as *Zermelo-Fraenkel set theory (ZF)*. Set theory's formalization is particularly relevant in areas such as Belief Revision, where the structure of information states can be modeled as sets of possible worlds or propositions. Just as axiomatic restrictions in set theory ensure consistency in mathematical reasoning, similar constraints in Belief Revision prevent inconsistencies when integrating new information. The transition from naïve to axiomatic set theory mirrors the refinement of Belief Revision frameworks, ensuring that modifications to an agent's knowledge remain logically coherent and well-founded.

The Standard Belief Revision AGM framework identifies three primary operations for the dynamics of changes on the belief sets:

1. **Expansion:** Adding a new belief to the set without regard for consistency. For example $B \cup \{\phi\}$. Expansion simply augments the belief set but may introduce contradictions.

2. **Contraction:** Removing a belief from the set to ensure consistency. For example, if $\neg\phi$ conflicts with existing beliefs, contraction eliminates ϕ or related propositions to resolve the conflict.

3. **Revision:** Incorporating a new belief into the set while maintaining consistency. Revision replaces conflicting beliefs with the new proposition and its implications. Formally, $B * \phi = (B - \neg\phi) \cup \{\phi\}$.

Belief Revision is guided by rationality principles to ensure coherence, maintain logical consistency, and adapt to new information in a systematic manner. These principles are fundamental to any formal model of belief change and can be presented as follows:

- **Consistency Preservation:** The revised belief set must remain free of contradictions. This ensures that the introduction of new information does not lead to logical inconsistencies that would render the belief system unreliable [3].

- **Minimal Change:** The revision process should make the smallest possible adjustment to the original belief set. This principle, often referred to as the *principle of minimal mutilation*, guarantees that prior knowledge is preserved as much as possible while integrating new information [32].

- **Closure Under Logical Consequence:** The belief set must remain closed under logical entailment. That is, if a proposition is inferred from the belief set, it must also be included in the revised belief set, ensuring deductive closure [39].

- **Success:** If a new piece of information is introduced, it must be incorporated into the revised belief set, provided it is not explicitly contradictory [3].

- **Priority of New Information:** When a conflict arises between existing beliefs and new information, the revision mechanism must prioritize the new input unless there is a compelling reason to reject it [39].

- **Iterativity:** Belief Revision should be dynamically stable, meaning that repeated applications of revision should follow a rational pattern, avoiding cycles or unnecessary reversals [29].

These principles form the foundation of classical Belief Revision frameworks, such as the *AGM theory* (Alchourrón, Gärdenfors, and Makinson), which formalizes how rational agents should incorporate new information into their epistemic states [3]. The integration of Belief Revision into formal logical systems often involves constraints that parallel axiomatic structures found in set theory, reinforcing the necessity of logical coherence in epistemic updates. Moreover, those principles reflect the intuitive idea that rational agents aim to retain as much of their original beliefs as possible while accommodating new evidence.

Traditional Belief Revision frameworks assume that propositions are context-independent. However, beliefs involving Indexicals, such as *"I am safe"* or *"Here is shelter"*, require a more nuanced approach. These expressions depend on contextual parameters, such as the speaker's identity, location, or temporal perspective. Addressing such context-sensitive beliefs necessitates extensions to standard frameworks, paving the way for tools like Two-Dimensional Modal Logic. Belief Revision provides a foundation for understanding how agents rationally update their beliefs. This thesis extends these ideas by incorporating context-sensitivity and dynamic reasoning through Two-Dimensional Modal Logic. The next sections explore how this integration addresses the challenges of context-dependent beliefs and Belief Revision in dynamic environments.

While the AGM framework provides a robust foundation for understanding Belief Revision, its standard formulation assumes that beliefs are context-independent and static. This assumption, while useful in many formal applications, becomes problematic when dealing with context-sensitive expressions, such as Indexicals or dynamically changing contexts. Traditional Belief Revision frameworks treat propositions as absolute statements, independent of the context in which they are expressed. For instance, the sentence *"The sky is blue."* is treated identically regardless of the speaker, time, or location. However, many everyday beliefs are inherently tied to context. No consider the following statement: *"I am safe here."* The truth of this statement depends on:

- **The speaker:** Who is expressing the belief.
- **The location:** Where the belief is expressed.
- **The time:** When the belief is expressed.

Standard Belief Revision frameworks lack the tools to handle these contextual parameters, making them ill-suited for reasoning about Indexicals or dynamic environments, but we cannot ignore the fact that beliefs often change over time as the context evolves. For example, suppose we say *"Today is a good day to find food."* This belief is valid only for a specific temporal context: namely, the day the utterance took place. As time progresses, the belief must be revised to reflect the new context. For instance, we would have to utter the corrected sentence *"Yesterday was a good day to find food."*. However, Standard Belief Revision frameworks do not account for the dynamic nature of such beliefs, treating them as static propositions rather than context-dependent expressions.

Thus, Indexicals, such as *"I,"* *"here,"* and *"now,"* pose unique challenges for traditional Belief Revision. Their meaning is inherently tied to the context of utterance, and their referents can shift depending on the speaker, time, and location. Sentences such as *"I am at Blueberry-Musk Camp now."* requires additional contextual information to determine its truth value. Traditional frameworks struggle to incorporate this information, leading to ambiguities and inconsistencies in Belief Revision. Moreover, Belief Revision is often nonmonotonic, meaning that new evidence can invalidate previously held beliefs. For example, *"There is food at the camp."* might be revised if the agent discovers that the camp has been scavenged. Unfortunately, AGM Belief Revision frameworks do not fully capture this defeasibility and context-sensitiveness, relying instead on monotonic reasoning where adding information never negates earlier conclusions.

Another limitation arises when aggregating beliefs from multiple agents or contexts. Suppose I say *"I saw food here yesterday."*. That statement might conflict with another agent's statement of *"There was no food here yesterday."*. Resolving such conflicts requires a framework capable of integrating context-sensitive parameters, something traditional Belief Revision frameworks do not address. The limitations outlined above highlight the need for a more flexible approach to Belief Revision, one that incorporates context-sensitivity, indexicality, and dynamic reasoning. This thesis explores the potential of integrating Two-Dimensional Modal Logic into Belief Revision frameworks, providing tools to address these challenges and extend the scope of traditional systems. The next sections develop these ideas further, presenting formal systems that operationalize this integration.

1.1.6 Defeasible Reasoning

Reasoning is rarely a static process. In everyday life, beliefs are formed based on available information but are subject to change as new evidence emerges. Defeasible reasoning captures this dynamic nature by allowing conclusions to be withdrawn or revised when additional information contradicts prior assumptions. Unlike classical logic, which enforces strict monotonicity—where once a conclusion is drawn, it remains valid regardless of new premises—defeasible reasoning acknowledges that knowledge is often incomplete, contextual, and evolving. This makes it particularly relevant for fields that deal with uncertainty, competing information, and dynamic updates, such as artificial intelligence, epistemology, and formal logic. Defeasible reasoning is distinguished by several key features that set it apart from traditional deductive reasoning:

- **Nonmonotonicity:** In classical logic, adding new premises to an argument does not invalidate previously established conclusions. In contrast, defeasible reasoning permits the retraction of conclusions when conflicting evidence arises. For instance:

  Premise: "Birds can fly."

  Conclusion: "Penguins can fly."

  New Evidence: "Penguins are flightless." (Conclusion is revised).

- **Default Reasoning:** Many conclusions in defeasible reasoning are based on general rules that hold in typical cases but may be overridden by specific counterexamples. For example:

  "Typically, if it is cloudy, it will rain."

  (However, in certain regions, cloudy weather does not imply precipitation.)

- **Defeaters:** A defeater is new information that invalidates a previously accepted conclusion. This is a key mechanism in Belief Revision, as it ensures adaptability. Consider:

  Premise: "There is food at the camp." Conclusion: "I can eat."
  Defeater: The food is spoiled. (Conclusion is withdrawn).

- **Priority Among Conflicting Rules:** In cases where multiple rules apply but lead to contradictory conclusions, defeasible reasoning employs priority mechanisms to resolve conflicts.

  Rule 1: "If it is cloudy, it will rain."
  Rule 2: "If it is cloudy and the wind is strong, it will not rain."
  (The latter rule takes precedence in a stormy context.)

Defeasible reasoning diverges from classical logic in several fundamental ways. Unlike classical logic, which assumes that valid inferences are immutable, defeasible reasoning accommodates change in the face of new information, increasing flexibility. That makes it suitable to handle uncertainty: Classical logic is based on absolute truths, whereas defeasible reasoning operates in environments where conclusions are subject to revision. Therefore, where classical logic struggles with contradictions, defeasible reasoning provides structured methods to manage conflicting information and prioritize beliefs accordingly. The adaptability of defeasible reasoning has led to its integration into various disciplines including *Belief Revision* and its ability to revise beliefs in response to new information is central to epistemic logic and rational decision-making.

Moreover, the principles of defeasible reasoning align naturally with Belief Revision. Both frameworks seek to provide a rational mechanism for updating an agent's knowledge base when confronted with new, and sometimes conflicting, information. Defeasible reasoning allows conclusions to be retracted in light of stronger evidence, mirroring the way Belief Revision mechanisms reconsider prior beliefs when contradictions arise. This adaptability is crucial for maintaining epistemic coherence, particularly in dynamic environments where information is constantly evolving. Furthermore, both frameworks emphasize rationality constraints, ensuring that belief updates are not arbitrary but instead follow principled modifications based on logical consistency and minimal change. The iterative nature of belief updates within these systems allows for a structured approach to reasoning, where conclusions are revised systematically rather than through ad hoc modifications.

Since Belief Revision fundamentally deals with dynamic epistemic states, defeasible reasoning offers a powerful framework for handling uncertainty, conflicting information, and changing contexts. Consider, for instance, a belief statement such as *"I am safe."* The validity of this belief is context-dependent, influenced by factors such as time, location, and external circumstances. A Belief Revision system equipped with defeasible reasoning can dynamically assess new evidence—such as receiving a warning about a potential threat—and adjust the epistemic state accordingly. This ability to retract, modify, or reinforce beliefs based on new input ensures that rational agents can navigate complex environments with greater reliability. By integrating defeasible reasoning into Belief Revision frameworks, these systems gain the flexibility necessary to model real-world reasoning and decision-making processes with greater accuracy and responsiveness.

The interplay between defeasible reasoning and Belief Revision reflects broader philosophical concerns regarding rational agency, epistemic justification, and context-sensitive reasoning. Both frameworks embody principles of *fallibilism*, recognizing that knowledge is subject to revision in light of new evidence, and *coherence theories of justification*, ensuring that beliefs remain logically consistent. Moreover, their reliance on contextual factors resonates with the role of *indexicals* in epistemic updates, a central theme in this thesis. By integrating these concepts, the formal study of belief dynamics gains a richer foundation that aligns with both logical rigor and philosophical insights into human reasoning.

1.2 Key Philosophical Concepts

At the core of this thesis lie key philosophical concepts that are both intricate and foundational, namely *Belief*[1] and *Contexts*. These notions have been extensively explored in the philosophical literature, spanning epistemology, logic, and philosophy of language. However, this thesis will focus on specific formulations of these concepts that are particularly relevant to Belief Revision and context-sensitive reasoning. By drawing from established theories while refining their application to dynamic epistemic frameworks, we aim to provide a structured and precise analysis of how beliefs are formed, maintained, and revised in varying contexts.

[1]Some belief ascriptions in this work appear in the form of quoted sentences — for example, "the agent believes 'I will be rescued.'" This mode of presentation is meant to highlight the agent's perspective in context-sensitive situations, particularly where indexicals or temporal expressions play a central role. However, belief contents are not to be understood as linguistic entities but as structured propositions — objects evaluable across possible worlds. The quotation marks are used informally, for illustrative clarity, especially in narrative examples. In future versions, a more systematic notation will be adopted, using expressions such as Bel(a, [p]), where [p] designates propositional content abstracted from its surface form. This adjustment aims to reinforce the underlying assumption already in place: belief is modeled as a relation to content, not to sentences.

1.2.1 Context and Indexicality

The concepts of **context** and **indexicality** are central to understanding how language and reasoning interact in dynamic and context-sensitive environments. Context shapes the interpretation of many expressions in natural language, and Indexicals, in particular, depend heavily on the specific circumstances of their utterance. This subsection introduces these key concepts and their philosophical significance, laying the groundwork for their formal treatment in subsequent sections. In the Philosophy of Language, **context** refers to the situational parameters that influence the meaning of an utterance. These parameters typically include:

- **Speaker:** The individual producing the utterance.
- **Time:** The temporal moment at which the utterance is made.
- **Location:** The spatial position of the speaker.
- **Audience:** The intended recipients of the utterance.

For example, the sentence *"I am here now."* derives its meaning from the context in which it is uttered. The speaker's identity, their current location, and the time of utterance determine the referents of the Indexicals *"I," "here,"* and *"now."* Without these contextual details, the sentence remains indeterminate. Thus, **Indexicals** are expressions whose referents shift depending on the context of utterance. Traditional examples include:

Personal Pronouns such as *"I," "you," "he/she."*; *Temporal References* such as *"now," "today," "yesterday."*; *Spatial References* such as *"here," "there."*, and more. Kaplan's seminal work on the semantics of Indexicals provides a framework for understanding their behavior. According to Kaplan [46, p. 491], the meaning of an indexical expression involves:

- **Character:** The linguistic rule that determines the referent based on the context.
- **Content:** The actual referent of the expression in a given context.

For instance, the character of *"I"* is the rule that refers to the speaker of the utterance, while the content of *"I"* in a specific utterance might be "Ana," the author of this thesis. Indexicals are philosophically significant because they highlight the interplay between language, thought, and context. Perry [73, p. 3] emphasizes the sensitivity of thoughts and utterances to the context in which they occur: *"It is not only that language is sensitive to the user and his context; the thoughts we have, that our utterances express, are similarly sensitive."*[73, p. 3].

This context-sensitivity poses challenges for formal systems, particularly those attempting to model rational Belief Revision: How can beliefs expressed using Indexicals, such as *"I am safe"*, be revised when the context changes?; What mechanisms ensure that beliefs involving Indexicals remain consistent across different contexts? The concepts of context and indexicality are central to the investigation of rational Belief Revision in this thesis. By integrating these notions into formal systems like Two-Dimensional Modal Logic, this work aims to address the challenges posed by context-sensitive reasoning and belief updates.

1.2.2 Belief and Knowledge

Belief and knowledge are central concepts in epistemology, the branch of philosophy that studies the nature and scope of human understanding. In the context of this thesis, the distinction between belief and knowledge, as well as their interaction with context, plays a critical role in modeling rational reasoning and Belief Revision.

Belief is a mental state or attitude in which an agent holds a proposition to be true. Beliefs are:

- **Subjective:** They reflect the agent's perspective and may not necessarily align with objective facts.
- **Propositional:** They are directed at propositions, which can be true or false.

For example:

"I believe it will rain today."

Moreover, beliefs play a fundamental role in our reasoning processes and decision-making. They are not static but rather dynamic entities that are constantly updated in light of new evidence, shifting perspectives, or changing contexts. Consider the statement:

"It will rain today."

This statement expresses the agent's belief that the proposition *"It will rain today"* is true. However, beliefs are susceptible to change based on new information. If the agent later receives a weather report predicting clear skies or observes a shift in atmospheric conditions suggesting otherwise, they may revise their belief accordingly. This adaptability is characteristic of defeasible reasoning, where conclusions can be withdrawn when new, overriding evidence emerges. Understanding the mechanisms behind Belief Revision is crucial in this investigation, as it informs the broader discussion of context-sensitive reasoning and indexical expressions.

Traditionally, knowledge has been defined as **justified true belief** (JTB). According to this model, for an agent to claim knowledge of a proposition, three conditions must be met:

1. **Belief:** The agent must hold the proposition to be true. Without belief, knowledge cannot exist, as an agent cannot know something they do not believe.
2. **Truth:** The proposition must correspond to reality. A belief that is false, even if justified, does not constitute knowledge.
3. **Justification:** The agent must have sufficient evidence, reasons, or a rational basis for holding the belief.

For example, consider the following assertion:

"I know it is raining because I can see it through the window."

In this case, the agent believes that it is raining, their belief is true, and their justification—direct observation—supports their belief. The combination of these three elements allows the agent to claim knowledge under the JTB model. Despite the apparent robustness of the JTB model, Edmund Gettier [35] famously challenged its sufficiency by presenting cases where justified true beliefs fail to constitute knowledge. Gettier's problem arises when an agent holds a belief that is both justified and true, yet the justification is flawed or coincidental. For instance, suppose an agent believes:

"There is a clock on the wall, and it shows that it is 3:00 PM. Therefore, I know it is 3:00 PM."

If the clock has stopped but coincidentally shows the correct time when the agent checks it, their belief is both true and justified, but it does not amount to knowledge. This example illustrates how justification, in some cases, fails to ensure that a belief is genuinely knowledge. As a result, philosophers have sought alternative or refined models of knowledge to address the shortcomings of the JTB framework. We will acknowledge this debate and focus on the Propositional Beliefs, even when talking about Knowledge. While belief and knowledge are closely related, they are not synonymous. However, in this investigation, our primary focus is not to resolve the longstanding philosophical debates about the nature of knowledge but rather to examine how indexical expressions interfere in Belief Revision. Indexicals—such as "I," "here," and "now"—play a crucial role in shaping beliefs by anchoring them to specific contexts. When an agent updates their beliefs, they must navigate these contextual dependencies, ensuring that revisions account for shifts in temporal, spatial, or epistemic perspectives.

For example, an agent might say: *"Yesterday, I believed that the storm would arrive today."* If today arrives and the storm does not materialize, the agent must revise their belief while recognizing the role of indexical terms like "yesterday" and "today" in tracking changes over time. By integrating these considerations, we move toward a framework that accommodates the fluidity of Belief Revision in the presence of indexicality and context-dependent reasoning. Rather than treating beliefs as static entities, we recognize them as evolving cognitive structures that interact with shifting epistemic landscapes, requiring formal mechanisms for systematic and rational updates.

Beliefs involving Indexicals and context-sensitive expressions are particularly dynamic. For example: *"I am safe here."* may be revised if new information indicates a threat in the same location. This highlights the need for a framework that accommodates the interplay between belief, context, and change. Moreover, the distinction between belief and knowledge becomes blurred when dealing with context-sensitive reasoning. For instance, the belief: *"I am at Blueberry-Musk Camp."* might be considered knowledge if supported by sufficient evidence, but the same belief could lose its epistemic status if the context changes (e.g., the agent moves to a different location).

We focus on the formal modeling of beliefs, particularly context-sensitive and indexical beliefs, within a framework that integrates Belief Revision and Two-Dimensional Modal Logic. Knowledge, while not the primary focus, provides a useful contrast for understanding the limitations and challenges of Belief Revision. The next subsections explore additional philosophical concepts, such as possible worlds and rationality, which further discuss the dynamics of belief and context-sensitive reasoning. This discussion naturally leads to an exploration of **Possible Worlds and Centered Worlds**, which serve as essential tools for understanding how context influences belief formation and revision. By analyzing how beliefs are structured across different possible worlds and how centered perspectives shape epistemic commitments, we gain deeper insight into the fluid nature of belief dynamics.

1.2.3 Possible Worlds and Centered Worlds

Possible worlds are a cornerstone of philosophical and logical inquiry, providing a framework for understanding modality statements about what is possible or necessary. By imagining alternative ways the world could be, philosophers can explore questions about truth, counterfactuals, and the nature of reality. However, standard Possible Worlds Semantics, while powerful, is often insufficient for reasoning about context-sensitive beliefs. The introduction of centered worlds, which incorporate an agent's perspective within a possible world, addresses these limitations. This subsection explores the philosophical foundations of possible worlds, their limitations, and how centered worlds extend the framework to account for context and perspective.

The idea of possible worlds has deep roots in the history of philosophy but was systematized in modern Modal Logic by Saul Kripke [52] and David Lewis [60]. Possible worlds are often described as "ways the world might have been." They are used to evaluate the truth of modal statements such as:

"It is possible that the rescue team will arrive today."

In this framework, a statement is necessary if it is true in all possible worlds and possible if it is true in at least one.

Philosophers differ on the nature of these worlds. Kripke famously argued for an **actualist** interpretation, where possible worlds are abstract representations or descriptions of alternative scenarios. For Kripke, these worlds exist conceptually rather than concretely. As he states: *"Possible worlds are stipulated, not discovered by powerful telescopes."* [52, p. 44]

In contrast, David Lewis defended **modal realism**, a more controversial view that treats all possible worlds as equally real. According to Lewis:

> *"The worlds are many and varied. They are more numerous than the philosophers have dreamed, but they are no less real for being so abundant."* [60, p. 84]

Regardless of their metaphysical differences, both views agree on the utility of possible worlds for evaluating statements about necessity and possibility. For instance, in the Wilderness scenario, the belief:

"If I stay at Blueberry-Musk Camp, I will have food tomorrow."

can be understood as asserting that in all possible worlds consistent with the agent's evidence, food is available at the camp. This use of possible worlds enables agents to reason about hypothetical scenarios systematically.

While Possible Worlds Semantics is a powerful tool, it encounters significant challenges when applied to context-sensitive reasoning. First, possible worlds are often too coarse-grained to capture the nuances of context. Statements like:

"I am here now."

depend on the speaker's identity, location, and temporal situation details not inherently included in a possible world. Without additional contextual parameters, the meaning of such statements remains indeterminate. Moreover, possible worlds struggle to represent the perspective of an agent situated within a world. For example, in the Wilderness narrative, your belief:

"I am at Lily's Fish Pond."

cannot be evaluated purely within standard Possible Worlds Semantics, as it requires information about who *"I"* refers to and where they are in the world. These limitations reveal the need for a more refined framework capable of capturing the subjective perspective of agents. **Centered worlds** extend the framework of possible worlds by introducing a center that captures the perspective of an individual agent. A centered world is represented as a pair $\langle w, c \rangle$, where w is a possible world and c is a center, typically specifying the agent and their spatiotemporal location within the world. This refinement allows for a more precise analysis of context-sensitive expressions. For instance, in the Wilderness example:

- A possible world w_1 might represent a scenario where rescue arrives on Day 100.
- A centered world $\langle w_1, c_1 \rangle$ specifies that you, the stranded individual, are at Ox-Foxy Hill on Day 50.

In this sense, centered worlds are particularly useful for interpreting Indexicals such as *"I," "here,"* and *"now."* As David Lewis noted:

> *"What distinguishes me from others is not my being in this world, but my being at this time, in this place, and experiencing the world from this point of view."* [59, p. 516]

This emphasis on perspective allows centered worlds to model self-locating beliefs and dynamic contexts more effectively than standard possible worlds. The introduction of centered worlds has profound implications for understanding reasoning and belief:

- *Indexicals and Self-Locating Beliefs:* Statements like *"I am safe here"* can be evaluated naturally within centered worlds, where the center specifies the speaker, their location, and the time of utterance.
- *Dynamic Reasoning:* centered worlds enable agents to update their beliefs as contexts evolve. For example, discovering evidence of a predator near Lily's Fish Pond might prompt you to revise your belief: *"This is a safe place to stay."*

- *Subjectivity and Rationality:* centered worlds highlight the importance of subjective perspective in reasoning, challenging purely objective accounts of belief and knowledge.

The integration of centered worlds into Belief Revision frameworks allows for a more nuanced approach to rationality. Rational agents must revise their beliefs in response to new evidence, maintaining consistency and coherence while adapting to dynamic contexts. For example, *"I believe the rescue team will arrive soon."*, might need to be revised if you observe no signs of rescue after 100 days. centered worlds provide the structure to evaluate such beliefs and their updates, incorporating both the agent's perspective and the evolving context.

1.2.4 Agents and Rationality

The concept of an agent lies at the core of rational reasoning and Belief Revision. In this thesis, an agent is defined as a reasoning entity capable of forming, maintaining, and revising beliefs in response to changes in their environment. Unlike passive observers, agents actively engage with their surroundings, relying on their beliefs to navigate the world and make decisions. This dynamic interaction between the agent and their context is particularly crucial in scenarios involving uncertainty and incomplete information. For example, in the Wilderness scenario, the stranded individual acts as an agent striving to survive in a challenging and unpredictable environment. Their beliefs, such as *"Blueberry-Musk Camp is safe"* or *"Lily's Fish Pond has reliable food sources"*, directly influence their actions and outcomes. As new evidence emerges such as signs of predators or dwindling food supplies these beliefs must be revised to ensure continued survival.

Rationality is the guiding principle that ensures an agent's beliefs remain coherent and adaptable in the face of change. In the context of Belief Revision, rationality is characterized by several core principles that govern how agents update their beliefs. The first principle is **consistency**, which requires that an agent's beliefs do not contradict one another. For instance, an agent cannot simultaneously believe *"Lily's Fish Pond is a safe place"* and *"Lily's Fish Pond is unsafe"* without violating logical coherence. As Gärdenfors points out: *"A belief system must be internally consistent to be useful for making rational decisions. Contradictory beliefs can lead to irrational choices or inaction."* [32, p. 32]. Consistency is a cornerstone of rational belief systems, ensuring that the agent's reasoning process is logically sound.

The second principle is **minimal change**, which dictates that when revising beliefs, agents should strive to make the smallest possible adjustment to their existing belief set. This principle reflects the intuitive idea that beliefs should not be abandoned or altered unnecessarily. For example, if the agent discovers that fish are scarce at Lily's Fish Pond today, they might revise the belief *"Lily's Fish Pond always has fish"* to *"Lily's Fish Pond usually has fish"* without discarding the entire belief system related to the pond. Alchourrón, Gärdenfors, and Makinson formalize this principle in the AGM framework, stating:

> *"When revising a belief set, the changes should be as minimal as possible to accommodate the new information while retaining as much of the original set as consistent."* [2, p. 516]

A third critical principle is **contextual sensitivity**, which emphasizes the importance of adapting beliefs to reflect the current context accurately. Many beliefs are inherently tied to specific times, places, or perspectives, and failing to account for these dependencies can lead to irrational conclusions. Consider the belief *"I am safe here now"*. This belief is valid only within a specific spatiotemporal context, and its truth depends on factors such as the agent's current location and the presence of potential threats. If the context changes such as the discovery of a predator nearby the belief must be revised accordingly to maintain rationality. As Lewis explains:

> *"Rationality is not merely a matter of holding consistent beliefs; it requires responsiveness to the context and the evidence at hand."* [59, p. 518]

centered worlds provide a formal framework for representing these context-sensitive beliefs, allowing agents to reason effectively in dynamic environments.

Agents operating in dynamic and uncertain contexts face additional challenges that complicate the application of rational principles. One significant challenge is the cognitive limitations of human agents, who often have finite resources for processing information and reasoning. These limitations can result in suboptimal decisions or inconsistencies in Belief Revision. Another challenge is the nonmonotonic nature of reasoning in dynamic environments. Unlike classical logic, where adding new premises cannot invalidate existing conclusions, everyday reasoning often requires agents to retract beliefs when confronted with contradictory evidence. For example, if the agent initially believes *"There is food at Blueberry-Musk Camp"* but later discovers that the camp has been scavenged, they must withdraw their original belief and adjust their plans accordingly. In multi-agent scenarios, conflicting beliefs further complicate the rationality of individual agents. Consider a situation where two agents provide contradictory reports:

Agent A:*"I saw food at Blueberry-Musk Camp yesterday."*

Agent B:*"There was no food at Blueberry-Musk Camp yesterday."*

Resolving such conflicts requires a mechanism for prioritizing evidence, reconciling differences, and updating beliefs in a way that preserves consistency while respecting the available information.

To address these challenges, rational agents rely on formal processes of Belief Revision. Belief Revision provides a systematic framework for updating beliefs in response to new evidence or changing contexts. By adhering to the principles of consistency, minimal change, and contextual sensitivity, Belief Revision ensures that agents can adapt to dynamic environments while maintaining logical coherence. In the Wilderness example, Belief Revision allows the agent to integrate new observations such as the presence of predators or the scarcity of food into their reasoning process, enabling better decision-making and improved chances of survival.

The integration of centered worlds into Belief Revision frameworks further enhances an agent's ability to reason about context-sensitive beliefs. By explicitly representing the agent's perspective within a possible world, centered worlds allow for more precise modeling of beliefs tied to specific contexts, such as *"I am at Blueberry-Musk Camp now"*. This integration not only addresses the limitations of standard Possible Worlds Semantics but also provides a more robust foundation for understanding rationality in dynamic and uncertain environments. This discussion of agents and rationality sets the stage for a deeper exploration of Belief Revision frameworks, which formalize the principles and processes discussed here.

1.2.5 Kinds of Beliefs

Before delving into belief changes, it is essential to explore what constitutes a belief. In the philosophical literature, belief is often considered a *propositional attitude*, which refers to an agent's mental state concerning the content of a proposition [55, p. 23]. A rational agent holds beliefs that she considers true and consistent, and these beliefs influence her actions. For instance, in the Wilderness example, if the agent believes that the front door is open, she may choose to close it to secure shelter. There are various types of beliefs, but they can generally be divided into two main categories: *qualitative* (or *all-or-nothing*) beliefs, and *quantitative* beliefs, or *degrees of belief*.

Suppose, for example, I believe that *"my front door is open"* is true. In this case, I would not simultaneously believe that *"my front door is not open"* is true, meaning that I consider the proposition *"my front door is not open"* to be false. In this case, we are dealing with a qualitative belief there is no middle ground: the belief is either true or false. However, suppose that I am uncertain about whether the front door is open and would prefer to check rather than act immediately. In this case, my belief that *"my front door is open"* is true might be less than certain, and I might assign it a degree of belief, such as *"probably/possibly true"*. Therefore, I would not hold that the opposite proposition, *"my front door is not open"*, is necessarily false, and my belief is more flexible.

Qualitative beliefs can be expressed by saying *"the agent x believes that proposition φ is true"* or *"the agent x believes that proposition φ is false"*, or even *"the agent x does not believe that proposition φ is true or false"* [55, p. 23]. In contrast, quantitative beliefs are usually represented using probabilistic approaches, assigning a degree of belief within the interval $[0,1]$, where 0 represents complete disbelief, and 1 represents certainty. Some authors argue that a qualitative belief in a proposition φ corresponds to assigning the maximum probability value of 1 to φ [54]. However, reducing all-or-nothing beliefs to quantitative ones is a more complex issue [54].

Lewis distinguishes between two types of propositional attitudes: *De Dicto* and *De Se* beliefs [56, p. 67]. The term "proposition," for Lewis, refers to a set of possible worlds, a region in logical space. When an agent believes in a proposition, they identify themselves as inhabitants of the possible world where the proposition holds. However, some beliefs cannot be understood as propositional; instead, they represent self-ascriptions of properties [56]. For example, beliefs about one's location in time and space are not merely propositions but are *De Se* beliefs, reflecting self-locating attitudes.

Another term for *De Se* beliefs, as used by Lewis, is *Self-Locating Beliefs*, a concept developed by Perry [73, p. 102]. These beliefs represent an agent's self-awareness of their identity within a specific context. To represent self-locating beliefs in logical space, we move beyond simple sets of possible worlds and introduce the notion of Centered Possible Worlds, which accounts for attitudes *De Se*. Lewis [56] attributes the original concept of centered worlds to Quine, although Quine did not fully develop a systematic framework. For Lewis, Quine's [2] ideas of replacing propositions with properties as objects of belief and knowledge lay the groundwork

[2] Quine's influence on the concept of centered worlds stems from his broader critique of intensional semantics and his preference for treating belief ascriptions in terms of properties rather than propositions. In Word and Object [82], Quine challenges traditional accounts of meaning and reference, arguing that belief attribution cannot be straightforwardly understood as a relation between an agent and a fixed proposition. Instead, he suggests that beliefs are better captured in terms of dispositions or properties an agent instantiates. Later, in The Roots of Reference [84] and Quiddities [85], Quine further develops his skepticism toward intensional entities and highlights the role of context in shaping epistemic states. Lewis [56] builds on this insight, interpreting Quine's ideas as a precursor to the centered-worlds framework. While Quine did not explicitly formulate a theory of centered possible worlds, his rejection of unqualified de dicto belief attributions in favor of context-sensitive accounts of meaning and reference provided a foundation for later developments in de se attitudes and self-locating beliefs.

for centered worlds.

The difficulty of expressing *De Se* beliefs highlights the challenges in formalizing context-sensitive beliefs. Stojanovic [103] points out that it is neither necessary nor sufficient for speakers to use the same sentence for their beliefs to be understood as "same-saying." This aligns with Perry's account of Indexicals and underscores the issues related to context-sensitivity in the expression of self-locating beliefs.

Beliefs about propositional objects, on the other hand, are classified as *De Dicto* beliefs, and these can be easily represented by sets of possible worlds without needing to account for contextual information. For example, scientific beliefs about established theories, which hold true across all possible worlds that adhere to the same physical laws, can be expressed in terms of *De Dicto* beliefs.

Nevertheless, there is no single, universally accepted way to represent a belief in formal systems [34]. The formalization of indexical beliefs presents significant challenges, especially when considering Schiffer's *Hidden Indexical Theory* [94]. This theory proposes that belief is a three-place relation $B(x, p, m)$ between a believer x, a proposition p, and a mode of presentation m. Schiffer's theory highlights the critical role of the mode of presentation, which shapes the belief's content. Ludlow [63] critiques Schiffer's theory by pointing out that the mode of presentation is often implicitly determined by the context, even when it is not explicitly stated.

Stojanovic [104] further emphasizes the distinction between holding beliefs about oneself (*De Se*) and beliefs about others (*De Dicto*). She suggests that understanding self-locating beliefs requires attention to the character and context of the expression used. This issue reflects the broader challenge of managing context-sensitive beliefs in formal systems, especially when addressing beliefs about spatial-temporal location and identity within a possible world.

1.3 Uncertainty and Belief Change

Human reasoning is full of uncertainties. Even when we believe ourselves to be certain about something, new evidence often forces us to reconsider our beliefs. The concept of *certainty* is an epistemic property of beliefs. Certainty is typically associated with the belief that a proposition α is true, but this certainty carries varying levels of confidence, ranging from psychological certainty to epistemic and moral certainty.

- A belief is *psychologically certain* when the agent is indubitably convinced of its truth;
- A belief is *epistemically certain* when it is supported by the highest possible epistemic justification;

- A belief is *morally certain* when it is sufficiently certain for everyday practical purposes.

While psychological and moral certainty may provide a sense of security in everyday life, epistemic certainty is a higher standard, often requiring rigorous evidence. Nonetheless, certainty does not guarantee truth, which underscores the importance of understanding and accounting for *uncertainty* in Belief Revision. Uncertainty, in this sense, plays a central role in human reasoning, as we rarely possess complete information.

In everyday situations, uncertainty is not only pervasive but often necessary for rational action. For example, when financing a house based on expected future income, we act with a certain level of certainty that our income will remain stable, even though external events such as job loss might cause the belief to change. Similarly, in scientific reasoning, theories that were once considered certain can be revised when new data emerges. For example, a well-established theory in science might be replaced with a more accurate theory as evidence accumulates.

Belief change due to uncertainty requires rational mechanisms for ensuring that beliefs remain consistent and coherent. In the case of a detective investigating a crime or scientists revising theories based on new evidence, the revision of beliefs depends on a robust system that ensures consistency. However, not all belief change is rational. A belief change should be justified by strong, credible evidence, and there must be mechanisms in place to avoid contradictions.

Belief Revision theories, such as those proposed by Rott [90], emphasize the need for a rational process when changing beliefs. According to Rott, belief is part of a larger pattern of attitudes that are ascribed to an agent by external observers, with the goal of rationalizing the agent's actions. This idea reinforces the need for logical consistency in Belief Revision.

> *"Belief is part of a larger pattern of attitudes ascribed to an agent by external observers that try to rationalize the agent's behavior. It follows that the belief component itself comes out as rational and coherent, and well-balanced to the degree that would never be obtained if Belief were only what the agent explicitly declares to believe, or would explicitly declare to believe if persistently interviewed."* [90, p. 11]

The role of uncertainty in Belief Revision and the necessity for a rational, coherent framework that allows agents to adapt their beliefs when confronted with new, uncertain information is essential in our investigation. It also highlights how probabilistic and qualitative approaches can be applied to better model the complexity of belief change in everyday scenarios.

1.3.1 Uncertainties and Contexts

Uncertainty is an inherent part of human cognition and decision-making, particularly in environments where information is incomplete or where the agent's perspective is limited. In the Wilderness scenario, the agent faces multiple forms of uncertainty: they do not know when they will be rescued, whether they are safe, or what day it is. These uncertainties affect the agent's beliefs, which must be formed and revised continuously as new information becomes available. The challenge for Belief Revision systems is to model this process, particularly when the uncertainty is tied to the context in which the belief is formed and the agent's shifting perspective.

Context plays a crucial role in shaping uncertainty, as it determines how an agent interprets information and updates their beliefs. In the Wilderness, for example, the agent's understanding of time and safety is contingent on their immediate surroundings and evolving knowledge. The belief *"I will be rescued in 100 days"* becomes increasingly uncertain as time passes, but it remains anchored in the context of the agent's current state. Similarly, the belief *"I am safe here now"* changes when new information such as the discovery of a predator alters the agent's understanding of their safety. Context-sensitive expressions, such as those involving Indexicals like *"I"*, *"here"*, and *"now"*, are integral to how agents form and revise their beliefs in the face of such uncertainty.

> *"Uncertainty, in the context of Belief Revision, arises when an agent must update or modify their beliefs due to the lack of information, which in turn requires adjustments in the agent's reasoning based on the evolving context."* [20, p. 12]

Context-sensitive reasoning refers to the idea that the meaning of certain expressions, and thus the beliefs they help form, is determined by the specific circumstances in which they are used. This feature of reasoning is particularly important when dealing with self-locating beliefs beliefs about one's own location, safety, or temporal position that depend on the agent's current context. A belief such as *"I am here now"* cannot be understood independently of the context in which it is expressed. Similarly, the passage of time and the agent's changing knowledge play a critical role in shaping their beliefs about future events, such as *"I will be rescued in 100 days"*.

In formal logic, uncertainty and context sensitivity introduce challenges that are not typically addressed in classical systems. Classical systems, such as propositional and First-Order Logic, treat beliefs as static propositions, with fixed truth values that do not depend on the agent's context. However, beliefs involving uncertainty and context-sensitive terms cannot be adequately represented in these frameworks. This motivates the need for a more expressive formal system that can capture how beliefs evolve as new information becomes available and how an agent's perspective shaped by their context affects the content and truth of their beliefs.

One of the key issues in Belief Revision is how uncertainty is handled. As new evidence is presented, beliefs may need to be revised, expanded, or contracted. The classical approach, as formalized in the AGM model of Belief Revision, treats belief sets as static collections of propositions that are updated in response to new information. However, this model does not account for the complex interplay between uncertainty and context. Defeasible reasoning, which allows for conclusions to be withdrawn or revised in the face of new evidence, is necessary to handle these situations, especially when the evidence affects beliefs that are context-sensitive.

As Gärdenfors notes:

> *"A belief system must be able to cope with changing information, especially when the information is uncertain or incomplete. In this context, defeasible reasoning plays a critical role in updating beliefs in response to new evidence."* [32, p. 15]

Defeasible reasoning provides a mechanism for revising beliefs in uncertain and dynamic environments, such as the one described in the Wilderness scenario. This process is essential for Belief Revision, particularly when the beliefs in question are influenced by changing contexts, new evidence, or evolving circumstances. Thus, the combination of uncertainty and context sensitivity introduces significant challenges for formal systems of Belief Revision. These challenges must be addressed in order to develop a framework that can model how agents revise their beliefs in dynamic and context-dependent environments.

1.3.2 Navigating Context in Possible Worlds Semantics

Understanding how beliefs evolve, especially in dynamic and uncertain environments, requires a robust framework that accounts for the fluidity of meaning and reference. One powerful tool in this regard is David Lewis' analysis of counterfactuals within the realm of Possible Worlds Semantics. Counterfactuals are statements of the form *"If X had happened, Y would have occurred"*, where X and Y describe potential events that did not actually happen. These conditionals allow us to reason about alternative worlds, capturing the idea that the truth of a statement can vary depending on which world we consider. For instance, in the Wilderness scenario, the belief *"I will be rescued in 100 days"* depends not just on the actual world but also on possible worlds where different outcomes unfold.

Lewis' work on counterfactuals introduces the concept of evaluating these conditionals based on the "closest possible worlds," which are alternative scenarios that are most similar to the actual world. The truth of a counterfactual is determined by examining these worlds and seeing whether the antecedent (X) holds and, if so, whether the consequent (Y) follows. This approach helps to formalize how beliefs change when the agent considers different possibilities, as in the case of the Wilderness example, where the agent revises their belief about being rescued based on new possible worlds (alternative rescue plans, delayed schedules, etc.). Moreover, according to Lewis, *"A counterfactual conditional is true if and only if in the nearest possible world where the antecedent is true, the consequent is true as well."* [58, p. 2].

This framework provides the foundation for understanding how counterfactual reasoning operates within possible worlds, and by extension, how we can model Belief Revision in dynamic environments. The key insight here is that counterfactuals depend on the structure of possible worlds, which in turn depends on how terms are interpreted and what conditions are considered relevant for evaluating their truth. In this way, Possible Worlds Semantics connects counterfactuals with intension and extension, two key concepts in understanding the dynamics of meaning in context-sensitive beliefs.

1.3.3 Intension and Extension: The Semantic Building Blocks

To fully appreciate how counterfactuals work in Possible Worlds Semantics, it is essential to understand the relationship between intension and extension, two fundamental concepts that govern how terms refer to objects in different possible worlds. These distinctions are vital for understanding the meaning of beliefs, especially those involving Indexicals like *"I"*, *"here"*, and *"now"*. They can be presented as follows:

Intension The intension of a term refers to the rule or function that determines how its reference is understood across different possible worlds. It captures the meaning of a term and its possible interpretations, depending on the context. For example, the term *"I"* in the Wilderness scenario refers to the agent, but the meaning of this term (its intension) is linked to the agent's perspective at a given moment. If the agent moves, the intension of *"I"* remains the same, as it still refers to the agent, but the extension (the actual referent) changes with the agent's location.

In counterfactual reasoning, the intension of terms determines how they refer to objects or situations in alternative worlds. When evaluating a counterfactual like *"If I had stayed at the lake, I would have been safe"*, the intension of *"I"* refers to the agent, while the intension of *"safe"* refers to a concept of safety that depends on the agent's situation. The truth of the counterfactual depends on the intension of the terms involved in the context of the nearest possible world.

Extension The extension of a term refers to the set of actual objects or events that it refers to in a given world. While the intension determines the meaning of the term, the extension represents the actual referent in the context of a specific world. For instance, in the Wilderness, the extension of the term *"safe"* refers to the areas where the agent is truly safe at any given moment. If the agent moves to a new part of the wilderness and discovers a predator, the extension of *"safe"* changes, but the intension of the term remains fixed.

The connection between intension and extension is crucial for understanding how Belief Revision works. In the case of counterfactuals, the intension of a term remains constant across worlds, but its extension may vary depending on the specific world in question. For example, in a world where the agent stays at the lake, the extension of *"safe"* would include that area, but in the actual world, the agent's understanding of safety might change based on new evidence. This relationship helps us formalize how beliefs involving Indexicals and self-locating terms are formed and revised as the agent's context shifts.

> *"The intension of a term is a function that determines its reference across possible worlds, while the extension refers to the actual referents in the actual world. This distinction is essential for understanding how Belief Revision works in dynamic contexts."* [20, p. 115]

Moreover, the relationship between Counterfactuals, intension, and extension provides a powerful tool for understanding Belief Revision in dynamic contexts. In the Wilderness scenario, the agent's belief *"I will be rescued in 100 days"* is contingent on their knowledge of the future, which is expressed through counterfactual reasoning. As the agent encounters new evidence (e.g., a delayed rescue operation), they revise their belief by considering alternative worlds where different rescue outcomes occur. In counterfactual reasoning, the intension of terms like *"I"* and *"rescued"* determines how these terms are understood across different possible worlds. The extension changes as the agent updates their beliefs, reflecting the evolving understanding of their safety and rescue prospects. This interplay between intension and extension is central to formalizing how beliefs are revised in response to new evidence, especially in cases where the agent's beliefs are contingent on context-sensitive terms.

The formalization of these concepts allows us to address the central problem of Belief Revision in dynamic, uncertain environments. By utilizing Possible Worlds Semantics and incorporating intension and extension, we can better understand how self-locating beliefs evolve over time and how agents revise their beliefs in light of changing contexts.

1.4 Dynamics of Context-Sensitive Reasoning: an issue

The central questions explored in this thesis concern the formation and revision of beliefs, particularly in environments that are dynamic and context-dependent. Context-sensitive reasoning, especially involving Indexicals and self-locating beliefs, poses a fundamental challenge to formal logic and Belief Revision systems. These challenges arise in scenarios where beliefs are not static, but instead shift depending on the agent's evolving understanding of their circumstances. As illustrated by the Wilderness scenario, beliefs about one's own position such as *"I am here now"* or *"I will be rescued in 100 days"* depend crucially on the agent's context, including their location, time, and knowledge. This thesis investigates the formalization of these beliefs and how they can be revised as new, context-relevant information becomes available.

Let us briefly discuss the key issues underpinning this research, focusing on how context influences reasoning and belief formation. We will examine the limitations of classical Belief Revision frameworks when applied to context-sensitive beliefs and highlight the role of formal logic in addressing these challenges. By introducing and exploring formal systems that can accommodate these complexities, this thesis provides a foundation for understanding how agents revise beliefs in response to dynamic and shifting contexts.

1.4.1 Context-Sensitive Reasoning in Belief Formation

Context-sensitive reasoning is an essential component of belief formation, as individuals often rely on terms whose meaning is directly tied to the context in which they are used. Expressions such as *"I," "here,"* and *"now"* are prime examples of Indexicals terms whose reference depends on the context of the utterance. In formalizing belief systems, the challenge is to capture how these terms contribute to the formation and revision of beliefs. As Kaplan explained:

> *"The referent of a pure indexical depends on the context, and the referent of a demonstrative depends on the associated demonstration."* [46, p. 491]

Kaplan's distinction between character and content is essential to understanding the role of Indexicals in belief formation. The *character* of an indexical is the rule that determines its reference in any given context, while its *content* is the specific referent it takes in a particular context. For example, the statement *"I am here now"* is context-dependent, with both the meaning of *"I"* and *"here"* changing depending on the speaker and the time of the utterance. These shifts in meaning require a formal system that can adapt to changes in context, particularly in cases of self-locating beliefs.

Perry's example of *"I"* further illuminates the role of Indexicals in belief formation. He argues that Indexicals such as *"I"* not only refer to a specific individual but also incorporate the individual's perspective and experiences:

> *"It is not only that language is sensitive to the user and his context; the thoughts we have, that our utterances express, are similarly sensitive."* [73, p. 3]

This underscores the point that beliefs involving Indexicals are shaped not only by the objective content of the proposition but also by the agent's perspective. Understanding this requires a formal framework that accounts for both the context in which the belief is held and the shifting reference of Indexicals.

1.4.2 Challenges in Belief Revision

Classical systems of Belief Revision, such as those based on propositional and First-Order Logic, are effective for reasoning in domains where beliefs are context-independent and the meaning of propositions remains fixed. However, these systems face significant limitations when dealing with context-sensitive beliefs, particularly those involving Indexicals. In classical systems, the meaning of a statement is treated as static: a belief is either true or false, independent of the agent's context.

For instance, in Propositional Logic, a simple belief such as *"It is windy"* can be represented as a proposition with a truth value. However, this approach fails to address the dynamic nature of Indexicals. When an agent says *"I am safe here now"*, the belief is not about a static fact; it is influenced by the agent's position and knowledge, which are constantly shifting. The belief may need to be revised as the agent's understanding of their environment evolves, making it difficult to represent within a classical framework. As Gärdenfors points out:

> *"A belief system must be internally consistent to be useful for making rational decisions. Contradictory beliefs can lead to irrational choices or inaction."* [32, p. 32]

While consistency is a crucial requirement, classical Belief Revision systems do not address the challenges posed by beliefs that depend on the agent's perspective. Context-sensitive beliefs, like those involving Indexicals, often require revision when new information or evidence emerges, and this process must account for the shifting context in which the belief was originally formed.

1.4.3 The Need for a More Expressive Formal Framework

The limitations of classical Belief Revision systems in handling context-sensitive beliefs have led to the development of more expressive frameworks, capable of modeling beliefs that change depending on context. Two-Dimensional (2D) Semantics, as introduced in earlier sections, is one such framework that distinguishes between the primary intension (the meaning across all possible worlds) and the secondary intension (the meaning relative to the agent's context). This approach allows for the formalization of beliefs that depend on the agent's perspective and the shifting nature of their environment.

The integration of defeasible reasoning into Belief Revision systems is another important development. Defeasible reasoning refers to the type of reasoning where conclusions can be withdrawn or revised in light of new evidence. This is particularly relevant when beliefs are subject to change based on the agent's context. As Antonelli explains: *"Defeasible reasoning reflects the adaptive nature of human cognition, where conclusions are revised or retracted based on new information that may contradict previous assumptions."* [4, p. 27] This type of reasoning is crucial for modeling Belief Revision in dynamic environments, where new, context-sensitive information continually forces the agent to reconsider their beliefs.

1.4.4 The Dynamics of Context-Sensitive Reasoning

The formalization of context-sensitive reasoning and Belief Revision has far-reaching implications beyond philosophy and logic. In fields such as Artificial Intelligence (AI), decision-making, and legal reasoning, agents must navigate uncertain, dynamic environments. A formal system that incorporates context-sensitive reasoning will allow AI systems to make more accurate decisions by modeling how beliefs evolve in response to changing environments. Similarly, in legal reasoning, where the context of a case often shapes how the law is applied, a more flexible model of Belief Revision is essential.

Moreover, this work contributes to the philosophy of language and epistemology by providing a clearer understanding of how self-locating beliefs are formed and revised. It allows for a more detailed exploration of how humans navigate the world and adjust their beliefs as their context changes. Thus, let us focus on the main questions underlying this investigation:

1.4.5 Main Questions of the Thesis

The primary goal of this thesis is to address the complex relationship between context-sensitive reasoning, Belief Revision, and formal systems. The Wilderness scenario, which serves as a guiding example throughout the thesis, encapsulates the central challenges that the research seeks to resolve. These challenges concern how beliefs, particularly those that depend on shifting contexts, can be rationally formed, revised, and maintained in the face of uncertainty. The key questions that drive this thesis are as follows:

1. How do Indexicals influence belief formation and revision?

A central issue explored in this thesis is the role of Indexicals expressions such as *"I," "here,"* and *"now"* in belief formation and revision. Indexicals are context-sensitive terms whose meaning shifts depending on the situation in which they are used. These terms are highly dependent on the context of the speaker, the time, and the location of the utterance. For instance, when an agent in the Wilderness scenario says *"I am safe here now"*, the meaning of *"I"* refers specifically to the agent, *"here"* refers to the location the agent is in at that moment, and *"now"* refers to the immediate temporal context. However, as time passes, or as new information becomes available such as the discovery of a predator nearby these beliefs must be revised. For example, the belief *"I am safe here now"* might shift to *"I am no longer safe here now"* if the agent realizes they are in danger. This revision is based on new information that changes the agent's context and the interpretation of their belief.

In this thesis, we investigate how beliefs formed through Indexicals are influenced by such changing contexts. The key challenge lies in understanding how agents maintain rational consistency when their beliefs are tied to shifting references. In the Wilderness example, the belief *"I am safe here now"* is intimately connected to the context the agent's position in the wilderness, the knowledge they have, and their temporal awareness. When one of these contextual factors changes such as realizing that the predator's presence alters their safety the belief must be revised. This raises important questions: How do agents rationally update their beliefs when the very references in those beliefs change over time? How do they revise beliefs when new contextual information alters their interpretation? For instance, in the Wilderness, when the agent is presented with evidence (a predator), does their belief about their safety immediately change, or is there a process through which the belief evolves over time? These questions are central to understanding the role Indexicals play in belief formation and revision.

To understand this dynamic, it is crucial to develop a formal system that can address how belief changes based on the evolving context. The use of Indexicals complicates Belief Revision because the meaning of terms changes as the context shifts. The challenge here is to formalize the process by which beliefs about self-location, time, and space beliefs that depend heavily on the context are revised. How can logic accommodate these changes in meaning that occur when the references shift? In the case of the Wilderness agent, this would involve creating a framework that can track how the meaning of *"I," "here,"* and *"now"* evolve with new information, and how these shifting meanings lead to the revision of beliefs. This thesis aims to explore how such formal systems can be constructed to model the context-sensitive nature of Belief Revision effectively, ensuring that these systems remain consistent and rational as the context shifts.

2. What are the limitations of classical systems in handling Indexicals and context-sensitive beliefs?

Classical formal systems, including Propositional Logic and First-Order Logic, assume that beliefs are static, meaning that they can be treated as fixed truth values. In these systems, a belief is either true or false, and the reasoning process is based on this binary framework. While this approach is effective in many contexts, it fails to account for the dynamic nature of beliefs that depend on context. The Wilderness scenario provides a vivid illustration of this limitation: when the agent in the wilderness believes *"I will be rescued in 100 days"*, this belief is highly dependent on the agent's current understanding of their situation, their temporal knowledge, and the surrounding environment. As new evidence emerges such as the appearance of a dangerous predator or the realization that the rescue operation is delayed the belief must evolve. However, classical formal systems struggle to accommodate such changes, as they are not equipped to handle beliefs that change with context.

For instance, in the Wilderness, the belief *"I will be rescued in 100 days"* initially holds as true, given the agent's current knowledge and expectations. But when the agent encounters new information, such as hearing that the rescue team is delayed or that they are in a more dangerous position than previously thought, the belief must be revised. This shift in belief moving from *"I will be rescued in 100 days"* to something like *"I might not be rescued in 100 days"* cannot be captured effectively by classical systems, which operate under the assumption that beliefs are static and fixed.

This limitation of classical systems in handling Indexicals and context-sensitive beliefs is a central issue explored in this thesis. Indexicals, such as *"I," "here,"* and *"now"*, introduce complexity because their meaning depends on the context in which they are used. In the Wilderness, the meaning of *"I"* changes as the agent's position or identity shifts, and the meaning of *"here"* and *"now"* changes depending on the agent's location and the moment in time. Classical systems, like Propositional Logic and First-Order Logic, assume fixed meanings for propositions, which makes it difficult to incorporate the context sensitivity of terms like Indexicals.

The challenge lies in how to extend or modify these classical systems to handle beliefs that are inherently dynamic and context-dependent. Classical systems treat propositions as having fixed truth values, but beliefs involving Indexicals beliefs like *"I am safe here now"* or *"I will be rescued in 100 days"* are highly context-dependent. These beliefs change as new information is encountered, and their meaning is influenced by the agent's temporal and spatial position. The question is how to develop a formal system that accounts for this fluidity and captures the way beliefs evolve as the agent's context shifts.

This thesis aims to investigate these limitations of classical Belief Revision systems and explore how they can be extended to handle context-sensitive beliefs. In particular, the research focuses on how to formalize Belief Revision in a way that maintains rational consistency despite the changing contexts. How can we ensure that beliefs revised in light of new, context-shifting information remain logically coherent? Classical systems do not address this challenge directly, and thus we need to explore new frameworks that can integrate the contextual dynamics inherent in everyday reasoning.

By addressing this question, the thesis seeks to highlight why classical systems fall short when applied to dynamic, context-sensitive scenarios. It aims to lay the groundwork for proposing formal alternatives that can better handle the complexities of Belief Revision in uncertain, evolving environments. Through this exploration, we will seek a more adequate representation of beliefs that change over time and in response to the agent's shifting context, bridging the gap between formal reasoning and everyday applications.

1.4.6 Shifting Sands: Context and the Dynamics of Belief

This thesis addresses several key issues related to Belief Revision, Indexicals, and context-sensitivity, proposing new frameworks for understanding the evolving nature of belief in everyday reasoning. The following core theses guide this work, each contributing to the development of a more comprehensive system for Belief Revision that takes into account the fluidity and context-sensitivity inherent in human cognition and reasoning.

1. Indexicals are Core to Rational Reasoning

One of the fundamental theses of this thesis is that Indexicals expressions such as *"I," "here,"* and *"now"* are essential for rational reasoning. These terms are not just linguistic conveniences but are integral to how we form and revise beliefs. Indexicals are inherently context-sensitive, meaning their meaning shifts depending on the speaker's identity, the location, and the time of utterance. In everyday life, Indexicals are critical tools for making sense of the world, especially when we are navigating uncertain or changing situations.

For example, in the Wilderness scenario, the agent's belief *"I am safe here now"* is formed based on their immediate context: their location, their knowledge of the environment, and their temporal awareness. Over time, the agent may encounter new information that challenges their previous understanding, such as the discovery of a predator or a sudden change in weather, requiring a revision of their belief. This is a quintessential example of how Indexicals shape belief formation and revision: the reference of *"I," "here,"* and *"now"* changes as the context changes, directly influencing the agent's beliefs.

This thesis argues that Indexicals are not simply linguistic tools but are foundational to how rational agents reason and make decisions in dynamic, uncertain environments. Therefore, any formal system that aims to model Belief Revision must account for the shifting references of these terms as part of the revision process. This is a new contribution to the field, as it emphasizes the necessity of incorporating context-sensitivity directly into the logic of Belief Revision.

2. Classical Formal Systems are Insufficient for Context-Sensitive Beliefs

A critical thesis of this work is that classical formal systems, such as Propositional Logic and First-Order Logic, are fundamentally insufficient for modeling context-sensitive beliefs. Classical systems assume that beliefs can be treated as static, with fixed truth values. While these systems work well for certain types of reasoning, they fail to accommodate the fluid nature of beliefs that depend on changing contexts, particularly those that involve Indexicals.

For example, in the Wilderness scenario, the agent's belief *"I will be rescued in 100 days"* is tied to the agent's knowledge of their situation, including their knowledge about the rescue operation and the current context. As new information arises, such as the discovery of a delay in the rescue mission, the agent must revise their belief. However, in classical systems, where beliefs are fixed and absolute, such changes cannot be adequately captured.

This thesis shows that classical Belief Revision frameworks based on fixed truth values cannot handle beliefs that are inherently dynamic and context-dependent. Beliefs formed through Indexicals do not remain constant; they change as the context shifts. Thus, the thesis explores how classical systems can be extended or modified to better represent the evolving nature of beliefs, introducing the need for more expressive formal systems that can accommodate beliefs that depend on the agent's K-context (knowledge context) and S-context (self-context). The introduction of these concepts is a new contribution of this work, bridging the gap between formal logic and the lived experience of belief formation and revision in uncertain environments.

3. Two-Dimensional Semantics (2D Semantics) is Necessary for Representing Context-Sensitive Beliefs

The third thesis defends the idea that Two-Dimensional Semantics (2D Semantics) is an essential extension to formal systems to properly handle context-sensitive beliefs. 2D Semantics introduces two dimensions for evaluating the truth of propositions: the actual world, where the agent's beliefs hold, and a counterfactual or hypothetical world, where the beliefs are evaluated based on potential scenarios. This approach allows for the formalization of beliefs that depend on the agent's context, including Indexicals.

In the Wilderness scenario, the belief *"I will be rescued in 100 days"* initially holds true based on the agent's context both their knowledge (what they believe about the rescue) and their temporal context (the 100-day timeframe). However, as the context shifts say, the rescue is delayed the belief must be revised. The 2D framework allows us to formalize this process by using two dimensions to represent K-contexts (the agent's knowledge at a particular moment) and S-contexts (the agent's self-locating beliefs). This provides a more accurate model of how beliefs change over time and in response to new contextual information.

This thesis proposes that 2D Semantics is a necessary tool for understanding how beliefs evolve in dynamic environments. It allows us to model beliefs that are context-dependent, particularly beliefs expressed through Indexicals, in a way that remains logically consistent even as the context changes. The use of 2D Semantics represents a key contribution of this thesis to the literature on Belief Revision, as it extends classical logic to accommodate context-sensitive reasoning.

4. A New Framework for Belief Revision is Required to Accommodate Dynamic Contexts

Finally, this thesis argues that a new framework for Belief Revision is required to integrate context sensitivity and defeasible reasoning into Belief Revision processes. Classical Belief Revision models often fail to account for the fluid nature of belief formation, especially when beliefs are based on Indexicals and change with the agent's evolving context. The challenge is to create a formal system that accommodates the inherent fluidity of beliefs while maintaining rational consistency.

One of the key contributions of this thesis is the introduction of a framework that combines defeasible reasoning reasoning that allows for beliefs to be withdrawn or revised in light of new, contradictory evidence with context-sensitive beliefs. This framework incorporates the agent's K-context and S-context, allowing beliefs to evolve in a rational and consistent manner. The idea is to create a system that can handle the uncertainties and changing contexts of everyday reasoning, while maintaining logical consistency and coherence.

This thesis explores how defeasible reasoning can be integrated with 2D Semantics to create a robust framework for Belief Revision. This new framework will be able to account for the dynamic nature of belief in uncertain environments, providing a more accurate model of how beliefs are formed, revised, and maintained in the face of shifting contexts and new information.

These core theses serve as the foundation for this work, offering both a critique of existing formal systems and a proposal for a new framework that integrates the complexities of context-sensitive reasoning. In the following sections, we will unpack these theses in greater detail, exploring the underlying philosophical concepts, formal systems, and the new framework for Belief Revision that this thesis proposes.

1.5 Thesis Structure

The structure of the thesis goes as follows. These initial remarks intend to crave the ground from which we take our definitions. We will divide this thesis into four main parts:

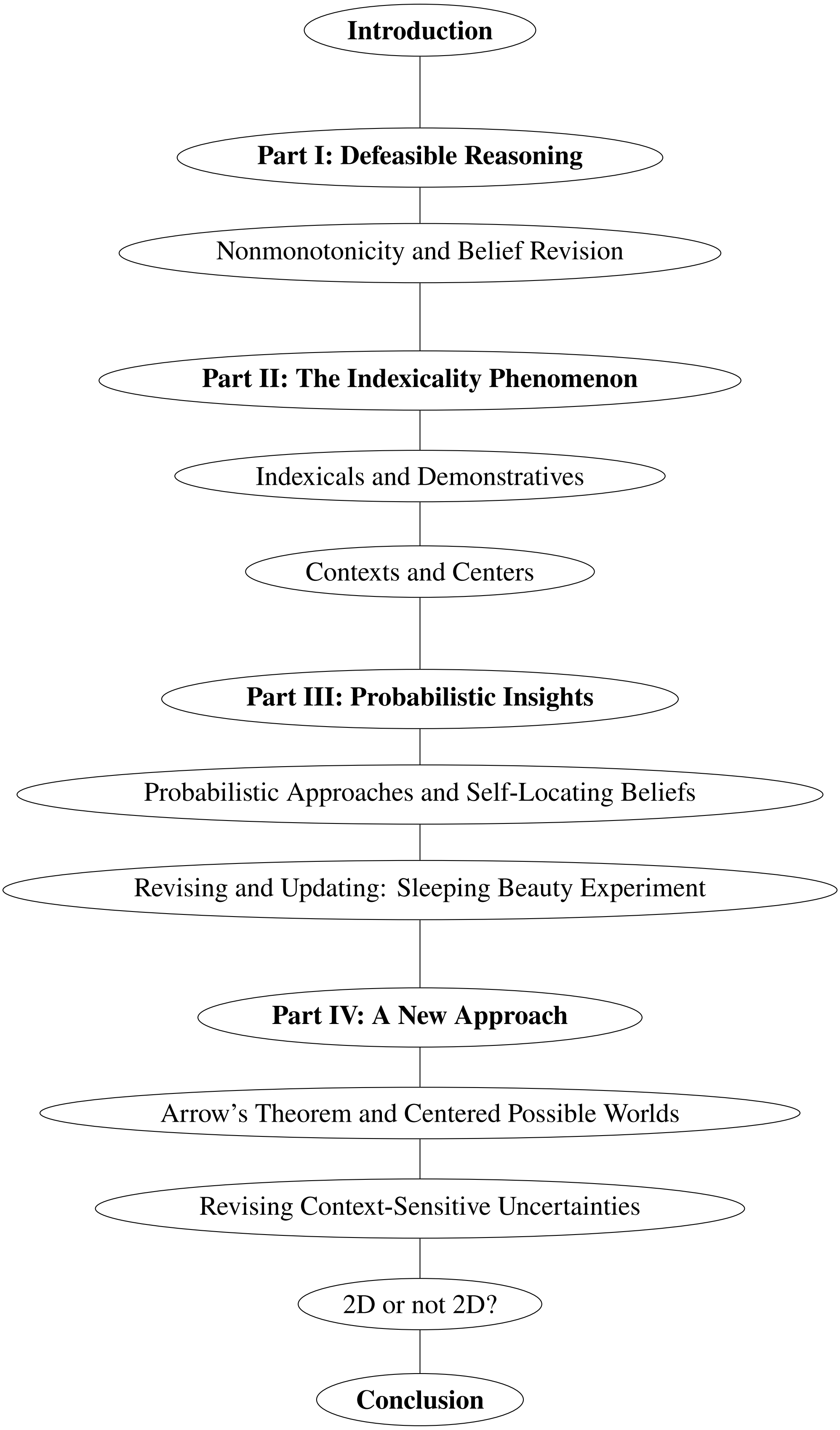
In *Part I* we will investigate formal treatments to defeasible reasoning. Furthermore, in *Chapter 2*, we review the Nonmonotonic Logic and Belief Revision literature. We start with Nonmonotonic systems and their close relationship with Belief Revision. This review intends to provide information about their connections from which we can revise our beliefs about defeasible reasoning. The main goal is to see if any already proposed formal systems are enough to deal with context-sensitive claims. In *Chapter 3* we will investigate the Sleeping Beauty Experiment in more detail and how Updating and Revising can lead to different results. This investigation can help us understand better why Belief Revision might be the best option for our purposes.

To make sense of the challenges indexical expressions give to the formal systems, in *Part II* we will review and investigate the available theories on Indexicals and Contexts. In *Chapter 4*, we explore Kaplan's literature about Indexicals and his proposed *Logic of Demonstratives*, as well as other suggested formal systems. Once more, we aim to use this literature review to investigate if either of the proposed formal systems can deal with defeasible reasoning. In *Chapter 5*, we will focus on the notion of *Contexts*. Because there are several possible definitions of what a *context* is, we will review the possible definitions, define different kinds of contexts, and investigate which ones are more aligned with our purposes. Moreover, we discuss the distinction between Kaplan's and Stalnaker's accounts on *contexts* and if they coexist within the same framework. Considering the notes on contextual features, we will present the *centered world's approach* as a possible alternative representation for indexical statements.

In *Part III, Chapter 6*, we will review how probabilistic approaches deal with self-locating beliefs to explore the quantitative alternatives for revising with context-sensitive information. We can develop analogous definitions for the qualitative case using some of the valuable insights. However, the two-dimensional aspect of indexical Language yields a non-unique solution to examples such as the Sleeping Beauty Experiment.

In *Part IV* we aim to construct Belief Revision Operators for Indexical Information, and, for that, we have to make choices. Thus, we need to rationally aggregate individual orderings, which come with challenges also faced by Social Choice Theory. In *Chapter 7*, we investigate Arrow's Impossibility Theorem and some desirable conditions for that aggregation. In *Chapter 8*, we will look at some examples to apply what we have learned so far. We will use graphic representations to guide the resolution of the examples intuitively. In *Chapter 9*, we will investigate the necessity of two-dimensional approaches and possible strategies to guide us.

Finally, in conclusion, we will summarize the whole story and present some possible further interesting applications. The literature review follows our attempt to find a proper way of dealing with all the examples. We found out that the proposed formal systems are inadequate to deal with different kinds of uncertainty when context-sensitive claims play a part. In order to make the flow of the thesis and literature review, this thesis is structured as follows:


Introduction
Part I: Defeasible Reasoning
Nonmonotonicity and Belief Revision
Part II: The Indexicality Phenomenon
Indexicals and Demonstratives
Contexts and Centers
Part III: Probabilistic Insights
Probabilistic Approaches and Self-Locating Beliefs
Revising and Updating: Sleeping Beauty Experiment
Part IV: A New Approach
Arrow's Theorem and Centered Possible Worlds
Revising Context-Sensitive Uncertainties
2D or not 2D?
Conclusion

Part I

Defeasible Reasoning

Chapter 2

Nonmonotonicity and Belief Revision

We begin this chapter by acknowledging a fundamental challenge to classical logic: its assumptions about the static nature of beliefs. Classical logics, particularly Propositional Logic and First-Order Logic, are based on the idea that once a belief is formed, it remains fixed, and new information cannot alter the set of beliefs. This idea is rooted in the *monotonicity* principle, which states that adding new information to a belief set Γ will not reduce the set's content or change its conclusions. While this principle works well in many contexts, it fails in the face of conflicting or uncertain information. In the real world, we frequently encounter situations where beliefs must be revised in light of new evidence, and this is where monotonicity falls short.

Consider the example of a detective, Jane, investigating a crime in a forest reservation. She must gather information from various sources and form conclusions based on what she learns. As Jane uncovers new details, some of them may contradict her prior conclusions. For instance, she might initially believe that a certain suspect was present at the crime scene based on initial evidence. However, later findings could suggest that the suspect was elsewhere at the time of the crime. Jane's beliefs must be revised in response to this new, conflicting information. This is a typical example of nonmonotonic reasoning, where new evidence forces the revision of previous beliefs, and the set of conclusions changes.

In the case of classical systems, such revisions would be difficult to formalize. These systems are designed to maintain logical consistency by adding new facts without altering old beliefs, but they cannot handle the inherent uncertainty and conflict that arises when beliefs are revised. This thesis aims to explore how nonmonotonic systems address this challenge by allowing for belief changes and updates that account for new, conflicting information.

2.1 Classical Systems and the Monotonicity Challenge

Monotonic reasoning assumes that once something is true, it remains true regardless of what new information is added. This static view of Belief Revision contrasts sharply with the dynamic nature of human reasoning, where beliefs often evolve in response to new, contradictory, or context-sensitive information. Consider Jane, the detective investigating a case in a forest reservation. Each new clue she uncovers might challenge her existing assumptions, forcing her to reassess prior beliefs. She may initially believe a suspect was nowhere near the crime scene, but discovering fresh footprints near the evidence site compels her to revise her belief. Classical logics, such as Propositional Logic and First-Order Logic, lack the flexibility to model this process of Belief Revision in response to conflicting information.

This chapter investigates nonmonotonic reasoning and Belief Revision, exploring how these approaches overcome the limitations of classical systems. Nonmonotonic reasoning allows conclusions to be retracted or adjusted when new evidence contradicts previous assumptions. This adaptability is crucial for modeling everyday reasoning, where beliefs are not always static, and where context and new information significantly influence our epistemic landscape. By incorporating nonmonotonic reasoning, we can better represent the way agents dynamically update their beliefs, accommodating uncertainty, context shifts, and evolving knowledge structures.

2.2 Defeasible Reasoning and Nonmonotonicity

Nonmonotonic reasoning, also known as defeasible reasoning, allows agents to draw conclusions based on incomplete or tentative information, while remaining open to revising those conclusions as new evidence becomes available. This kind of reasoning is particularly important when dealing with generalizations rules or beliefs that hold in most cases but not necessarily in all. For example, we might believe that birds can fly in general, but we know that certain birds, such as penguins, do not fly. This generalization is defeasible because the conclusion (that birds fly) can be retracted in light of new information (the case of penguins). In Jane's investigation, for instance, she may believe that the suspect was at the scene based on witness testimony, but later find out that the testimony was unreliable, necessitating a revision of her belief.

The concept of defeasible reasoning has been widely studied in both artificial intelligence and philosophy. In AI, defeasible reasoning is often used to model reasoning in uncertain environments, where conclusions can be withdrawn as more information becomes available. Philosophers, however, are also deeply interested in defeasible reasoning because it closely mirrors how humans revise their beliefs in response to new evidence. As Jane's investigation unfolds, she must be able to revise her conclusions in light of new, possibly conflicting information. The process of Belief Revision is essential for ensuring that reasoning remains rational and consistent, even as the context evolves.

There are two main approaches to defeasible reasoning: an epistemological approach and a logical approach. Both approaches are connected because defeasible reasoning has epistemological implications it involves making inferences based on uncertain or incomplete knowledge. In the epistemological approach, defeasible reasoning is concerned with how agents handle beliefs in the face of uncertainty. In the logical approach, the focus is on developing formal systems that capture how agents revise their beliefs logically. Both approaches are relevant to this thesis, as we aim to bridge the gap between the philosophical exploration of defeasible reasoning and its formalization in logic.

> *"Defeasible reasoning reflects the adaptive nature of human cognition, where conclusions are revised or retracted based on new information that may contradict previous assumptions."* [4, p. 27]

The flexibility of defeasible reasoning is crucial for handling the kind of uncertainty that arises in situations like Jane's investigation, where new evidence may force a reconsideration of earlier conclusions. Unlike classical logics, which maintain a rigid structure of fixed beliefs, nonmonotonic reasoning allows for the dynamic revision of beliefs based on new evidence, making it a powerful tool for modeling how beliefs evolve over time.

2.3 Nonmonotonic Systems: A Look at Formal Approaches

Nonmonotonic logics have been developed to model defeasible reasoning and Belief Revision more effectively. These systems are designed to handle cases where the introduction of new information can contradict or qualify previously held beliefs. Classical logic assumes that once a belief is established, it remains true, regardless of new evidence. However, nonmonotonic systems reject this assumption by allowing for beliefs to be revised in light of new, often conflicting, information.

For example, consider the detective Jane in the Wilderness scenario. Jane may initially believe that a suspect is guilty based on early evidence, but as new facts emerge perhaps a new witness statement this belief may need to be revised. Nonmonotonic reasoning allows for the revision of this belief, ensuring that Jane's conclusions evolve as new evidence is acquired. In this scenario, classical logic would fail to accommodate the changing nature of Jane's beliefs because it assumes that once something is believed, it is fixed.

Nonmonotonic reasoning is not merely about rejecting the monotonicity principle. It is a family of logics that allow for conclusions to be revised or withdrawn when new evidence arises, a property that is essential when dealing with complex reasoning processes like Belief Revision. As Makinson [64, p. 27] notes, there are various nonmonotonic consequence relations, and these logics extend beyond just the rejection of monotonicity. The family includes systems that allow for conclusions to shift dynamically in response to new, sometimes conflicting, information.

One of the primary contributions of nonmonotonic systems is their ability to incorporate contextual features that is, how beliefs change depending on the agent's knowledge and the context in which the information is gathered. In the Wilderness scenario, Jane's beliefs are not static but constantly shift as she uncovers new clues or faces new challenges. Classical systems are inadequate because they assume that beliefs remain unchanged when new information is introduced. Nonmonotonic logics, on the other hand, offer a more nuanced approach, where beliefs are not fixed but are subject to change based on evolving contexts.

This thesis will explore how different formal approaches to nonmonotonic reasoning can capture the complexities of Belief Revision. By examining these systems, we aim to determine if they offer a more accurate model for Belief Revision than classical logics, particularly in everyday situations, like Jane's investigation, where new, conflicting information regularly forces agents to revise their beliefs.

2.4 Reiter's Default Logic

One formal approach to defeasible reasoning is Reiter's Default Logic [88, p. 34], which provides a framework for handling nonmonotonic reasoning in uncertain environments. The core idea behind Reiter's Default Logic is the Default Rule, an inference rule that allows agents to make conclusions based on available information, even when that information is incomplete or potentially contradictory. Informally, a default rule states that "if a condition γ holds and there is no evidence to the contrary, then it is reasonable to assume τ." For example, Jane might initially assume that the suspect was at the crime scene, but as new evidence emerges, she may need to revise her belief.

> *"Default reasoning models the human tendency to make conclusions based on incomplete knowledge, but it must be capable of revising those conclusions when new, contradicting information becomes available."* [88, p. 78]

In Reiter's framework, the credibility and consistency of the premises play crucial roles. Nonmonotonic reasoning, like the reasoning Jane uses in her investigation, often arises when the premises fail to be fully credible or consistent. For instance, Jane's initial belief about the suspect might be based on unreliable testimony or incomplete evidence, and as new facts emerge, her belief must be adjusted. The consistency of the premises, and the context in which those premises are evaluated, is essential in determining which conclusions are valid. in the framework of Default Logic, context helps define the set of premises that can be assumed true and the inferences that can be drawn from those premises.

However, Reiter's Default Logic also has limitations. For example, it does not always guarantee that a default extension exists for every theory, and there are situations where the default consequent may contradict the justification, which the system cannot resolve. This raises important questions about the consistency of beliefs in nonmonotonic reasoning. What does it mean for a proposition to be consistent in the framework of defeasible reasoning? As Reiter [88, p. 45] points out, one of the most challenging tasks in developing Default Logic is providing a clear, formal definition of consistency in justifications.

In everyday reasoning, especially in cases like Jane's investigation, agents may hold conflicting beliefs at the same time. Jane might believe both that the suspect was present at the scene and that they were not, depending on the available evidence. This conflicting information must be carefully managed within the logic. Another approach to resolving conflicts in default reasoning is found in McDermott's nonmonotonic modal system [65, p. 60]. McDermott introduces an operator M, which denotes the majority of evidence or reasoning in favor of a particular belief. This approach helps reduce conflicts by prioritizing the majority of evidence, thus offering a more robust way of managing Belief Revision when new, conflicting information arises.

> *"In nonmonotonic modal systems, conflicts are resolved by considering the majority of evidence, thus providing a more adaptive approach to defeasible reasoning."* [65, p. 112]

Ultimately, the introduction of a majority operator like M helps avoid contradictions when new information is integrated, making this a useful addition to the framework of nonmonotonic reasoning. By incorporating such mechanisms, we can model how agents like Jane adapt to new, conflicting evidence while maintaining rational consistency in their reasoning. This is crucial for ensuring that the Belief Revision process remains coherent and aligned with the agent's evolving understanding of the situation.

Thus, Reiter's Default Logic and the nonmonotonic modal systems contribute important tools for addressing the challenges of Belief Revision in uncertain and dynamic environments. These systems provide a formal means of revising beliefs in response to new, conflicting, or incomplete information, making them valuable for modeling everyday reasoning, such as the complex decisions faced by a detective like Jane in the Wilderness.

2.5 Modal Nonmonotonic Logics

Nonmonotonic logics have been developed to address the challenges of defeasible reasoning and Belief Revision, offering a more flexible framework than classical logics for situations where new, conflicting information necessitates Belief Revision. Classical logics, such as Propositional Logic and First-Order Logic, are based on the assumption that once a belief is established, it remains fixed, regardless of new evidence. This principle of monotonicity states that adding information to a belief set will not change the conclusions drawn from it. However, this assumption fails in everyday reasoning, where beliefs are often revised in light of new, contradicting evidence.

For instance, in the Wilderness example, Jane, the detective investigating a crime in a forest, is presented with new pieces of evidence that may contradict her prior conclusions. Initially, she may believe that a suspect was at the scene of the crime, based on a witness statement. However, as Jane gathers more information, this belief may need to be revised, especially if new evidence (e.g., a different witness statement or an alibi) contradicts the initial one. This is an example of how beliefs are defeasible subject to retraction or modification as new information becomes available.

The modal systems used to model belief in classical logic, particularly Epistemic Modal Logic, incorporate modal operators such as B (for belief) and K (for knowledge) to formalize concepts of belief and knowledge in relation to possible worlds [20]. However, these systems fall short when attempting to capture the dynamics of belief that is, how beliefs change in response to new evidence, especially in complex, uncertain environments. Classical Epistemic Modal Logic can formalize static beliefs, but it does not provide an adequate framework for handling the revision of beliefs in light of conflicting evidence or context-sensitive information.

To address these limitations, several nonmonotonic modal logics have been proposed. One of the key approaches in this area is the Modal Nonmonotonic Logic proposed by McDermott [65], which seeks to model Belief Revision in a way that accommodates defeasible reasoning reasoning that allows for beliefs to be withdrawn or modified when new, conflicting information arises. The central assumptions of

McDermott's Modal Nonmonotonic Logic are as follows:

- (A1): The agent has full introspective powers; that is, if an agent believes a proposition φ, then the agent believes in $\diamond\varphi$ (i.e., it is possible that φ is true), and if φ is not believed, then the agent believes in $\neg\diamond\varphi$ (i.e., φ is not possible).
- (A2): The agent reasons from her initial set of beliefs, which includes her assumptions, previously accepted beliefs, and what is consistent to assume given them.
- (A3): The agent assigns belief sets only to theories that are consistent with the underlying formal system she is considering.

These assumptions help formalize the agent's reasoning process, allowing her to revise her beliefs dynamically in response to new information. The concept of context is crucial in Modal Nonmonotonic Logics, but it is framed differently than in the later discussions on indexical beliefs. In McDermott's approach, contexts are treated as theories T_c that serve as a basis for checking the consistency of reasoning. The theory T_c dictates the set of formulas the agent considers when making inferences.

In this framework, the notion of a context-dependent proof emerges: a proof is considered context-dependent if it is derived with respect to a given theory T_c. This means that an agent's reasoning process is context-sensitive, and the inferences she makes depend on the theory she adopts at a particular moment. In McDermott's formalism, context-dependent proofs are used to derive the formulas that an agent should believe, based on the theory she considers valid in a given context.

The idea of expansion in Modal Nonmonotonic Logics parallels the concept of Belief Revision in the AGM theory of Belief Revision. In both approaches, beliefs can be expanded with new, consistent information, provided that the new information does not contradict previous beliefs. However, the key distinction is that Modal Nonmonotonic Logics focus on how beliefs evolve when the agent encounters new, potentially contradictory evidence. When beliefs are expanded, the new set of beliefs is consistent with the agent's prior beliefs, but this is not always the case in defeasible reasoning, where new information may lead to retraction of previous conclusions.

Despite these advancements, McDermott's Modal Nonmonotonic Logic has faced criticism. While it captures some of the dynamics of Belief Revision, it struggles to adequately represent the context-sensitivity required for beliefs involving indexicals. For instance, in the Wilderness example, Jane's belief about the suspect's guilt might change not only based on new evidence but also depending on contextual features such as her location or the time at which new information is discovered.

McDermott's logic does not provide a sufficient formalization of the contextual factors that affect Belief Revision when indexical expressions like "I", "here", and "now" are involved. This is a significant limitation for modeling everyday Belief Revision processes, where beliefs often depend on contextual parameters such as spatial and temporal factors.

An alternative approach to modeling defeasible reasoning in a more context-sensitive manner is the Autoepistemic Logic [69]. Unlike Modal Nonmonotonic Logics, which focus on beliefs about the world, Autoepistemic Logic allows the agent to reason about her own beliefs what she knows and what she does not know. It provides a way for the agent to introspectively reason about her own belief set, incorporating knowledge about knowledge (i.e., epistemic reasoning) into Belief Revision. However, this system has been critiqued for its normative nature: it assumes that the agent is ideal, rational, and omniscient, which is unrealistic for modeling natural agents' behaviors. While Autoepistemic Logic offers valuable insights, it is not well-suited for modeling Belief Revision in agents who must operate with uncertain, incomplete, or conflicting information.

> *"To characterize the status of presumptions, we need a general framework for representing the dynamics of acceptance the policies that an agent has for changing what it accepts in response to new information, including information that conflicts with what it accepts."*
> [100, p. 192]

Ultimately, nonmonotonic modal logics including McDermott's and Autoepistemic Logic offer important tools for understanding Belief Revision in dynamic contexts. These systems show how beliefs can evolve over time, especially when agents face conflicting evidence or must reason about their own beliefs. However, they are still limited in fully addressing the context-sensitivity that is required when beliefs depend on indexicals and self-locating information. This thesis will explore these limitations and propose new frameworks for Belief Revision that can accommodate the shifting, context-dependent nature of everyday reasoning.

In addition, an alternative approach worth considering is the Adaptive Logics framework, which uses adaptive strategies to handle inconsistencies within a belief set [10]. These strategies allow agents to reason about their beliefs while taking into account the dynamism and fluidity of reasoning in everyday contexts. Although these approaches are less developed than the nonmonotonic modal logics, they provide promising avenues for future research on Belief Revision in uncertain environments.

2.6 Autoepistemic Logic

Default reasoning and autoepistemic reasoning are distinct notions, though both deal with revising beliefs in light of new evidence. Default reasoning typically deals with incomplete evidence, while autoepistemic reasoning is concerned with how an agent's beliefs relate to one another, particularly in terms of self-reflection and introspection. Consider the following reasoning principle:

> Suppose $\varphi_1, \varphi_2, \ldots$ are instances of φ. We have the following reasoning principle:
>
> If φ_i were true, I would know φ_i; since I do not know φ_i, it cannot be accurate. [53, p. 136]

In Autoepistemic Logic, an agent is assumed to know the propositions that are true within their belief set. Thus, the focus is on the agent's introspective abilities to reflect on their own beliefs. If an agent does not know something, it implies that the belief does not hold in the current context. Autoepistemic reasoning cannot be considered defeasible in the same way as default reasoning because it assumes that all occurrences are known by the agent. The agent's belief set is static in that it only contains what is known or logically implied by the agent's introspection. However, it is nonmonotonic because beliefs can change or be retracted in light of new evidence, depending on the context of the reasoning process.

In contrast, Default reasoning deals with incomplete or insufficient information and allows for the retraction or modification of conclusions as new evidence emerges. For instance, in the Wilderness example, Jane, the detective investigating a crime in a forest, may initially believe that a suspect is guilty based on available evidence. However, as new facts come to light, her belief about the suspect's guilt may change. Default reasoning provides a way to accommodate these changes, whereas autoepistemic logic focuses more on the introspective consistency of Jane's own beliefs, given the evidence she already holds.

Let's examine the formalism of Autoepistemic Logic in more detail. Let $\mathcal{L}_{\dashv\updownarrow}$ be a language with modal operators. An autoepistemic theory $T_{al} \subseteq \mathcal{L}_{\dashv\updownarrow}$ is a set of formulas representing the agent's belief set about her own beliefs. This theory has two main properties: stability and groundness [53].

Definition 2.6.1 *An autoepistemic theory $T_{al} \subseteq \mathcal{L}_{\dashv\updownarrow}$ is said to be* ***stable*** *if it satisfies the following conditions:*

- *(1) If $\alpha_1, \ldots, \alpha_n \in T_{al}$ and $\alpha_1, \ldots, \alpha_n \vdash \beta$, then $\beta \in T_{al}$.*
- *(2) If $\alpha \in T_{al}$, then $\Box\alpha \in T_{al}$.*

- *(3) If $\alpha \notin T_{al}$, then $\neg\Box\alpha \in T_{al}$.*

The stability property characterizes the consistent behavior of belief set elements. It establishes which formulas must co-occur within a belief set to maintain internal coherence. However, stability is limited in that it cannot determine which formulas should be excluded from the belief set. This is where the second property, groundedness, comes into play.

Definition 2.6.2 *An autoepistemic theory $T_{al} \subseteq \mathcal{L}_{\dashv\updownarrow}$ is **grounded** in a set of premises $S \subseteq \mathcal{L}_{\dashv\updownarrow}$ if every formula in T_{al} is included in the tautological consequences of $A \cup \{\Box\alpha : \alpha \in T_{al}\} \cup \{\neg\Box\alpha : \alpha \notin T_{al}\}$.*

Groundedness is an important property that is often used in relevant logics to determine the set of premises that ground possible rational conclusions. It helps ensure that the belief set is consistent with the initial assumptions, providing a way to define what must be included in the set of beliefs. While stability tells us which formulas need to be in a belief set, groundedness defines which formulas should not belong to the belief set. Together, these properties aim to describe a rational agent's Belief Revision process.

Definition 2.6.3 *An autoepistemic theory $T \subseteq L$ is a stable expansion of a set of premises $A \subseteq L$ if it satisfies the following conditions:*

- *(1) $A \subseteq T$,*
- *(2) T is stable,*
- *(3) T is grounded in A.*

A stable expansion characterizes the belief set of an ideally rational agent concerning a given premise set. It includes the premises and ensures that the belief set is both stable and grounded in those premises. Thus, an ideal belief set contains all and only those formulas deemed rational, given the premises. Stability determines what formulas need to be included, while groundedness ensures that the belief set is minimal, avoiding unnecessary expansions.

> A stable expansion characterizes the belief set of an ideally rational agent concerning a given premise set. It includes the premises and is stable and grounded in the premises. Thus, an ideal belief set contains all and only those formulas deemed "rational" given the premises. Stability determines what formulas need to be included, and groundedness ensures that the premise set's extension to achieve stability is minimal. [53, p. 140]

As in McDermott's Nonmonotonic Modal Logic [65], Autoepistemic Logic also treats the notion of context as a consistent theory. In this formalism, the evaluation of a belief is done within a context of interpretation [101]. However, in contrast to the way context is used in indexical reasoning, here the context is defined by a set of well-formed formulas, rather than being linked to specific expressions like "I", "here", or "now".

The context of interpretation in Autoepistemic Logic can be understood as a fragment of the set of formulas. Furthermore, T_c admits an expansion operation. If there is a sentence A that is consistent with T_c, then A can be added to T_c without introducing inconsistencies, which mirrors the concept of expansion in Belief Revision.

While Autoepistemic Logic is useful for capturing introspective reasoning, it has limitations in dealing with context-sensitive beliefs particularly when it comes to indexicals. For example, Jane may believe that "I am close to the crime scene" in the Wilderness scenario, but this belief could change if new information about her actual location comes to light. Autoepistemic Logic does not provide a way to directly handle such dynamic, context-dependent beliefs in the way that Nonmonotonic Modal Logics might. Thus, while Autoepistemic Logic offers valuable insights into Belief Revision based on introspection, it is not sufficient to handle belief changes driven by external evidence or context-sensitive information. To fully capture the complexities of everyday Belief Revision, including the challenges posed by indexicals, a more flexible framework is needed one that accommodates the fluidity and dynamic nature of beliefs as they evolve in response to both introspective and external factors.

2.7 Nonmonotonicity, Beliefs, and Changes

A rational process for revising our beliefs is essential because our decisions are often based on the belief set we hold and what we consider true about the world. It is not reasonable to accept any formula as true without thoroughly reflecting on whether it remains consistent with the beliefs we have previously accepted. Consistency is a critical rationality condition; holding contradictory beliefs may lead to undesirable or explosive conclusions.

Learning involves the consistent process of adding or removing information from our belief set. In some cases, this addition or removal can occur without any issues, maintaining the consistency of the belief set. However, when new information introduces potential contradictions, modifications to the belief set may be required. Logic alone cannot determine which formulas should be removed or revised, especially when multiple conflicting alternatives exist.

The topic of rational belief change is precisely what the AGM Postulates [2] aim to address. The AGM framework identifies three fundamental types of belief change: expansion, contraction, and revision [32]. However, some scholars argue that revision and contraction should be treated as two aspects of the same process. According to this view, the three belief change operations would be expansion, revision, and update [48].

Although Belief Revision and Nonmonotonic Reasoning are often thought to be closely related, given their shared focus on defeasible reasoning and uncertainty, their objectives differ significantly. Nonmonotonic reasoning focuses on nonmonotonic inference relations, dealing with situations where conclusions can be retracted based on new information. On the other hand, Belief Revision is concerned with the consistent outcomes of belief changes. As illustrated in the example of Autoepistemic Logic, these two domains address different problems. The relationship between Nonmonotonic Reasoning and Belief Revision can be summarized, prima facie, as follows:

$$A \vdash B \text{ iff } T * A \models B$$

where T represents a classical theory. In this formulation, Belief Revision operations generate nonmonotonic inference operations [48]. While both fields address dynamic reasoning processes, the distinction in focus makes Belief Revision a more general approach for handling the evolution of beliefs in the face of new, potentially contradictory information. Nonmonotonic Reasoning focuses on the conditions under which beliefs can be retracted or revised in a specific logical framework, whereas Belief Revision encompasses the broader task of ensuring that belief sets evolve rationally while maintaining consistency.

For example, in the Wilderness scenario, Jane, as a detective, initially holds the belief that a suspect is guilty. As new evidence is uncovered, Jane must revise her belief. However, this revision is not simply the process of retracting an earlier belief; it involves ensuring that the revised belief set remains consistent, incorporating new evidence while respecting the coherence of the belief system as a whole. In this context, Belief Revision provides the necessary structure to ensure that Jane's beliefs evolve in a rational and consistent manner, despite the conflicting information she encounters.

This distinction highlights the importance of a formal system capable of integrating the dynamics of belief change while maintaining consistency, a task that Belief Revision frameworks are specifically designed to handle.

2.8 AGM Postulates for Belief Revision

Belief Revision is a central concept in nonmonotonic systems, referring to the process of modifying a knowledge base when new information is acquired. The process of Belief Revision involves three main operations: expansion, contraction, and revision. These operations allow us to adjust our belief set in response to new evidence. The concept of Belief Revision plays a crucial role in nonmonotonicity and sheds light on the nature of defeasible reasoning [5]. Roughly speaking, Belief Revision can be broken down as follows:

- **Expansion**: In this operation, new information is added to the belief set without removing any of the previously accepted beliefs. This operation is useful when the new information does not conflict with the existing set.
- **Contraction**: This operation involves eliminating some beliefs from the set. The challenge of contraction lies in determining which information to discard, especially when faced with conflicting beliefs.
- **Revision**: Revision is a more nuanced operation, aiming to incorporate new information by adjusting the existing belief set as little as possible to preserve consistency. In this case, new information is integrated while maintaining the coherence of the belief set.

These three operations allow an agent to update their beliefs in response to new, potentially contradictory evidence. For example, in the Wilderness scenario, Jane, the detective investigating a crime, may initially believe a suspect is guilty based on evidence. As new, contradictory evidence emerges, Jane must revise her belief to account for this new information, updating her knowledge base to maintain consistency.

There are various interpretations of what constitutes a belief set or current knowledge base, which is an important epistemological issue. In nonmonotonic logics, belief sets are typically interpreted as rational, coherent propositions that represent information about the world (e.g., beliefs, opinions, or facts) [18]. Adaptive Logics use similar interpretations for the set of premises. However, the interpretation of a current knowledge base as a set of justified true beliefs can be problematic. Instead, we might accept the notion of a current knowledge base as a set of defeasible propositions expressing information that is subject to revision.

2.8.1 Revision Postulates

Let ψ and μ represent propositions in the language of the agent, and let $\psi \circ \mu$ denote the revision of ψ by μ. The AGM Postulates for Belief Revision are as follows:

- **(R1)** $\psi \circ \mu$ implies μ: The new information μ must be included in the revised belief set.

- **(R2)** If $\psi \wedge \mu$ is satisfiable, then $\psi \circ \mu \leftrightarrow \psi \wedge \mu$: If the conjunction of the old belief ψ and the new information μ is satisfiable, the revised belief set is equivalent to the conjunction of the old and new beliefs.

- **(R3)** If μ is satisfiable, then $\psi \circ \mu$ is also satisfiable: The revised belief set must be consistent with μ.

- **(R4)** If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$, then $\psi_1 \circ \mu_1 \leftrightarrow \psi_2 \circ \mu_2$: The revision of equivalent beliefs should lead to equivalent revised belief sets.

- **(R5)** $(\psi \circ \mu) \wedge \varphi$ implies $\psi \circ (\mu \wedge \varphi)$: If the revision of ψ by μ and φ is consistent, then the revision of ψ by the conjunction $\mu \wedge \varphi$ is also consistent.

- **(R6)** If $(\psi \circ \mu) \wedge \varphi$ is satisfiable, then $\psi \circ (\mu \wedge \varphi)$ implies $(\psi \circ \mu) \wedge \varphi$: If the conjunction of the revision and φ is satisfiable, the revised belief set must satisfy both the original revision and φ.

These postulates are designed to formalize the process of revising beliefs when new information is received. However, the challenge lies in how to handle conflicting information and ensure the revised belief set remains consistent.

2.8.2 AGM Axioms for Revision

Katsuno and Mendelzon, drawing inspiration from the work of Alchourrón, Gardenförs, and Makinson, extended the AGM framework to handle Belief Revision formally. Let K and H be belief sets, and let A and B be formulas in the language. The AGM Axioms for Belief Revision are as follows [32]:

- $[K*1]$ K^*_A is a belief set.

- $[K*2]$ $A \in K^*_A$: The belief set after revision contains the new information A.

- $[K*3]$ $K^*_A \subseteq K^+_A$: The revised belief set is a subset of the expanded belief set.

- $[K*4]$ If $\neg A \notin K$, then $K^+_A \subseteq K^*_A$: If the negation of A is not in the belief set, the expanded belief set will be a subset of the revised belief set.

- $[K*5]$ $K^*_A = K_\perp$ iff $\vdash \neg A$: If the revision leads to a contradiction, the revised belief set becomes inconsistent.

- **[$K * 6$]** If $\vdash A \leftrightarrow B$, then $K^*_A = K^*_B$: If A and B are logically equivalent, revising by A and B will yield the same revised belief set.

In addition to these postulates, two further axioms for the revision of belief sets are given:

- **[$K * 7$]** $K^*_{A \wedge B} \subseteq (K^*_A)^+_B$: If the agent revises by $A \wedge B$, the belief set should expand to include the new belief.
- **[$K * 8$]** If $\neg B \notin K*_A$, then $(K^*_A)^+_B \subseteq K^*_{A \wedge B}$: If B is consistent with the belief set, the belief set should reflect this new information.

These axioms serve to formalize the operations of expansion, contraction, and revision as they apply to belief sets. In the case of revision, the axioms guarantee that the belief set is updated in a way that is consistent with the agent's existing knowledge, even in the face of new information.

2.8.3 Challenges in Contraction

Contraction in Belief Revision presents some of the most difficult and contentious challenges in formal epistemology. in the framework of Belief Revision, contraction refers to the process of eliminating some beliefs from an agent's belief set in order to accommodate new information that conflicts with the previously held beliefs. Unlike expansion, where new beliefs are added without removing any existing ones, contraction involves a delicate balancing act. The agent must decide which beliefs to discard to maintain the consistency of the belief set while retaining as much of the original belief set as possible. The challenge lies in determining which beliefs should be removed, especially when multiple conflicting pieces of evidence arise.

At the heart of contraction is the issue of epistemic entrenchment, which refers to the relative strength or degree of belief associated with each proposition in an agent's belief set. Some beliefs are more deeply entrenched than others and thus should be retained when faced with new, contradictory information. Others may be less entrenched, making them more easily revised or discarded. The concept of epistemic entrenchment was highlighted by Rott [89] and has been central to the debate about how beliefs should be contracted in a rational way.

Epistemic Entrenchment and Belief Prioritization

The principle of epistemic entrenchment suggests that when deciding which beliefs to contract, agents should prioritize the beliefs that are most entrenched. Rott [89] defines two versions of the informational economy principle in the framework of belief contraction:

1. Minimal Change: An agent should aim to minimize the change in their belief set when incorporating new information. In other words, the agent should keep as many of their original beliefs as possible.

2. Retreat to Less Entrenched Beliefs: When multiple belief changes are possible, the agent should discard the least entrenched beliefs, that is, those beliefs that are least certain or strongly held.

The first version, minimal change, follows the intuitive idea that agents should make the least disruptive modifications to their beliefs. However, the second version raises an interesting question: how do we determine which beliefs are least entrenched? This becomes even more complicated when we consider indexical beliefs beliefs about self-location or perspective that are inherently context-sensitive and prone to revision based on new, context-shifting information.

Consider, for example, Jane, the detective in the wilderness, who initially believes that a certain trail leads to the suspect's hideout. As she investigates further, she discovers conflicting evidence that places the suspect elsewhere. Jane now faces a dilemma: she must contract some of her beliefs to accommodate the new information. Should she discard her belief that the trail leads to the hideout, or should she revise her belief about the suspect's location? A key issue in belief contraction is the non-uniqueness of the contraction process. In many cases, there are multiple ways to contract a belief set. If new evidence contradicts two or more beliefs in the set, the agent must choose which belief(s) to discard. This raises several philosophical issues related to the rationality of contraction. How does the agent make this choice? What criteria should guide the decision-making process? Should the agent contract the belief that is less supported by evidence, or should they choose to keep the belief that is more strongly held?

This problem of multiple contraction choices has been extensively discussed in the literature. One approach to resolving this issue is to adopt a defeasible reasoning framework, where beliefs are not treated as fixed truth values, but as conclusions that can be withdrawn in the light of new evidence. This approach allows the agent to retract beliefs in a way that respects the context-sensitive nature of reasoning, particularly when the agent's beliefs depend on shifting references or changing perspectives. For example, Jane, upon learning new evidence, may initially have several beliefs that are equally likely to be true, but some of those beliefs will need to be given up to avoid contradiction. The challenge is determining which belief to contract without introducing irrationality or inconsistency in the process. This becomes especially difficult when the beliefs involve indexicals terms like "I," "here," and "now" which change depending on the agent's context. The issue becomes even more pressing when considering beliefs about self-location (de se beliefs), which are highly sensitive to the agent's perspective and may need to be revised when new, context-sensitive information arises.

The issue of contraction in Belief Revision also touches upon broader philosophical debates in epistemology. One key debate is between constructivist and foundationalist views of knowledge and belief. Constructivists argue that beliefs are formed through processes of negotiation and revision, meaning that contraction is an essential part of how we update our beliefs in light of new evidence. Foundationalists, on the other hand, emphasize the idea of epistemic stability, arguing that certain core beliefs should remain unchanged even when new evidence challenges them.

This tension between stability and change is at the heart of the problem of contraction. Foundationalists might argue that beliefs about basic facts such as "The sun rises in the east" should be immune to revision, while constructivists might argue that even seemingly stable beliefs can be revised when new information arises. The challenge for Belief Revision systems, then, is to find a way to balance these two extremes by providing a formal mechanism for belief contraction that maintains rational consistency and logical coherence.

in the framework of nonmonotonic reasoning, belief contraction plays a crucial role in handling conflicting beliefs. As Jane the detective revises her beliefs, she must not only incorporate new evidence but also ensure that her revised belief set remains consistent and coherent. The nonmonotonic nature of Belief Revision means that the agent's belief set is not closed new evidence may invalidate previous beliefs, requiring the agent to contract or revise their beliefs. The key challenge here is determining which beliefs to retain and which to discard, ensuring that the agent's reasoning remains consistent and rational.

Contraction is inherently linked to the broader process of Belief Revision, as it represents one of the three core operations in the revision of belief sets: expansion, contraction, and revision. While expansion involves adding new information to the belief set, contraction requires the removal of inconsistent or less entrenched beliefs. The revision process, on the other hand, seeks to reconcile conflicting beliefs while minimizing the disruption to the overall belief set. Belief Revision systems, such as AGM (Alchourrón, Gardenförs, and Makinson) theory, have been developed to formalize these operations and provide a structured framework for updating beliefs in response to new evidence. However, as we have seen, contraction presents a particular challenge, especially when dealing with context-sensitive beliefs and indexicals. The next section will explore the relationship between nonmonotonic reasoning and Belief Revision, further developing the philosophical and formal tools needed to address these challenges.

The concept of *epistemic entrenchment*, as discussed by Rott [90], plays a crucial role in the Belief Revision process, particularly when dealing with the challenge of contraction. Epistemic entrenchment refers to the relative importance of propositions in an agent's belief set, determining which beliefs are more deeply embedded and which can be revised or discarded when faced with new information. This idea is central to understanding how agents prioritize information when revising their beliefs. Gardenförs presents the concept of epistemic entrenchment as follows:

> The propositions accepted by an agent in a state of belief have different *epistemic entrenchments* for the agent some propositions are more valuable than others in inquiry and decision making. [32, p. 75]

In this context, epistemic entrenchment helps guide the process of belief contraction by identifying which beliefs are more "entrenched" or foundational to the agent's knowledge base. When confronted with new, contradictory information, beliefs with higher epistemic entrenchment are typically preserved, while those with lower entrenchment are more likely to be revised or discarded.

Let us illustrate this with an example. Imagine Jane, our detective, investigating a crime in a forest reservation. She initially believes that the suspect, Mr. Green, is guilty based on a set of circumstantial evidence she has gathered. This belief is entrenched in her mind, as it aligns with her initial assumptions about the case. However, as Jane continues her investigation, she discovers a new piece of evidence a reliable eyewitness report suggesting that Mr. Green was elsewhere at the time of the crime. This new evidence directly contradicts her initial belief.

According to the principle of epistemic entrenchment, Jane must decide which beliefs to revise. Her belief that Mr. Green is guilty is now challenged by the new information. Given the strength of the new evidence, Jane may decide to discard or revise her entrenched belief. However, if her belief in Mr. Green's guilt was deeply rooted in a broader pattern of circumstantial evidence (e.g., fingerprints found at the scene), she might choose to revise only her belief about the alibi rather than discarding the entire theory of guilt.

This decision is guided by the epistemic entrenchment of her beliefs. The more firmly entrenched the belief (i.e., the more supporting evidence she has), the harder it may be for her to revise or discard it. Conversely, beliefs with less entrenchment, such as the credibility of a single eyewitness, might be more easily adjusted in light of new, more compelling evidence.

The AGM Postulates for Contraction formalize this process by ensuring that the agent can remove conflicting beliefs while maintaining as much consistency as possible within the revised belief set. Let's examine the AGM Postulates for Contraction in this context. The AGM Postulates for Contraction formalize how an agent should handle the removal of inconsistent beliefs from their set:

- $(K_1^-)K_A^-$ *is a belief set: The agent's belief set before and after contraction is defined as a set of propositions.*

- *(K_2^-)$K_A^- \subseteq K$: The contracted belief set is a subset of the original belief set. This reflects the idea that beliefs must be removed or adjusted without introducing contradictions in the remaining beliefs.*
- *(K_3^-)$If \ \neg A \notin K$, then $K_A^- = K$: If the belief $\neg A$ (the negation of some proposition A) is not in the belief set, no revision occurs.*
- *(K_4^-)$If not \vdash A$, then $A \notin K_A^-$: If a belief cannot be logically deduced, it will not be present in the contracted belief set.*
- *(K_5^-)$If A \in K$, then $K \subseteq (K_A^-)_A^+$: If a belief A is present, the contraction of the belief set must expand to include A as consistent with the new theory.*
- *(K_6^-)$If \vdash A \leftrightarrow B$, then $K_A^- = K_B^-$: If A and B are logically equivalent, revising by A or B will yield the same contracted belief set.*

In Jane's case, these postulates help guide her through the process of contracting her belief set. She must decide whether to discard certain beliefs such as her initial belief in Mr. Green's guilt while maintaining others. The epistemic entrenchment of each belief helps Jane determine which beliefs are most important to retain and which can be modified or eliminated. The AGM Postulates for Contraction provide a formal framework for guiding Belief Revision when new, conflicting information arises. They help agents like Jane make rational decisions about which beliefs to revise and which to retain, ensuring that their beliefs remain consistent with new evidence. The principle of epistemic entrenchment provides a guiding intuition for this process, emphasizing the relative importance of beliefs and how they should be adjusted when new information challenges the existing belief set.

In the next section, we will discuss how Grove's Spheres models extend these ideas and offer a more detailed approach to modeling Belief Revision and defeasible reasoning in dynamic contexts, such as the one Jane encounters in her investigation.

2.9 Grove's Spheres Models and Belief Revision

One model for Belief Revision is Grove's System of Spheres, which was inspired by Lewis's system for counterfactuals [see 58, p. 256]. While Lewis's counterfactuals focus on possible worlds, Grove's system centers on a belief set of possible worlds. This model allows for the representation of beliefs in a structured way, accounting for how beliefs are formed, revised, and retracted based on new information. Instead of simply relying on a single possible world, Grove's system orders a set of possible worlds according to their consistency with an agent's current belief set, allowing for the revision of beliefs as new, potentially contradictory information arises.

Grove's system of spheres [37] is an important framework for formalizing Belief Revision, as it captures the dynamic nature of belief sets and how they evolve in light of new information. Let K represent a belief set, and let $[K]$ be a subset of the set M_L, which consists of all maximal consistent extensions of a language L. Formally, $[K]$ is defined as follows:

Definition 2.9.1 $[K] = \{M \in M_L : K \subseteq M\}$

This definition encapsulates the idea that the belief set $[K]$ is a subset of all possible worlds M_L where every sentence in the belief set K is included. This formulation of belief sets allows the agent to reason about their beliefs in the framework of all possible worlds that are consistent with K.

Let **S** represent a collection of subsets of M_L. This collection is called a system of spheres centered on $[K]$ if it satisfies the following conditions:

- **(S1)** **S** is totally ordered by $\subseteq$; that is, if $X, Y \in \mathbf{S}$, then either $X \in Y$ or $Y \in X$.
- **(S2)** $[K]$ is the $\subseteq$-minimum of **S**; that is, $[K] \in \mathbf{S}$, and for any $X \in \mathbf{S}$, $[K] \in X$.
- **(S3)** M_L is in **S**, making it the largest element in the system.
- **(S4)** If A is a sentence and there is a sphere in **S** intersecting $[A]$, then there is a smallest sphere in **S** that intersects $[A]$.

These conditions provide a way to model Belief Revision in a way that is grounded in the consistency of the belief set. The system of spheres formalizes how an agent's beliefs can be revised as new information emerges. Grove's system introduces a function that selects the closest worlds to $[K]$, in which a given sentence A holds. This function captures the idea that some worlds are "closer" to an agent's belief set, based on their consistency with that belief set.

Let's consider Jane, the detective investigating a crime in a wilderness reserve. Jane initially believes that the suspect, Mr. Green, is guilty based on circumstantial evidence. She holds this belief in her mind, and it forms part of her initial belief set K. As Jane continues her investigation, she uncovers a new piece of evidence an eyewitness report that suggests Mr. Green was elsewhere at the time of the crime. This new evidence contradicts Jane's belief. Using Grove's System of Spheres, Jane's belief set K is updated to account for the new evidence. The belief set $[K]$ represents all the possible worlds where Jane's initial belief in Mr. Green's guilt holds. However, with the new evidence, Jane's belief set expands to include new worlds that are consistent with the eyewitness report. These new worlds form part of a new system of spheres, where the closest worlds are those where the new evidence is true, and her initial belief is revised or retracted.

For example, suppose Jane now holds the belief that Mr. Green is not guilty, based on the eyewitness report. The belief set $[K]$ will now be centered on this new belief. The closest worlds will be those where the new evidence (the alibi) holds, and Jane's belief in Mr. Green's guilt is no longer part of the belief set. In this case, Jane's belief set has been revised, and the relevant worlds in M_L reflect this revision. This process of revision is governed by Grove's system, which allows for the identification of the closest worlds and the revision of Jane's beliefs accordingly. The system of spheres formalizes this process, ensuring that Belief Revision respects the structure of the belief set and the consistency of the information involved.

To formalize the process of Belief Revision, Grove introduced an ordering on the sentences of the language, denoted by $\leq_G$. This ordering is crucial for guaranteeing that the Belief Revision operator maintains consistency and respects the agent's ordering of beliefs. The postulates for $\leq_G$ are as follows:

- ($\leq_G$ 1) $\leq_G$ is connected.
- ($\leq_G$ 2) $\leq_G$ is transitive.
- ($\leq_G$ 3) If $\vdash A \rightarrow B \vee C$, then $B \leq_G A$ or $C \leq_G A$.
- ($\leq_G$ 4) $\neg A \notin K$ if and only if $A \leq_G B$ for all B.
- ($\leq_G$ 5) $\vdash \neg A$ if and only if $B \leq_G A$ for all B.

These postulates ensure that the Belief Revision operator maintains consistency and respects the agent's ordering of beliefs. The function that selects the closest worlds and the corresponding revision operator is defined as follows:

Definition 2.9.2 *(C $\leq_G$) $B \in K{*}_A$ if and only if $A \wedge B \leq_G A \wedge \neg B$*

This operator formalizes how belief sets are revised when new information A is introduced. The operator ensures that the revised belief set remains consistent and that the closest worlds are selected, preserving the logical structure of the Belief Revision process. Grove's Spheres model provides a formal, structured approach to Belief Revision, offering a way to select the closest worlds based on the consistency of an agent's belief set. By introducing an ordering on the language and a revision operator that respects AGM postulates, Grove's model allows for nonmonotonic reasoning and defeasible Belief Revision in a way that accounts for new, potentially conflicting information. Jane's investigation demonstrates how this model can be used to revise beliefs when confronted with new evidence.

2.10 Belief Revision Considerations and Challenges

Rott [90] highlights that the idea of modeling nonmonotonic reasoning using Belief Revision was first suggested around 1980 and then formally developed in the AGM Theory. The central idea behind Belief Revision is to model beliefs based on an agent's acceptance set, with the mechanisms for belief change being sensitive to the agent's belief states and the content of those beliefs. The basic principle is that the resulting belief set, after a revision, should maintain the same format as the original set. This ensures that revisions are rational and coherent, allowing agents to make informed decisions based on updated belief states. If we conceptualize the belief set of an agent as an epistemic state, the revised belief set will also be represented by epistemic states, ensuring coherence between the theories.

The core challenge in Belief Revision lies in ensuring that the beliefs remain consistent and logically coherent, even when new, potentially conflicting information is added. This is why Rott [90] discusses *Maxims of Coherence for Belief Representation and Revision*, which guide how belief changes should occur:

(i) **Consistency:** An agent's beliefs should be consistent. This is a fundamental requirement for rationality, as inconsistent beliefs can lead to paradoxical conclusions.

(ii) **Inferential Closure:** If a sentence φ may be inferred from an agent's beliefs, then φ should be believed by the agent. This ensures that beliefs are logically coherent and the agent is willing to accept inferences drawn from them.

(iii) **Minimal Change:** The amount of information lost in a belief change should be minimal. This is the principle of Informational Economy, where we seek to preserve as much of the agent's original belief set as possible while incorporating new information.

(iv) **Epistemic Entrenchment:** As some beliefs are considered more important or entrenched than others, one should retract the least entrenched beliefs first when revising beliefs.

The first maxim, consistency, is crucial as it upholds the rationality criterion we discussed earlier, asserting that a rational agent should maintain a consistent belief set. The second maxim, inferential closure, ensures that the agent adopts all logical consequences of the beliefs they hold. The third and fourth maxims focus on minimizing the disruption to the agent's belief system, ensuring that Belief Revision is as non-disruptive as possible. However, these maxims also present challenges in everyday reasoning, particularly when new information contradicts deeply held beliefs.

2.10.1 Informational Economy and Minimal Change

One of the most debated aspects of Classical Belief Revision is the *Informational Economy* or *Minimal Change*. This principle states that when revising beliefs, an agent should change as little as possible in their belief set. Originally, Informational Economy was based on the idea that information is costly, and unnecessary losses of information should be avoided [90]. However, in contemporary contexts, particularly with the vast availability of information, it may seem that the principle of minimal change is less relevant. Some of the agent's original beliefs could be based on faulty or uncertain information. Moreover, in some cases, new information might require substantial revisions to a belief set, especially when earlier beliefs were based on incomplete or outdated knowledge.

For instance, consider Jane, the detective investigating a crime in a forest reserve. Initially, she believes that the crime occurred in a specific area based on witness testimony. However, when new evidence emerges, suggesting that the crime occurred in a different location, Jane must revise her belief. According to the Informational Economy principle, Jane would ideally minimize the revision of her belief set, but the new evidence may necessitate a substantial shift in her reasoning. In this case, minimal change would involve revising only those beliefs that are inconsistent with the new evidence, while keeping other beliefs intact.

Gärdenfors [32] provides the following formulation for the constraint of Minimal Change: "When changing beliefs in response to new evidence, you should continue to believe as many of the old beliefs as possible "**?** , 75. This formulation underlines the importance of retaining the core of an agent's belief set, even as they integrate new, sometimes contradictory, information.

The principle of Minimal Change is closely related to *epistemic entrenchment*, which determines which beliefs are more "entrenched" in an agent's belief set. Epistemic entrenchment reflects the relative importance or stability of beliefs within a belief set. Some beliefs are deemed more essential and should be preserved, while others, deemed less critical, may be revised or discarded when new information arises.

2.10.2 Epistemic Entrenchment and Ordering

Epistemic entrenchment refers to the degree to which a belief is held firmly within the belief set. Beliefs that are more entrenched are considered more central to an agent's reasoning process and are less likely to be revised in the face of new evidence. Conversely, beliefs with lower epistemic entrenchment are more vulnerable to revision. This concept plays a crucial role in the revision process, as it helps guide which beliefs should be retained and which can be discarded when new information is incorporated.

To illustrate the concept of epistemic entrenchment, consider the following example using Jane the detective. Imagine that Jane, after conducting an investigation, believes that the suspect's alibi is solid. However, after receiving new evidence that suggests the alibi might be false, she must revise her belief. The belief in the suspect's alibi was initially entrenched, as it was based on concrete evidence. However, upon receiving new evidence, this belief is subject to revision. The newly acquired evidence, which challenges the entrenched belief, would be considered more relevant and would likely lead Jane to update her belief set accordingly.

To formalize epistemic entrenchment, we define an ordering $\leq_e$ on beliefs, where A is considered more entrenched than B if A is more central to the agent's reasoning. The relationships between beliefs can be characterized as follows:

- $A \leq_e B$: Belief A is at least as epistemically entrenched as belief B.
- $A <_e B$: Belief A is less epistemically entrenched than belief B.
- $A \equiv_e B$: Beliefs A and B are equally epistemically entrenched.

We can also define several properties of epistemic entrenchment orderings:

- **Transitivity**: If $A \leq_e B$ and $B \leq_e C$, then $A \leq_e C$.
- **Dominance**: If $A \models B$, then $A \leq_e B$.
- **Conjunctiveness**: Either $A \leq_e (A \wedge B)$ or $B \leq_e (B \wedge A)$.

The agent's belief set can be revised according to this ordering. In cases where multiple beliefs must be revised, those beliefs with the lowest epistemic entrenchment are typically the first to be discarded. This is because they are considered less crucial to the agent's overall reasoning and decision-making process.

> The epistemic entrenchment of a sentence is dependent on the belief state in which it occurs. Different belief sets may thus be associated with different orderings of epistemic entrenchment; even if the sets of sentences contained in the belief sets overlap, the orderings of epistemic entrenchment need not agree on the common parts. [32, p.88]

Thus, epistemic entrenchment provides a formal mechanism for determining which beliefs should be retained or discarded in response to new evidence. This is particularly useful when the agent is faced with conflicting or contradictory information. The ability to prioritize beliefs based on their epistemic importance ensures that the revision process remains coherent and rational.

In summary, when agents encounter new information that contradicts their existing beliefs, the process of Belief Revision is guided by the principles of consistency, minimal change, and epistemic entrenchment. The goal is to revise beliefs in a manner that maintains the coherence of the belief set while incorporating new, potentially contradictory, information. The next section will explore the formal framework for Belief Revision, focusing on how AGM postulates can be applied to model this process.

2.11 Further Notes on Revision and Contraction Functions

In Belief Revision, *Expansion* is often considered an unproblematic case, as it simply involves adding new information to a belief set without removing any previously held beliefs. However, the more complex operations, such as *Revision* and *Contraction*, are the primary focus of standard AGM approaches to Belief Revision. These functions are crucial because they allow an agent to update their belief set in a rational way when faced with new, often contradictory, information. Understanding how to revise or contract a belief set is central to modeling how agents change their minds over time.

The challenge with Revision and Contraction is determining how to update the belief set in a way that preserves consistency while incorporating new, potentially conflicting information. There are two main equivalent proposals to determine revision and contraction functions in Belief Revision theory [32]. The first approach relies on the concept of epistemic entrenchment, which provides a way to order the beliefs within a set according to their "strength" or "importance." In this model, revision and contraction operators are defined based on the ordering of epistemic entrenchment. The second approach uses the *Informational Economy* principle, which focuses on selecting the *maximal subsets* of a belief set K that do not include a particular proposition A.

Epistemic entrenchment reflects the importance or stability of a belief within a belief set. To explain this further, let us define the set of belief sets that do not contain a given proposition A. We denote this as $K \perp A$, which represents the collection of all maximal subsets of K where K' fails to imply A. This leads to the definition of a selection function S_{ch}, which picks out the most "relevant" belief sets from $K \perp A$ according to some rational ordering.

Definition 2.11.1 (Selection Function S_{ch}) *Let S_{ch} be a selection function that picks the most relevant belief set from the collection $K \perp A$ by selecting the belief set with the lowest epistemic entrenchment, i.e., the least entrenched belief sets within*

K that exclude A.

$$S_{ch}(K \perp A) = \{K' \in K \perp A : K^n \leq_e K'' \text{ for all } K^n \in K \perp A\}$$

This selection function ensures that when revising a belief set, the least entrenched beliefs are the ones most likely to be discarded.

Once the selection function is defined, we can proceed to define the various contraction functions that apply when a new proposition A is introduced. The contraction functions determine how to remove inconsistent beliefs while keeping the belief set as intact as possible.

Definition 2.11.2 (Maxichoice Contraction Functions) *The Maxichoice Contraction function selects the union of all maximal subsets of K that exclude A when such a set is non-empty. If no such subset exists, the original belief set K remains unchanged.*

$$K^{-}A = \begin{cases} \bigcup(K \perp A) & \text{whenever } K \perp A \text{ is not empty,} \\ K & \text{otherwise.} \end{cases}$$

Definition 2.11.3 (Partial Meet Contraction Functions) *The Partial Meet Contraction function takes the intersection of all the subsets in $S_{ch}(K \perp A)$, which selects the most relevant subsets of K that exclude A. This ensures that the most coherent belief sets are retained, even when a large portion of the set must be revised.*

$$K^{-}A = \bigcap S_{ch}(K \perp A)$$

Definition 2.11.4 (Full Meet Contraction Functions) *The Full Meet Contraction function selects the intersection of all maximal subsets in $K \perp A$. If there are no such subsets, the original belief set is retained. This approach ensures that the least amount of information is discarded, but it may result in a more significant revision of the belief set.*

$$K^{-}A = \begin{cases} \bigcap(K \perp A) & \text{whenever } K \perp A \text{ is non-empty,} \\ K & \text{otherwise.} \end{cases}$$

To make these concepts more intuitive, let's apply them to Jane, our detective in the Wilderness scenario. Imagine that Jane has already formed several beliefs about the crime scene based on initial evidence. She holds the belief that "the crime occurred near the west edge of the forest," which she considers an entrenched belief due to the weight of the evidence. However, when new information emerges from another witness suggesting that the crime may have taken place closer to the east side of the forest, Jane must revise her belief set.

1. Maxichoice Contraction In this case, Jane's belief set would be revised by considering all possible belief sets that exclude the new proposition ("the crime occurred near the east side of the forest"). The Maxichoice function would take the union of those belief sets, focusing on the most relevant subsets that exclude the contradictory information. This would involve keeping beliefs that still align with the west-side hypothesis while discarding beliefs that conflict with the new information.

2. Partial Meet Contraction Alternatively, Jane could use the Partial Meet Contraction function, which would select the intersection of the most relevant subsets of her belief set that exclude the new proposition. This would allow Jane to retain the most coherent beliefs while discarding the least relevant ones, ensuring that her revised belief set is as rational and consistent as possible. In this case, Jane would prioritize the beliefs that still make sense given both the west-side and east-side possibilities, rejecting those that no longer hold.

3. Full Meet Contraction If Jane takes the Full Meet Contraction approach, she would select the intersection of all maximal subsets of her belief set that exclude the new proposition. This approach would result in the greatest degree of Belief Revision, discarding more of her original beliefs in order to accommodate the new evidence. This might happen if the new information about the east-side location of the crime is overwhelming and clearly more consistent with the available evidence.

One of the main challenges in using these functions is determining the ordering of epistemic entrenchment, or which beliefs should be discarded first when new information contradicts the current belief set. As we discussed earlier, epistemic entrenchment reflects the importance or stability of a belief within the belief set, but the process of determining this ordering is not always straightforward. It is particularly difficult when the agent faces conflicting information and must decide which beliefs are more entrenched or important than others.

This issue can lead to the *non-uniqueness problem*, where different reasonable orderings of beliefs can lead to different revision outcomes. The lack of a single, universally agreed-upon ordering of beliefs can make it difficult to construct a single, objective contraction or revision function. This challenge is one of the reasons why Belief Revision is a complex and dynamic process, requiring careful consideration of the agent's epistemic state and the context in which the revision occurs.

2.12 Graphic Representation of Belief Revision

There are several possible ways to represent Belief Revision visually. Those graphic representations intend to explicate what is going on with the agent's sets of beliefs during the revision processes. We will briefly show some examples in order to make the explanation more visual. Consider the following representation, for instance. Here we have an ordering of belief sets (sets of propositions the agent believes in). While revising with A, we take all the maximal subsets of $M(K)$ that implies A. Now, we have the resulting set $K * A$.

Moreover, we can represent the Belief Revision operation like Gärdenfors [32] suggests. Suppose we accept a proposition A. In the revised set $K * A$, we might give up some of our beliefs. If we further on accept B and B is consistent with $K * A$, the resulting set would be the same as revising K with $A \wedge B$ (or $A\&B$). This is a possible worlds representation of the relation between K*A and K*AB, according to Gärdenfors:

This graphic representation might be accurate to the case described, but there are other ways to represent the beliefs within the Belief Revision. The idea is that those representations can show what is going on during the revision process. For instance, for a system of spheres centered on $[K]$, Gärdenfors [32] give the following representation:

This representation, however, can lead to misunderstandings. The problem is that it can leave the impression that the sphere-centered set is contained in the other sets. It is even regarded as *smaller*. However, that does not represent the fact that our ordering of beliefs sometimes does not have a smaller set in terms of how many propositions it has. Sometimes, the first graphic representation can be the case, leaving more room to intuitively understand that the belief sets might not share all the propositions believed in the smaller sphere-centered set [K].

For any proposition $[A]$, by the conditions given by the system of spheres, if $[A]$ intersects any sphere in $\mathcal{S}$, there is some sphere S_A that intersects $[A]$, and it is smaller than any other sphere with that property. The set $C(A) = \{[A]\} \cap S_A$ is the set of *closest* elements in $M(K)$. Thus, the revision $K * A$ can be represented by the set $C(A)$ of the closest possible worlds like the following:

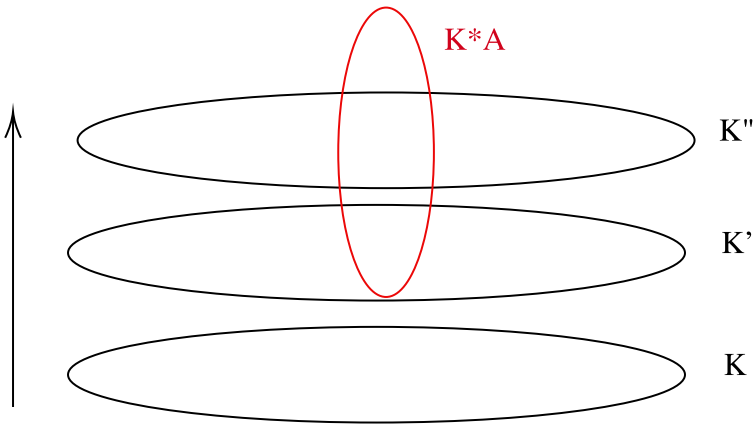


Figure 2.1: Representation of Belief Revision: Initial Knowledge Set and its Revisions. A visual representation of the belief revision process, where the agent's knowledge set (K) undergoes revision upon accepting a new proposition (A), resulting in K*A.

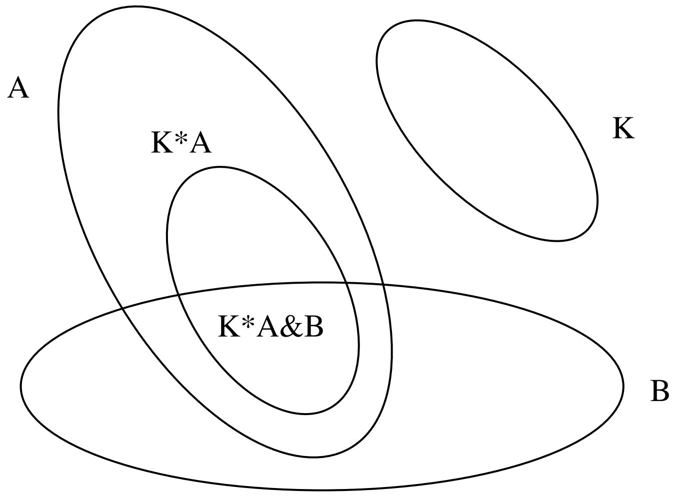


Figure 2.2: Possible Worlds Representation of Belief Revision. This diagram illustrates the relationship between the revised knowledge set K*A and its further revision with another consistent proposition B, leading to K*AB, following Gärdenfors' approach.

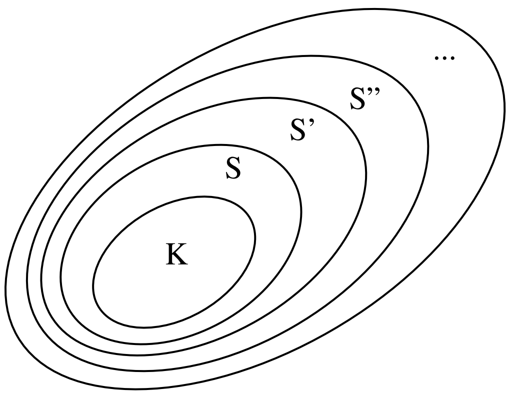


Figure 2.3: System of Spheres Representation of Belief Revision. A system of spheres representation of belief revision, where the ordering of spheres around K illustrates how revision is determined by the closest possible worlds in the model.

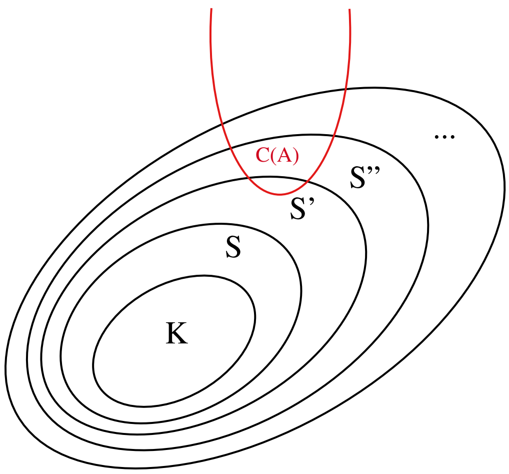


Figure 2.4: Closest Possible Worlds Representation in Belief Revision. The system of spheres shows how a proposition $[A]$ intersects spheres in $\mathcal{S}$, leading to the selection of a minimal sphere S_A that contains the closest possible worlds for the revision process.

2.13 Choosing Belief Revision

In our exploration of formal systems, we have examined several approaches, such as nonmonotonic logics, to model defeasible reasoning. However, these systems are not always suited for capturing the complexities of human reasoning, especially when faced with new indexical information. Indexical expressions, such as “I” or “here,” depend on context, and the human mind often revises beliefs based on changing information. Nonmonotonic logics, though insightful, are generally tailored to models where beliefs evolve in response to contradictory or missing information, rather than to the dynamic nature of context-sensitive beliefs.

We have learned from the formal systems we reviewed that while consistency remains a central goal, they fall short of addressing situations involving complex, everyday reasoning, particularly those requiring self-locating beliefs or adjustments based on shifting contexts. For instance, Reiter’s Default Logic emphasizes the importance of maintaining consistency but remains limited by its inability to handle context-sensitive Belief Revision particularly when beliefs are grounded in indexical or self-locating information. This is crucial when we consider situations where the agent’s knowledge or belief must adapt to new contextual cues or changing perspectives. Moreover, Autoepistemic Logic, which explores how agents reason about their own beliefs, provides valuable insights into introspective reasoning. While it offers a framework where agents can derive new conclusions based on reflective analysis, it still lacks the ability to fully incorporate indexical beliefs or accommodate changes in epistemic state in response to new, context-dependent evidence. Both of these frameworks though useful in particular contexts do not provide a comprehensive solution to Belief Revision as it apply to everyday scenarios involving shifting knowledge and context-sensitive reasoning. And that is one of the main issues we identify in our thesis.

At this point, we shift our focus to a more flexible and intuitive framework for modeling belief changes Belief Revision . Unlike nonmonotonic logics or Autoepistemic Logic, Belief Revision offers a more structured and pragmatic approach to defeasible reasoning. It provides a formal mechanism to revise an agent’s beliefs when confronted with new, sometimes contradictory, information, in a way that preserves consistency and rationality. The AGM Postulates , developed by Alchourrón, Gardenförs, and Makinson, offer a set of rationality criteria for how beliefs should change. These postulates guide the agent’s epistemic state changes while ensuring that the resulting belief set is as coherent as possible.

Belief Revision operates on ordinary propositions, which are assumptions or statements about the world that do not inherently include context-sensitive elements like indexicals. In other words, standard Belief Revision handles beliefs that are fixed, timeless, and not dependent on an agent's current location, time, or perspective. However, as we will explore further in subsequent chapters, this framework, while powerful, does not fully address the issue of context-sensitive beliefs, which often arise in everyday reasoning. This gap is especially evident when indexical expressions such as "I am here" or "this event" play a role in belief formation and revision. These expressions are inherently tied to context, and as such, they introduce a layer of complexity that ordinary propositions cannot capture.

To illustrate this, consider the example of Jane, the detective . Initially, Jane may believe that "The crime occurred near the west edge of the forest," based on the first piece of evidence she receives. However, as Jane gathers more information, she learns that the witness's testimony suggests the crime may have occurred closer to the east side. This new evidence challenges her initial belief, requiring revision. If Jane is to update her belief set rationally, she must incorporate the new information while preserving consistency within her belief system. In Jane's case, the revision function will help her balance her current beliefs with the new evidence. But there's an additional challenge: Jane's beliefs are not just ordinary propositions. The shifting perspectives and context-dependent nature of the new information introduce indexical elements such as the location of the crime scene. As new testimony becomes available, Jane's belief set is no longer just a collection of fixed, timeless propositions. Instead, it must account for the shifting context, with beliefs changing depending on how they are situated in time and space.

This is where indexical expressions complicate the standard Belief Revision framework. Standard AGM Postulates assume that beliefs can be represented as static propositions. However, to apply Belief Revision effectively in cases like Jane's, we need to extend the framework to handle indexical information. This means we must develop a way to represent context-sensitive beliefs, where the truth of a belief may depend on the agent's current location, time, or even their perspective on the evidence. The standard AGM Postulates , while useful for revising beliefs about static propositions, do not account for this dynamic, context-sensitive aspect of Belief Revision. Therefore, to apply Belief Revision to everyday scenarios like Jane's, we must adapt the AGM framework to handle indexical expressions. This extension will allow us to represent beliefs that change not only due to new evidence but also because of shifts in context such as Jane's changing understanding of where the crime occurred.

Belief Revision offers a promising foundation for formalizing defeasible reasoning, and the AGM Postulates provide a solid starting point for guiding rational belief changes. However, to fully capture the complexities of human reasoning, especially in situations where context-sensitive beliefs are involved, we need to extend the standard framework to include indexical expressions. By doing so, we can better model the process through which agents like Jane revise their beliefs in response to new, context-dependent information. This will be the focus of the next chapters, where we will explore how to incorporate context-sensitive beliefs into the formal system of Belief Revision , making it more suitable for everyday applications like detective work, scientific inquiry, and everyday reasoning.

We learned, so far, that although nonmonotonic systems have initially been relevant to our investigation, the formal proposals presented are not intended to reflect cases in which flawed human reasoning faces new indexical information. Each of the systems we reviewed provided valuable insights, but they do not fully address the complexities of Belief Revision in such contexts. Reiter's Default Logic teaches us that consistency is a key principle in Belief Revision, and this aligns with classical logic. However, while Autoepistemic Logic provides a mechanism for introspective reasoning, it still lacks the capacity to handle self-locating beliefs. Nonmonotonic approaches are adequate when the closed world assumption is relevant, and they are effective within artificial intelligence frameworks, but they are not well-suited for modeling the shifting, context-sensitive nature of Belief Revision we encounter in everyday human reasoning.

We, then, turn our attention to a closely related framework: Belief Revision . Belief Revision offers a more intuitive and effective way to model defeasible reasoning, particularly when it comes to belief changes. The AGM Postulates offer a robust foundation for formalizing Belief Revision and provide clear guidelines on how an agent's epistemic state should evolve when confronted with new information. However, as we have seen, standard Belief Revision focuses on ordinary propositions and does not fully accommodate context-sensitiveness . Context-sensitiveness is crucial when indexical expressions like "I" or "here" play a role in belief formation and revision. In the next chapters, we will explore how to extend Belief Revision to incorporate these expressions and apply the framework to situations involving self-locating beliefs. By doing so, we can develop a more robust and nuanced model for Belief Revision, one that is applicable to everyday scenarios like detective work, scientific inquiry, and daily decision-making.

The connection between Belief Revision and the debate on contexts becomes particularly evident when we consider how beliefs are formed and revised in dynamic environments. Contexts play a central role in determining the truth conditions of many beliefs, especially those involving indexical expressions, such as "I" or "here." These expressions inherently rely on the context in which they are used, and any shift in context can alter the truth of the belief. In the case of Belief Revision , the introduction of new evidence that is context-dependent requires not only the updating of beliefs but also a reconsideration of the context in which those beliefs were formed. This brings us to a fundamental issue in Belief Revision: how to formalize and represent contexts in a way that accounts for the shifting nature of beliefs. As AGM Postulates rely on static, timeless propositions, they are insufficient to capture the complexities of Belief Revision in everyday scenarios where context shifts over time, as seen in the case of Jane the detective . Therefore, the extension of Belief Revision to include context-sensitive beliefs, and a proper formalization of context, is essential to make the framework applicable to everyday decision-making and reasoning.

This challenge opens up broader philosophical debates about the nature of context and its role in reasoning. Traditionally, contexts have been treated as external to the agent's belief system, providing the backdrop against which beliefs are evaluated. However, as we move toward integrating context-sensitive beliefs into formal systems like Belief Revision , we must reconsider the relationship between beliefs and the contexts in which they are situated. The key question here is: how do we model the interaction between the agent's epistemic state and the context in which they are embedded? This question ties directly to ongoing discussions in philosophy, particularly in areas like indexical epistemology and contextual semantics , where scholars examine how context influences meaning and belief formation. By exploring these issues in conjunction with Belief Revision , we can develop a more refined framework that accommodates not only static beliefs but also the dynamic, context-sensitive nature of human reasoning. This will allow us to model belief changes that are driven by both new information and shifting contexts, offering a more robust understanding of how beliefs evolve in a world that is constantly changing.

Part II

The Indexicality Phenomenon

Chapter 3

Indexicals and Contextual Information

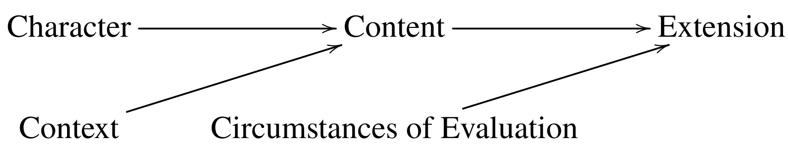
In everyday communication, we often use sentences that contain indexical expressions, which are context-sensitive in nature. *Indexicals* are expressions that depend on the context in which they are uttered. According to Kaplan [45], indexicals have two components: *character* and *content*. The character of an indexical is the function that assigns content in a given context, while the content is the specific context-dependent meaning of the expression.

For example, consider a statement made at a specific time t. The content of the statement will depend on the context, such as the speaker's location or the time at which the statement is made. The character, however, remains independent of the context and provides the structure for how the content is derived. In simpler terms, while the content is determined by the specific context, the character gives us a consistent method of interpreting the content across different contexts. Expressions involving demonstratives, such as *this*, *that*, or *I*, will typically convey different content depending on the context in which they are used. As Kaplan explains, the *context* of an indexical expression is defined as the "possible occasion of use" of a sentence [45].

Kaplan [45] distinguishes between two primary types of indexicals: *pure indexicals* and *genuine demonstratives* (or simply *demonstratives*). It is important to note that Kaplan uses the term *demonstratives* to refer to all types of indexicals, including both pure indexicals and genuine demonstratives, which can lead to some confusion. The key difference between pure indexicals and genuine demonstratives lies in their requirement for an associated demonstration. Pure indexicals do not require an associated demonstration, whereas genuine demonstratives do [45]. In this work, the term *indexicals* will be used to refer to pure indexicals unless otherwise specified, though it is possible to extend the results to genuine demonstratives as well.

The study of indexicals has led to several significant issues in the philosophy of language and formal semantics. Incorporating context-sensitivity into formal systems poses challenges because context is dynamic and difficult to represent in a rigid framework. This issue led Kaplan to develop a formal system to handle a language containing demonstratives, known as the *Logic of Demonstratives* (LD), which will be briefly explored later in this chapter.

In addition to Kaplan, John Perry also made substantial contributions to the study of indexicals, albeit with a slightly different approach. Perry [75] conceptualizes the act of speaking, writing, or any communicative act as an *utterance*, which he views as an *intentional act*. Perry is influenced by Reichenbach's work, and he posits that tokens such as utterances or written words are traces left by these intentional acts [75]. Russell [93] provides a framework for understanding Kaplan's two-dimensional theory of indexicals, presenting the following diagram:



Perry further distinguishes between *automatic* and *intentional* indexicals. Automatic indexicals function like pure indexicals, where the content is automatically attributed by the rules of use. For instance, the indexical expression "I" automatically refers to the speaker, without the need for further contextual determination. On the other hand, *intentional indexicals* require consideration of the speaker's intention to establish the content of the utterance. The interpretation of these indexicals depends not only on the context but also on the speaker's intended meaning, which introduces a layer of complexity in their interpretation. A well-known example that Perry uses to illustrate his theory is the case of Rudolf Lingens, as discussed by Lewis [56]:

> An amnesiac, Rudolf Lingens, is lost in the Stanford Library. He read a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost. He still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, "This place is aisle five, floor six, of Main Library, Stanford. I am Rudolf Lingens" [74, 497].

This example is particularly useful for understanding how indexicals can be problematic in cases of self-identification and location, as the context plays a pivotal role in the meaning of the utterance. Despite the fact that Lingens can understand sentences like "You are hungry" or "It is time to leave *now*" , his lack of personal identity and knowledge of his location prevents him from properly interpreting indexicals such as “I” or “here” . According to Mellor [66], this phenomenon is influenced by the role that causal continuity plays in our subjective beliefs about time and identity. The referent of indexicals like “I” and “now” is fixed by the context, which includes the agent’s own psychological state.

Context is, indeed, a central notion in the debate surrounding indexicals. Typically, a context is required to specify an individual, a moment in time, a temporal location, and a possible world. However, it is tempting to think of context as a set of presuppositions, which can be problematic because context is not merely a set of assumptions but a dynamic feature of communication. The debate over the nature of contexts is extensive, and in this chapter, we will focus on Kaplan’s definition of *Context* [45]. In the next chapter, we will explore various definitions of context and how they can be formally represented. For now, we will primarily investigate Kaplan’s *Logic of Demonstratives*, as originally proposed by Kaplan [46], and consider how it incorporates context-sensitivity. It is also important to note that the development of a formal system to handle indexicals inherently involves the treatment of temporal features, which are generally less problematic than other forms of context. Thus, Kaplan’s *Logic of Indexicals* draws on Tense Logics, with Temporal Indexicals serving as a good starting point to understand his theory.

3.1 Temporal Indexicals and Tense Logic

There is an interesting phenomenon concerning time and indexicality: the temporal indexicals. We know that *indexicals* are linguistic expressions whose reference shifts from context to context. Temporal indexical expressions, such as “now” , “today” , and others, refer to moments or periods in time, depending on the instant of the utterance of a sentence in which the indexical occurs. Temporal indexicals are closely tied to the context in which they are used, as they depend on the time of utterance and provide information about when the speaker is referring to. According to Perry [77], temporal indexicals differ from *dates*, as dates are considered to behave like descriptions that pick up a specific instant within the temporal space, whereas temporal indexicals are context-sensitive and depend on the moment of utterance.

Dates are the most common way we refer to specific times, especially when talking about past events. According to Perry [77], periods of time are related to one another in an orderly way, and this order is incorporated into our system of dates to mark specific times. This system can be seen as a figurative way of talking about time, but it does not necessarily reflect the reality of time or its position within a possible world. Dates are special types of descriptions that help us refer to specific periods, but they do not behave like indexicals. Kaplan's formal theory of indexicals, as we will see, does not account for dates.

Perry [77] offers a useful distinction of temporal indexicals and their role in utterances. He provides a shortlist of temporal and other indexicals to clarify how these expressions play a role in an utterance:

- An utterance *u* of "today" refers to the day on which *u* occurs.
- An utterance *u* of "yesterday" refers to the day before the day on which *u* occurs.
- An utterance *u* of "tomorrow" refers to the day after the day on which *u* occurs.
- An utterance *u* of "present" or "now" refers to the time at which *u* occurs.
- An utterance *u* of "here" refers to the place at which *u* occurs.
- An utterance *u* of "I" refers to the speaker of *u*.

Perry [77] extends his notion of *temporal indexicals* to include *tense* as an indexical phenomenon. However, Prior [81] disagrees, asserting that tense is not an indexical phenomenon. For Prior, the past-tense marker indicates that the event occurred *before* the time of the utterance, and the future-tense marker indicates that the event occurred *after* the time of utterance. This view separates tense from indexicals and places it in a different category of expression. According to Prior [81], temporal markers are more about the relation of events to a given moment in time, rather than shifting their reference from context to context.

Because of the critical role that tenseness plays in language, it is natural that logicians would attempt to formalize temporal information. Formal systems based on Classical Propositional Logic that do not include temporal information representation are insufficient for formalizing tensed propositions. Therefore, tense formal systems were developed to address the representation of temporal information within a logical framework, and these systems are known as *Temporal Logics*. Specifically, *Tense Logic* was introduced by Arthur Prior [20] and forms the underlying system for many subsequent formalisms. Prior was particularly concerned with the question of whether time flows or passes. He argued that "genuine flowing or passage is something that occurs *in* time and *takes* time to occur" [81]. This view of time flows challenges the metaphorical use of time passage, which suggests that time moves as a simple process.

Prior, like many philosophers, was also concerned with the notion of change and its relationship to time. For him, things change, but events do not change; rather, events *happen*. He justifies his view by stating that to say something happened is to say that something changed in some way [81]. Moreover, existing things can change, but it raises an interesting issue: if something is becoming more past relative to a near past, what does it mean for something to be more in the past if the event no longer exists [81]?

In order to address these philosophical concerns, Prior turned to formal language. He argued that temporal operators should receive similar treatment as modal operators [67]. Temporal operators could prefix a sentence and be combined to represent different kinds of tensed structures in a language. This approach led to Prior's development of *Tense Logic*, which focused on tense, not indexicality. One limitation of Prior's system was its inability to handle the shifting reference of context-sensitive beliefs, such as those involving indexical expressions, which were later addressed in Kaplan's work.

Thus, while Prior's *Tense Logic* provided a framework for formalizing temporal propositions, it did not tackle the complexities introduced by context-sensitive indexicals. This limitation led to the development of Kaplan's formal system, which sought to handle indexicals and context shifts in a more comprehensive way. Because the development of **LD** (Logic of Demonstratives) has its origins in Prior's *Tense Logic*, understanding the treatments towards temporal indexicals starts by understanding the foundations of Prior's system. Furthermore, we will see that most formal treatments of indexicals consider temporality as a main component. This is because it is relatively easier to formalize a sequence or ordering of instants than to formalize sets of individuals and how they relate to each other. This highlights the importance of temporal indexicals in formal systems, as they allow us to represent time in a structured way. Let us take a closer look at Prior's temporal operators to understand this framework better. Prior [31] originally presented the following temporal operators:

- P: can be read as "It was at some time the case that...".
- F: can be read as "It will at some time be the case that...".
- H: can be read as "It has always been the case that...".
- G: can be read as "It will always be the case that...".

These temporal operators allow us to express statements about time, such as past, future, permanence, and necessity. Equivalence between these operators can be intuitively given as follows:

- $Pp \equiv \neg H \neg p$; and
- $Fp \equiv \neg G \neg p$.

These equivalences show how the temporal operators P and F can be related to the operators H and G, allowing us to express different types of time-related statements. Temporal operators increase the expressivity of a system, allowing it to formalize different statements involving time. For instance, Prior [31] used these operators to build formulas that express various philosophical theses about time. Some examples of such formalization include:

- $Gp \rightarrow Fp$ represents "What will always be, will be."
- $G(p \rightarrow q) \rightarrow (Gp \rightarrow Gq)$ represents "If p will always imply q, then if p will always be the case, so will q."

In *Tense Logic*, a *temporal frame* consists of a set T of entities called times (or instants), together with a strict partial ordering relation $<$ on T. This structure defines the metaphorical flow of time, which allows us to use the temporal operators like the ones mentioned above. By formally ordering times, we can represent the passage of time and the relationships between different moments in time.

Along with the truth-conditions for Propositional Classical Logic for propositional fragments of Tense Logic, the truth-conditions for the tense operators can be given as follows [31]:

- Pp is true at a given time t if and only if p is true at some time t' such that $t' < t$.
- Fp is true at a given time t if and only if p is true at some time t' such that $t \leq t'$.
- Hp is true at a given time t if and only if p is true at all times t' such that $t' \leq t$.
- Gp is true at a given time t if and only if p is true at all times t' such that $t \leq t'$.

These truth-conditions show how tense operators can be applied to temporal propositions, representing different time-related relations in a logical framework. The ability to express relationships like "before," "after," "always," and "eventually" makes Tense Logic a powerful tool for formalizing temporal statements. However, as we will see, while Tense Logic is highly effective in capturing relationships between times and events, it has its limitations. Specifically, Tense Logic fails to account for temporal indexicals, such as *now*, which present unique challenges in formalization.

In addition to formalizing temporal propositions, Tense Logic deals with the metaphorical nature of time. Philosophers like Schopenhauer have argued that *being* consists of the set of events that are simultaneous with the utterance of "now" [112]. This perspective brings us to the issues concerning the utterance of *now*, a temporal indexical. Temporal indexicals like *now* are context-sensitive, and their meaning depends on the specific time at which they are uttered. In contrast to the abstract formalism of Tense Logic, *now* refers to a specific, fleeting moment that is tied to the present.

This connection between temporal indexicals and time presents a significant challenge for formal systems like Tense Logic. While Tense Logic is successful at representing time as an abstract, ordered sequence of moments, it lacks the ability to handle the context-dependent nature of temporal expressions like *now*. This limitation is a key motivation for developing formal systems that can better account for indexical expressions, as we will explore further in the next section.

Thus, while Prior's *Tense Logic* provided a significant advancement in the formal treatment of time, its inability to capture the shifting, context-dependent nature of temporal indexicals like *now* leaves a gap that needs to be addressed. Kaplan's work, as we will see in the next chapter, provides a formal framework for dealing with indexicals, including temporal ones, and offers a more complete solution to this issue.

3.2 Logic of Demonstratives

Kaplan adopted the Language of First Order Logic as underlying language to develop his *Logic of Demonstratives*. Yet, Kaplan distinguishes between two kinds of variables [1]: *individual variables* and *positions variables*. The language $\mathcal{L}_{LD}$ of the Logic of Demonstratives also contains Modal Operators, 1-place and Temporal Operators [45]:

- Connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$;
- *Definite Description Operator* "the" ;
- Identity =;
- Modal Operators $\Box$ and $\Diamond$;
- Temporal Operators F ("it will be the case that"), P ("it was the case that"), G ("one day ago, it was the case that");
- 1-place Operators N ("it is now the case that"), A ("it is actual the case that"), Y ("yesterday was the case that");

[1] Besides the distinction between *free* and *bound* variables, this is a distinction about what kind of "object" those variables refers to.

- 1-place Functor "dthat" ;
- *m-n* [2]-place Predicates (in particular, the 1-0-place predicate "Exist" and the 1-1-place predicate "Located");
- Individual *n-m*-functors denoting individuals (in particular, the 0-0-*i*-functor is the individual constant "I"); and
- Position *n-m*-functors denoting positions (in particular, the 0-0-place *p*-functor "Here");
- Quantifiers ∃ and ∀.

A remark might be necessary: the *1-0*-place predicate "Exist" should not be confused with the quantifier ∃. On the one hand, the quantifier ∃ ranges over variables of its scope in a given formula where ∃ occurs and could not refer to an object. On the other hand, there is the 1-0-place predicate "Exist" . That allows us to state sentences such as "$\exists x$ Exist x " meaning "There is an x that exists" [46]. However, Kaplan does not present any ontological questions about using the predicate "Exist" . One possible interpretation is that the predicate states that the variable x denotes an existent referent or even could allow us to talk about the existence of some entities, while the quantifier ∃ rules the variables only on the language level.

The distinction between *individual* variables and *position* variables in **LD** explicit the intuitive distinction between places and objects/agents. Associated with times and worlds, those notions can provide us with detailed similarities and/or distinctions between sentences and what entities they are talking about. The attempt to include *demonstrations* and the use of "dthat" in the formal language can also be viewed as an attempt to include an extralogical intensional element that seems to be missing in some statements. The 1-place functor "dthat" indicates an associated demonstration in the sentence. Demonstratives, as we saw, require associated demonstrations to establish the referent. Using a bound variable in the sentence would be related to an extralogical element, although there is no consensus about what features a demonstration. However, our concern here is not about those statements requiring demonstrations but those referring to times.

It should be clear by now that sentences involving demonstratives will generally express different concepts when uttered in different contexts. However, the notion of a sentence α being *true* in a context does not require an utterance of α [45]. In First Order Logic, the truth-conditions do not depend on the instants and the

[2]The *m-n*-place is given as follows: n is the number of individual variables of the predicate and m, the number of position variables.

possible world: the time is just fixed as the time of the utterance. When talking about indexicals, however, we need to consider the possible world and the instant of the context and, perhaps, other contextual information.

Formally, the semantics of **LD** is given in terms of contexts and it is based on Kripke's Possible World Semantics. $\mathfrak{A} = \langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, \mathcal{T}, \mathcal{I} \rangle$ is an **LD-structure** if and only if there are $\mathcal{C}$ (set of contexts), $\mathcal{W}$ (set of possible worlds), $\mathcal{U}$ (set of individuals), $\mathcal{P}$ (set of positions), $\mathcal{T}$ (set of integers, representing instants), and $\mathcal{I}$ (interpretation function that assigns an intension to every predicate and individual constant of the language) such that [45]:

- $\mathcal{C} \neq \emptyset$; $\mathcal{W} \neq \emptyset$; $\mathcal{U} \neq \emptyset$; $\mathcal{P} \neq \emptyset$; $\mathcal{T} \neq \emptyset$
- If $c \in \mathcal{C}$, then:

  (i) $c_A \in \mathcal{U}$ (the agent of c)

  (ii) $c_T \in \mathcal{T}$ (the instant of c)

  (iii) $c_P \in \mathcal{P}$ (the position of c)

  (iv) $c_W \in \mathcal{W}$ (the world of c)

† is an alien entity that does not belong to $\mathcal{U}$ nor to $\mathcal{P}$ and represents an undistinguished value of the function; usually, we take † as $\{\mathcal{U}, \mathcal{P}\}$. In particular, $\mathcal{I}$ is the function that assigns to each predicate and functor a proper intension as follows [46]:

- (i) If π is a m-n-position predicate, then $\mathcal{I}_\pi$ is a function such that, for each $t \in \mathcal{T}$ e $w \in \mathcal{W}$, $\mathcal{I}_\pi(t, w) \subseteq (\mathcal{U}^m X \mathcal{P}^n)$;

  (ii) If η is an i-functor of m-n-position, then $\mathcal{I}_\eta$ is a function such that, for every $t \in \mathcal{T}$ and every $w \in \mathcal{W}$, $\mathcal{I}_\eta(t, w) \in (\mathcal{U} \cup \{\dagger\})^{(\mathcal{U}^m X \mathcal{P}^n)}$;

  (iii) If η is a p-functor of m-n-position, then $\mathcal{I}_\eta$ is a function such that, for every $t \in \mathcal{T}$ and every $w \in \mathcal{W}$, $\mathcal{I}_\eta(t, w) \in (\mathcal{P} \cup \{\dagger\})^{(\mathcal{U}^m X \mathcal{P}^n)}$.
- $i \in \mathcal{U}$ if and only if $(\exists t \in \mathcal{T})(\exists w \in \mathcal{W})(\langle i \rangle \in \mathcal{I}_{Exist}(t, w))$
- If $c \in \mathcal{C}$, then $\langle c_A, c_P \rangle \in \mathcal{I}_{Located}(c_T, c_W)$
- If $\langle i, p \rangle \in \mathcal{I}_{Located}(t, w)$, then $\langle i \rangle \in \mathcal{I}_{Exist}(t, w)$

Kaplan defines validity of a sentence α in **LD** as being true in every context and every **LD**-structure [45]. Consider an assignment function f such that A function f in $\{cftw\}$ is an *assignment* in respect to $\langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, \mathcal{T}, \mathcal{I} \rangle$ if and only if

$\exists f_1 f_2(f_1 \in \mathcal{U}_i^{\mathcal{V}} \& f_2 \in \mathcal{V}_p \& f = f_1 \cup f_2)$ [46]. A function f assigning x to α can be defined as $f_x^\alpha = (f \sim \{\langle \alpha, f(\alpha)\rangle\}) \cup \{\langle \alpha, x\rangle\}$.

If α taken in an context c is true in respect to the instant t and the possible world w, we denote as $\models^{\mathfrak{A}}_{cftw} \alpha$. We will assume that $\mathfrak{A} = \langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, \mathcal{T}, \mathcal{I}\rangle$, as previously defined. Now, consider that $c \in \mathcal{C}$, f an assignment function as above, $t \in \mathcal{T}$ and $w \in \mathcal{W}$. Then:

- If α is a variable, then $|\alpha|_{cftw} = f(\alpha)$
- $\models_{cftw} \pi\alpha_1...\alpha_m\beta_1...\beta_n$ if and only if $\langle |\alpha_1|_{cftw}...|\beta_n|_{cftw}\rangle \in \mathcal{I}_p i(t,w)$
- If η is not " I " or " Here ", then $|\eta\alpha_1...\alpha_m\beta_1...\beta_n|_{cftw} = \mathcal{I}_\eta(t,w)(\langle |\alpha_1|_{cftw}...|\beta_n|_{cftw}\rangle)$ if none $|\alpha_j|)cftw...|\beta_k|_{cftw}$ is †; otherwise, †.

Formally, Kaplan offers the following truth-conditions to some interesting well-formed formulas of **LD** (truth-conditions for first-order formulas and modal formulas in **LD** are given in the usual way)[46]:

- (i) $\models_{cftw} F\alpha$ if and only if $\exists_{t' \in \mathcal{T}}$ such that $t' > t$ e $\models_{cft'w} \alpha$

  (ii) $\models_{cftw} P\alpha$ if and only if $\exists_{t' \in \mathcal{T}}$ such that $t' < t$ e $\models_{cft'w} \alpha$

  (iii) $\models_{cftw} G\alpha$ if and only if $\models_{cf(t-1)w} \alpha$
- (i) $\models_{cftw} N\alpha$ if and only if $\models_{cfc_tw} \alpha$

  (ii) $\models_{cftw} A\alpha$ if and only if $\models_{cftc_w} \alpha$

  (iii) $\models_{cftw} Y\alpha$ if and only if $\models_{cf(c_{t-1})w} \alpha$

It could seem that Y and G are the same; indeed, in **LD**, $\models (Y\alpha \leftrightarrow G\alpha)$. However, Y is us a demonstrative while G is a temporal operator [45]. Saying something about an event in the past such as "One day before the election I was sick" will only be equivalent to "Yesterday I was sick" if happens that *today is the election day*. As can see, Kaplan's **LD** provides truth-conditions for temporal statements that are more complex than Prior's.

Intuitively, α in a context c is *true in respect to t and w* if and only if the sentence expressed by α *-in-the-context-c* is true at the instant t if w is the actual world [45]. We denote the content assigned to a sentence α as $\{\alpha\}^{\mathcal{U}}_{cf}$. Formally, α is true in the context c in a structure $\mathcal{U}$ if and only if for *every* assignment f, $\{\alpha\}^{\mathcal{U}}_{cf}(c_t, c_w) =$ true.

An interesting distinction made by Kaplan is the distinction between *Stable Character* and *Stable Content*. If α is a term or a formula in **LD**, the content of α in a context c in a structure $\mathcal{U}$ is said to be a *stable content* if and only if for every assignment f we have that $\{\alpha\}^{\mathcal{U}}_{cf}$ is a constant function. On the other hand, if α is a term or a formula in **LD**, the character of α is the function that assigns to each structure $\mathcal{U}$, each assignment f and each context c in $\mathcal{U}$, $\{\alpha\}^{\mathcal{U}}_{cf}$. If α is a term or a formula in **LD**, the character of α is said to be a *stable character* if and only if for every assignment f and every structure $\mathcal{U}$ the character of α is a constant function. Intuitively, a formula has a stable character if and only if has the same content in every context [45].

For Kaplan [45], a *context* can loosely be described as possible occasion of use while *circumstances of evaluation* are known as *possible worlds* in *Possible Worlds Semantics* and *instants* in *Tense Logics*. In **LD**, a possible context is $\langle \mathcal{U}, c\rangle$ [46]. In **LD**, a formula α is *logically true* if is true in every index of every structure [45]: this sense could also be expressed by the modal formula $\Box\alpha$. The truth-conditions to indexicals takes in consideration the notions of *true-at-a-time* and *true-at-a-world*. Kaplan's system requires that the agent of any context c exists in the world of c and is located at the time and position of c in the world of c. Similarly for "now" , given the "token-reflexive" aspect of some indexicals. The formal system **LD** is sometimes accused of being outdated but shows us that there is a problem with dealing with indexicals in the usual way. Since it does not provide a way of fixing a point of reference, the particularities of Y and G, for instance, are not adequately contemplated. Thus, a formal system to properly deal with indexicals (especially temporal ones) should be able to explicit the distinction between them.

As we can see, Kaplan offers a formal system to deal with indexicals. However, not only Kaplan was worried about the formalization of some indexical expressions. The indexicals "I" and "Now" aroused the interest of other philosophers and logicians because of their unique features. Usually, in Classical Logic (Propositional and First-Order), statements that do contain "now" are considered the same sort of statement that does not contain "now" . The same treatment comes from the necessity of having formal temporal systems to deal with different statements involving temporal expressions. This topic sometimes leads people to think that the occurrences of "now" might be vacuous [44], perhaps because in many cases, the use of the present continuous conjugation of verbs could substitute the use of the word "now" in a sentence. But concluding that *all* occurrences of the word "now" are vacuous is an inductive conclusion.

Kamp highlights that a formal analysis of "now" will take into account other temporal operators; moreover, according to him, many of the interesting non-vacuous occurrences of "now" are in sentences containing other temporal operators [44]. Mellor [66] prefers to call indexicals such as "now" and "I" *token-reflexive*, since they make the truth of a sentence depends on facts about those expressions. To add McTaggart's famous argument about time, Mellor [66] states that there is no such time as “now” (since time is unreal): reality cannot include a time or a property (of being now, for example) because that would lead to a contradiction (which is not welcome informal systems that wish to preserve classical properties). Someone could say that “now” could be a definite description; however, even if we do not have access to the definite description for the word “now” , we still understand that if someone uses “now” to refer to a specific instant t [66].

Therefore, “now” is quite different from temporal modifiers as “past” or “future” , and it is even more evident that “now” is not a date. Unfortunately, Tense Logics does not consider the unique features of “now ” in the system. Indeed, adding N to a system increases the expressivity power, as we can see from **LD**, but not in Tense Logic.

At first, it could seem that Prior did not give much attention to those temporal indexicals in his Tense Logic. However, he was already concerned with those expressions and their token-reflexive character: he suggests that this “apparent token reflexiveness” would be deceptive [81]. Although Prior’s and Kaplan’s approaches differ, both do not provide solutions to some kinds of paradoxes and situations involving different instants of utterance and instants of evaluation of the utterance. One example is given by the *Electronic Secretary Machine Paradox*, presented as follows:

If I record the utterance “I am not here now “in a message on my electronic secretary machine, according to Kaplan’s theory, the content would be fixed at the time of the utterance; however, this message is supposed to be evaluated in other times.

Unfortunately, **LD** does not allow us to work with the different instants on those contexts. Partly because there is a restriction to only *proper contexts* and partly because it does not provide a way to fix a point of reference. Nevertheless, this could find a possible treatment within a particular class of formal systems: the so-called *Hybrid Logics*. Prior’s work inspired those formal systems, and they are exciting systems to deal with temporal information formally and by Reichenbach’s work on the semantics of tense [14]. Moreover, Blackburn [14] suggests using Hybrid Logics to deal with Temporal Indexicals excitingly.

3.3 Temporal Indexicals and Hybrid Tense Logic

Prior’s *Tense Logic* is concerned with relations between two times: the time of the speech, when the sentence is uttered and evaluated, and the time at which the event described in the sentence takes place [67]. However, for many sentences, it is vital to consider a third time, which Reichenbach [14] called the *point of reference*.

A fascinating aspect of Reichenbach’s theory of tenses is the introduction of three particular times: the *point of speech*, which is the instant at which the sentence is uttered; the *point of event*, which is the instant at which the eventuality described in the sentence takes place; and the *point of reference*, which is not necessarily identical to the point of the event [14]. The *point of reference* is crucial in natural language because it allows us to talk about events independently of both the time of the utterance and the time we want to refer to, as exemplified in the case of discussing election day.

This distinction between points of speech, event, and reference helps formalize how temporal information is handled in language. For example, consider the detective Jane in the wilderness. At time t_1, she believes that the crime happened near the west edge of the forest (the point of event). However, at time t_2, she receives new information placing the crime near the east side. When she revises her belief, she is considering both the time of the utterance (the point of speech) and the time of the event (the point of event). The *point of reference* allows her to update her belief about the crime's location even though her beliefs are situated in different temporal contexts.

The points of speech and event fit Prior's formalization: prefixing a formula with the temporal operator P, for example, locates the point of the event in the past relative to the point of speech. Conversely, prefixing a formula with the temporal operator F locates the point of the event in the future relative to the point of speech [14].

In this context, *Hybrid Logics* emerge as an extension of Modal Logic. This extension involves adding propositional symbols called *nominals*, such that each nominal is true relative to exactly one point in Possible Worlds Semantics, which can be understood as an instant [16]. Hybrid Logics also incorporate satisfaction operators (@) and binders. Binders bind nominals, which can then have *free* or *bound* occurrences.

Semantics of Hybrid Logic The language of Hybrid Logic includes: propositional symbols such as $p, q, r, \ldots$; nominals such as $a, b, c, \ldots$; satisfaction operators (@); binders ($\forall, \downarrow$); connectives ($\wedge, \vee, \neg, \rightarrow, \leftrightarrow$); and modal operators ($\Box, \Diamond$) [17]. The semantics of Hybrid Logic is provided in terms of Kripke's Possible Worlds Semantics. Informally, a nominal a is *true relative to a point* w if and only if the reference of a is identical to w, while a *satisfaction statement* $@_a\alpha$ is *true relative to a point* w if and only if α is true relative to the reference of the nominal a [16]. For example, $@_a a$ asserts the reflexivity property. We say that a given formula is *pure* if its only atomic subformulas are nominals [43].

Hybrid Logic can also be translated into first-order logic with equality through truth-preserving translations. However, there are many reasons to prefer Hybrid

Logic (and Modal Systems in general) for philosophical and computational purposes [43].

A *Hybrid Tense Logic* is obtained by adding the temporal operators G and H (as Prior defined them) in place of the modal operator $\Box$. Specifically, we define $H\alpha$ as $\neg P \neg \alpha$ and $G\alpha$ as $\neg F \neg \alpha$. Hybrid Logic addresses some of the limitations of Tense Logic, particularly by considering the point of reference, which could be crucial for solving problems related to temporal indexicals.

It is intuitive that the indexical "now" picks out exactly the time of the utterance as content. Blackburn [14] suggests a hybrid approach to deal with four temporal indexicals: *now*, *yesterday*, *today*, and *tomorrow*. According to Blackburn, hybrid systems enable us to capture special features of temporal indexicals and their interaction with tenseness. Moreover, "now" can be considered as a special nominal that is true at a contextually determined utterance time, while "yesterday," "today," and "tomorrow" can be considered as propositional constants true at an aligned chain of points of time around this particular point [43].

Referring to Kaplan and Kamp, Blackburn [14] argues that the Indexical Hybrid Tense Logic represents the convergence of Hybrid Logic and *Kamp-Kaplan Semantics*.

To obtain an Indexical Hybrid Tense Logic, it is first necessary to add the atomic symbol *now* and then take into account a *contextual model* to deal with the particular features of "now." A *contextual model* is a 5-tuple $\mathcal{M} = \langle T, R, V, C, \eta \rangle$, where $\langle T, R \rangle$ is an ordinary modal frame, V is a valuation function, C is a non-empty set of contexts, and η is a mapping from contexts to points in T [43]. As Blackburn states, the function η is crucial: it specifies, for any context $c \in C$, what the time (or temporal location) of any utterance in that context is. It tells you, for any context, what your *now* moment is. Thus, η is exactly what Kaplan calls the *character* of *now* [43].

Since indexicals are context-sensitive, we say that a formula α is *satisfied in the context c* when α is satisfied in c at the utterance time of c, taking into account the time of the utterance. Moreover, a formula α is said to be *contextually true in a model* $\mathcal{M}$ if it is satisfied in every context of the model. “Now”, in that sense, is not logically valid but is contextually valid [43]. This aligns with Kaplan’s theory: “now” could never be *uttered* [3] falsely in a given context.

The language can be extended with three other symbols: $yesterday$, $today$, and $tomorrow$. Like "now," these are temporal indexicals. However, they are not nominals. Blackburn [43] considers them propositional symbols denoting sets of points positioned in the model relative to the instant of utterance. These symbols

[3] The key here is to suppose that it is necessary to have a context for an utterance to occur, including a time.

occur in formulas, not as subscripts of @ since they are not nominals.

The functions $yesterday$, $today$, and $tomorrow$ map contexts to specific sets of times [43], and there are constraints to guarantee that the sets are ordered correctly in respect to the context, the utterance time $\eta(c)$, and each other, while ensuring that they respect their property of being sufficiently "daylike" [43]. The structure is given *locally* and does not impose any global requirements on the accessibility relation R. Formally, considering R as an accessibility relation, for all $c \in C$, $yesterday(c)$, $today(c)$, and $tomorrow(c)$ are subsets of T such that:

- (i) $\eta(c) \in today(c)$;
- (ii) $\eta(c)$ is an R-successor of every point in $yesterday(c)$;
- (iii) $\eta(c)$ is an R-predecessor of every point in $tomorrow(c)$;
- (iv) $yesterday(c)$, $today(c)$, and $tomorrow(c)$ are pairwise disjoint;
- (v) Every point in $yesterday(c)$ R-precedes every point in $today(c)$;
- (vi) Every point in $today(c)$ R-precedes every point in $tomorrow(c)$;
- (vii) Every point in $yesterday(c)$ R-precedes every point in $tomorrow(c)$;
- (viii) $yesterday(c)$, $today(c)$, and $tomorrow(c)$ are all R-convex;
- (ix) If $t \in yesterday(c)$ and $t \in today(c)$ and tRs and sRt, then either $s \in yesterday(c)$ or $s \in today(c)$;
- (x) If $t \in today(c)$ and $t \in tomorrow(c)$ and tRs and sRt, then either $s \in today(c)$ or $s \in tomorrow(c)$.

Therefore, Indexical Hybrid Tense Logic incorporates the temporal indexicals *now*, *yesterday*, *today*, and *tomorrow* into Hybrid Logic. Blackburn [43] presents results for this system, including completeness theorems. The literature on hybrid logic offers an extensive list of results, and, according to Blackburn, it is possible to adapt these general results to deal with systems containing temporal indexicals such as "now" [43]. Hybrid systems are syntactically less complicated than **LD** while offering more expressive power than Tense Logic, though they are limited to temporal indexicals.

In summary, a formal system should incorporate the specific features of temporal indexicals to properly handle indexical expressions occurring within a sentence. Indexical Tense Hybrid Logic provides a promising formal system capable of distinguishing the different temporal points of discourse and handling the distinction between *contextually valid* and *logically valid* expressions. This approach is consistent with current debates about the structure of time [43] and allows us to engage with metaphysical claims about time while considering Reichenbach's theory of tense.

Extensions of First-Order Logic could be expressive and might allow us to formalize temporal information and some features of temporal indexicals. However, First-Order Logic is undecidable, presenting challenges from a computational perspective [14]. For these reasons, it seems preferable to opt for systems based on Modal Logic, such as Hybrid Logics. Moreover, from an ontological standpoint, realists about time may prefer Modal Tense Logic to make explicit the *earlier-later relations*.

3.4 Logic of Indexicals

Kaplan's **LD** is restricted to *proper contexts*, and that does not always correspond to conversational situations. Radulescu [87] presented a new formal approach to deal with indexicals. One of his main arguments is that Kaplan's Logic of Demonstratives is too limited to proper contexts and, therefore, unable to deal with the context shifts usually associated with indexicals. Suppose, for instance, the indexical "I" is used in two distinct propositions within an argument. Then, Kaplan's **LD** would force us to exclude this argument. Soames [87] agrees with Kaplan and claims that we should not allow contexts to change within the same argument.

However, changing contexts is common in natural language and when expressing one's beliefs, especially when we need to revise our belief set. The fact that we use indexical expressions to utter propositions in conversations, for instance, is what leads Radulescu [87] to propose the development of a new formal system: the *Logic of Indexicals*.

When reasoning about indexicals, it is crucial to consider the proposition itself and its context. However, Radulescu [87] claims that we cannot formally keep the context fixed in order to apply classical logical rules to propositions containing indexical expressions. Indeed, even in **LD**, some changing context situations might be dealt with by just changing the indexical expression to match the referring content. For instance, the day changes, so we move from "today" to "yesterday". For Radulescu [87], the contexts are intimately connected to the utterances since they are the environments and conditions in which some agent utters sentences. As in Kaplan's semantics, they are sequences of parameters for each indexical. The Logic of Indexicals is thought to be used in conversational environments, so a *context* is thought to be (S,D,W), where S is the speaker, A is the addressee, D is the day, and W is the world [87].

Definition 3.4.1 *An argument is a sequence of the form* $\langle[c_1,\varphi_1],[c_2,\varphi_2],...,$

$[c_n, \varphi_n]\rangle$, where $n \geq 1$, and each pair $[c_i, \varphi_i]$ is a pair of a context and a sentence. Given an argument, the sequence of its $\langle \varphi_1, \varphi_2, ..., \varphi_n \rangle$ is its conversational thread. The sequence of its contexts $\langle c_1, c_2, ..., c_n \rangle$ is its conversational situation.

It is supposed that in conversational environments, there are some similarities between the contexts and some rules governing those similarities [87]. For instance, the idea of successive days should help keep track of the conversational thread. At this point, it seems that some background knowledge is required. Radulescu [87] claims that "from all the relations that may hold between conversational situations, we must pick only the ones who matter for the relevant sense of validity". To that, the restriction is added that the worlds must stay constant; no changes in the possible worlds are contemplated. Indeed, since the contexts should be proper, i.e., both speaker and the addressee must exist in the world of the context on the day of the context, hypothetical scenarios and counterfactuals are not part of his system. Thus, the notion of the conversational situation must be defined as proper.

Definition 3.4.2 *A conversational situation $\langle (S_1, A_1, D_1, W_1)$, (S_2, A_2, D_2, W_2),..., $(S_n, A_n, D_n, W_n)\rangle$ is proper iff $(\forall i, j \in [1, n]) W_i = W_j \wedge$ (S_i and A_i exist at W_i on D_i).*

Because of these restrictions, the Logic of Indexicals also does not provide a solution for "paradoxes" such as the *Answering Machine Paradox*. By the time the addressee hears the utterance, the speaker is no longer present in the conversational situation. The Logic of Indexicals also deals with time as structured successive integers aligned, although it does not provide a possibility of fixing a point of reference, as in the Hybrid Logics. However, to account for the fact that conversational situations might be slightly different, Radulescu [87] introduces the notion of *similar conversational situations*. In particular, "we will require that, for two conversational situations to be similar, it must be the case that for any two contexts in one conversational situation if the difference between the days of those contexts is *n*, the difference between the corresponding contexts in the other conversational situation must also be *n*" [87]. Moreover, similarity requires identities and non-identities between speaker and addressee to be preserved.

Definition 3.4.3 *Two proper conversational situations $\langle c_1, c_2, ..., c_m \rangle$ and $\langle c'_1, c'_2, ..., c'_n \rangle$ are similar iff*

(a) $m = n$;
(b) $(\forall i, j \in [1, n]) A_i = A_j \leftrightarrow A'_i = A'_j$;
(c) $(\forall i, j \in [1, n]) S_i = S_j \leftrightarrow S'_i = S'_j$;
(d) $(\forall i, j \in [1, n]) A_i = S_j \leftrightarrow A'_i = S'_j$; and
(e) $(\forall i, j \in [1, n]) D_i = D_j \leftrightarrow D'_i = D'_j$.

That is intuitive to see in the case of temporal indexicals, with the notion of the passage of time. Temporal operators like "one week ago", "one day ago", etc., work as similarity relations.

Definition 3.4.4 $[c_1, \varphi]$ *is a logical truth iff for any* $\langle c'_1 \rangle$ *which is similar to* $\langle c_1 \rangle$*,* $[c'_1, \varphi]$ *is true.*

One exciting thing about this logical truth notion from the Logic of Indexicals is that while in Kaplan's Logic of Demonstrative the sentence has to be true in all possible contexts to be a logical truth, in **LI**, a sentence is said to be a logical truth if it is true in all contexts similar to that particular context [87]. Finally, we can define the validity of an argument in **LI**:

Definition 3.4.5 *An argument* $\langle [c_1, \varphi_1], [c_2, \varphi_2], ..., [c_n, \varphi_n]$*, where* $[c_n, \varphi_n]$ *is taken to be the conclusion, is valid iff for any conversational situation* $\langle c'_1, c'_2, ..., c'_n \rangle$ *similar to its conversational situation, if* $(\forall i \in [1, n-1])\varphi_i$ *is true in* c'_i*, then* φ_n *is true in* c'_n*.*

As in **LD**, any sentence containing an indexical expression in **LI** is evaluated with respect to a given context. However, in **LI**, the context is defined as the parameters in which the sentence was uttered, while the requirement for an utterance to occur is not shared with Kaplan's logic. In order to express a proposition, then, the Logic of Indexicals express the pairs between a sentence and a context. If a sentence represents a proposition, that is, a set of possible worlds, those pairs are similar to Lewis's centered possible worlds approach that will be discussed in more detail in the next chapter.

Radulescu [87] dismisses the issue about self-knowledge, as in the Rudolf Lingens example, to be a problem of the agent missing knowledge of a particular type: he does not know what the participants in a conversation are expected to know, for instance. Therefore, it is assumed that the speaker and the addressee share a background knowledge that allows them to follow the conversational thread. That view seems to be supported by Lewis's [57] *scorekeeping* idea, on which Lewis supposes that for any conversational situation, a certain amount of background knowledge is required and presuppositions play an important role.

What is interesting about the Logic of Indexicals is that Radulescu [87] introduces a logical treatment towards Communication with Indexical expressions. This everyday use is relevant when talking about indexicals since we frequently appeal to those words to express beliefs. Nevertheless, it does not consider that the contexts keep changing in conversational situations. For instance, one restriction of **LI** is that the instants only count as full days [87], which makes temporal indexicals such as "now" equivalent to "today". That is not the case for every situation.

Perhaps we must consider the insights we can extract from those formalisms towards an adequate approach to communication and testimony issues involving indexical statements. Each system dealing with indexical expressions is limited to a particular application, which is very interesting. However, our purpose is to investigate a logical system capable of dealing with how we can revise our beliefs in the presence of indexical expressions.

Both Kaplan's and Radulescu's logical systems fail when dealing with the so-called "Answering Machine Problem". Kaplan's **LD** is restricted to *proper contexts* $\langle x, p, t, w \rangle$ such that x is an individual located at a spatial position p at the instant t in the world w and do not provide a point of reference distinct from the point of evaluation [45]. In Radulescu's **LI**, it is required that both the speaker and addressee exist in the world w and in the conversational situation, which is not the case.

Deferred Utterances are exactly those utterances that are written or recorded to be read or listened to in a different context [92]. In many scenarios, those utterances can employ indexical expressions such as "I am not here now". Indeed, there are many examples of deferred utterances. For instance, when a professor leaves a note on the door saying 'I am not here now,' the instant of the utterance is not intended to be the same context of evaluation. The sentence is true for a given student that goes to the professor's office and learns that the professor is not there at that instant. That example goes against the apparent certainty that the utterance of the sentence "I am here now" is always true [79].

Predelli [79] argues that to get out of the Answering Machine Problem, it is important to consider the *intended context of interpretation*, which is not necessarily the same as the context of utterance. Moreover, to interpret some indexical expressions correctly, one should consider those intended contexts to be those relevant and accessible to the speaker and the addressee [13]. That requirement is aligned with Radulescu's notion of similarity between contexts and with Lewis [58] treatment for counterfactuals. To interpret possible intended contexts, one should consider some sort of ordering on possible worlds or even in the beliefs an agent holds towards the similarity between contexts. The issue is how to incorporate indexical beliefs into a theory since it should take into account beliefs about the content of the proposition and beliefs about the related contexts. Following Lewis [58], it is not necessarily the same as the context of utterance. Moreover, to interpret some indexical expressions correctly, one should consider those intended contexts to be those relevant and accessible to the speaker and the addressee [13]. That requirement is aligned with Radulescu's notion of similarity between contexts and with Lewis's [58] treatment for counterfactuals.

3.5 Indexical Beliefs

What exactly would count as an *indexical belief* is not well defined yet. As Stalnaker [99] points out, believing is an attitude towards propositions and the proposition's truth-value that does not vary or depend on contextual parameters. To understand how one can hold beliefs about context-sensitive propositions, we must first seek to define the nature of indexical beliefs. One possible approach is that context-sensitive sentences ultimately refer to a unique proposition. This approach, however, can only be done once the context is known and the content of the sentence is fixed. Thus, it involves a two-step process rather than a direct one, as the context must be clarified before a belief about its truth-value can be formed.

The *Two-Dimensional Semantics* is used to characterize context-sensitive expressions and takes into account both contexts of use and circumstances of evaluation [95]. In Kaplan's **LD**, for example, sentences have their truth-value relative to a context [93]. Beliefs involving context sensitivity raise a problem about how static the context is. Since time is always changing and agents can always move, fixing a context depends on our beliefs about where it is located in a possible world, which is not always straightforward. This introduces a further challenge: while we may hold beliefs about indexical propositions, these beliefs themselves might change as the context shifts.

One aspect of the **LD** presented by Radulescu [87] is that it does not provide a way to deal with situations where the agent fails to track what is assumed in a conversation. As the context shifts, so too must our understanding of the propositions involved, but in real-world conversational situations, agents often fail to maintain an accurate account of the context. This leads to questions about how such failures impact indexical beliefs, particularly in scenarios where an agent may be unaware of contextual shifts.

To address these issues, Schulz [96] proposed a framework for dealing with the dynamics of indexical belief. Schulz focuses on how agents maintain beliefs about context-sensitive propositions when they struggle to track the relevant context. Agents are often uncertain about their exact location in both time and space. For example, suppose an agent says something about their day, unaware that it is past midnight and technically the next day. Schulz [96] raises the question of how our indexical beliefs are affected when we cannot keep track of all the relevant contextual information. This issue is especially relevant when dealing with time changes, but it can also apply to other contextual parameters, such as spatial location or the presence of others.

Schwarz [97] highlights that as we exist in the world, we are constantly required to update our beliefs in response to our changing location, both spatially and temporally.

The challenge lies in ensuring that these updates occur in a way that reflects the current context of the agent. Schulz's proposal takes a probabilistic approach to address this challenge, utilizing *conditionalization* and *distance semantics*. These approaches provide a quantitative method for handling the shifting nature of indexical beliefs, particularly self-locating beliefs.

As we explore this topic further, we will revisit the details of probabilistic approaches for the problem of indexical belief, especially self-locating beliefs, in later chapters. These methods, which aim to quantify the uncertainty inherent in context-sensitive Belief Revision, will help refine our understanding of how agents manage belief changes when confronted with shifting contexts. The key insight here is that indexical beliefs cannot always be treated as static; rather, they are inherently dynamic and must be handled as such in any formal theory that seeks to model Belief Revision and context-sensitivity.

3.6 The Debate over the Essential Indexical

The problem of the *essential indexical*, as formulated by John Perry, highlights the indispensable role of indexical expressions in rational deliberation and action. The debate centers on whether indexical reference can be fully captured by standard propositional attitudes or whether a distinct kind of self-locating content is necessary. This discussion is particularly relevant for frameworks of *Belief Revision*, where agents must update their beliefs in light of new self-locating information. Traditional accounts of belief, rooted in Fregean or Kaplanian semantics, often assume that all propositional content can be captured without appeal to indexicals. However, Perry's challenge suggests that certain beliefs—those involving terms like "I," "here," and "now"—cannot be reduced to non-indexical descriptions without loss of crucial cognitive significance.

In Perry's influential example, a messy shopper realizes that he himself is the one causing the mess in a supermarket. The shift from recognizing "Someone is making a mess" to "I am making a mess" is not just a change in linguistic expression but a fundamental change in belief that enables action. This distinction challenges classical accounts of propositional attitudes that rely on the assumption that beliefs can be fully expressed in non-indexical terms. Perry argues that *self-locating beliefs* are essential for reasoning and decision-making, as they provide information that no amount of purely descriptive content can substitute.

David Lewis further develops this idea through *centered possible worlds*, where an agent's belief state is represented not just as a set of possible worlds but as a set of *centered worlds*, each including an agent's temporal and spatial position. This perspective aligns with the distinction made in this thesis between *K-intensions* and *S-intentions*. *K-intensions* (Kaplan-style) correspond to information that remains stable across different contexts, while *S-intentions* (Stalnaker-style) account for world-relative self-locating content that cannot be reduced to non-indexical descriptions. We will explore them later.

The challenge of self-locating belief becomes particularly evident in *Belief Revision* scenarios. Suppose an agent is uncertain about their location and receives a piece of information such as "You are in Room 101." This statement, while true in a given possible world, has no significance unless the agent relates it to their own self-location. Updating a belief system to accommodate self-locating information requires more than just incorporating new propositional content—it necessitates a shift in *S-intentions*.

In standard *AGM Belief Revision*, belief states are typically represented as sets of possible worlds, and revision involves selecting a minimal change that preserves consistency. However, when indexicals are involved, a purely possible-worlds-based approach is insufficient. Instead, Belief Revision must operate over *K-contexts* and *S-contexts*, where K-contexts encode general background knowledge, and S-contexts encode self-locating knowledge. The interaction between these two layers of representation is crucial for a dynamic and flexible update mechanism.

Two-dimensional semantics offers a structured way to account for the role of indexicals in reasoning. In this framework, a statement like "I am here now" has two levels of meaning:

- **Primary intension (K-intension)**: A general function mapping possible contexts to truth values.

- **Secondary intension (S-intension)**: A function mapping possible worlds to truth values, relative to an agent's actual location in that world.

This distinction clarifies why Belief Revision involving indexicals requires a two-dimensional approach. If an agent wakes up in an unfamiliar environment, their beliefs about their identity and location are not just a matter of evaluating which possible world they are in but also updating their self-locating position within that world. Thus, *Belief Revision must accommodate shifts in both K-intensions and S-intentions simultaneously*. The debate over the essential indexical has generated multiple responses. Some theorists, such as Kaplan, argue that while indexicals play a special role in belief, they do not require a fundamentally different semantic treatment. Instead, they propose that indexicals can be understood through structured propositions, where the meaning of "I" includes an implicit reference to the speaker.

Other philosophers, including Stalnaker and Lewis, emphasize that indexical beliefs are inescapably tied to an agent's self-locating knowledge. This is where the notion of *centered worlds* becomes crucial. According to Lewis, an agent's beliefs are best represented as sets of centered possible worlds, where each world is annotated with an individual's spatial and temporal location. This model captures the indexicality problem but raises new issues regarding how agents should update beliefs in dynamic environments.

An alternative approach, developed in contemporary epistemic logic, involves *hyperintensionality*. Proponents argue that standard modal logics fail to capture the fine-grained distinctions required for indexical reasoning. Instead, hyperintensional models allow for a richer representation of cognitive content, distinguishing between beliefs that are necessarily equivalent but cognitively distinct. Moreover, applications of two-dimensional semantics extend beyond epistemology into artificial intelligence and cognitive science. In AI, understanding how agents process self-locating information is crucial for designing autonomous systems that adapt to environmental changes. Similarly, cognitive scientists study how humans navigate indexical references in natural language processing and decision-making.

The debate over the *essential indexical* remains central to discussions in the philosophy of language, epistemology, and Belief Revision. Perry's challenge demonstrates that self-locating beliefs play a fundamental role in rational deliberation, a claim reinforced by the necessity of two-dimensional semantics in representing belief dynamics. By distinguishing between *K-intensions and S-intentions*, we can better model Belief Revision in context-sensitive settings, where an agent's reasoning is not only about *what is the case in the world* but also about *where and when they are in that world*. This integration of indexicals into Belief Revision theory provides a more nuanced framework for understanding rational belief updates in dynamic and uncertain environments.

3.7 Perry's Famous Examples: The Messy Shopper and Rudolf Lingens

One of John Perry's most well-known examples is that of the *Messy Shopper*, which he uses to illustrate the essential nature of indexical expressions in belief and action. The story unfolds in a supermarket, where an individual notices a trail of sugar spilling from a shopping cart. Concerned, he follows the trail, attempting to identify the person responsible for the mess. After some time, he comes to a striking realization: **he himself** is the one causing the mess.

This shift in perspective—from believing that "someone" is making a mess to the realization that "I am making a mess"—marks a fundamental transformation in the nature of belief. Perry argues that this change is not merely a substitution of one belief with another but a recognition of the *self-locating* nature of certain beliefs.

The agent's ability to act—to adjust the cart and stop spilling sugar—depends on this indexical shift. Without it, the knowledge that "someone is making a mess" remains insufficient for taking appropriate action. This example underscores Perry's central claim that indexicals such as "I," "here," and "now" play an ineliminable role in our reasoning and decision-making processes [73].

Another crucial example introduced by Perry is the case of *Rudolf Lingens*, an amnesiac lost in the Stanford library. Lingens reads a highly detailed biography about himself, learning extensive information about the life and characteristics of a certain Rudolf Lingens. Despite having all this descriptive content available, he does not come to the realization that **he himself** is Rudolf Lingens. The problem illustrated here is that while Lingens possesses a complete set of non-indexical facts about himself, he lacks the self-locating knowledge necessary to connect these facts to his own perspective. Perry uses this example to argue that beliefs about identity and self-location cannot be entirely captured by descriptive or referential content alone. Instead, indexicals provide a unique kind of cognitive significance that is crucial for practical reasoning [76].

Both of these examples demonstrate that indexical expressions are indispensable for understanding certain types of belief updates and actions. While traditional theories of belief and reference, such as Fregean descriptivism and Kaplan's theory of demonstratives, attempt to explain meaning without relying on indexicals, Perry's cases suggest that such attempts are insufficient. The knowledge that "Rudolf Lingens is in the library" or that "someone is making a mess" does not equate to the indexical realizations "I am Rudolf Lingens" or "I am making a mess." These insights have profound implications for areas such as epistemic logic, belief revision, and theories of self-locating knowledge. They suggest that formal systems for belief representation must account for both descriptive content and the role of indexicals in reasoning.

3.8 Further Considerations on Indexials

Understanding *indexical beliefs* is central to addressing the challenges posed by context-sensitive expressions in reasoning and Belief Revision. An *indexical belief* refers to a belief that is directly tied to a context-sensitive expression, such as "I," "here," or "now." These beliefs depend heavily on the context in which they are formed, and understanding how they evolve requires us to examine both the context of the utterance and the context in which the belief is held. As Stalnaker [99] points out, belief is generally considered an attitude towards propositions and their truth-values, but these truth-values do not typically vary based on contextual parameters. This creates a challenge when we try to understand how one can hold a belief about context-sensitive propositions. One possible approach is that context-sensitive sentences ultimately refer to a unique proposition, but only once the context is known and the content of the sentence is fixed. This makes the process of Belief Revision a two-step process: first identifying the context and then revising the belief accordingly.

The Two-Dimensional Semantics [95] provides a useful framework for dealing with context-sensitive expressions. This theory accounts for both the context of use and the circumstances of evaluation, which is particularly relevant when working with indexical expressions like “now” or “I.” In Kaplan’s **LD** system, sentences have their truth-values relative to a context [93], but beliefs involving context sensitivity complicate matters, especially when the context itself is not fixed. Time is constantly changing, and agents can move between different contexts, making it difficult to pin down a single context in which to evaluate an indexical expression. For example, an agent’s belief about “now” is often tied to their position in time and space, and keeping track of these changes requires continuous adjustment of their belief set.

One particular challenge arises when agents fail to track the changes in contexts, a scenario that Radulescu [87] highlights in his formal approach to indexicals. Radulescu notes that traditional systems, like Kaplan’s **LD**, struggle with the changing nature of context, especially in conversational environments. He proposes that we must allow for the possibility that contexts can shift even within the same argument. This is particularly important in everyday conversations where indexical expressions such as “I,” “here,” and “now” are constantly in flux. Radulescu’s solution to this problem involves defining *similar conversational situations* to account for these shifts. This idea of context shift is central to understanding how beliefs are revised in natural language.

The Lingens example, where an amnesiac, Rudolf Lingens, is unable to recognize his surroundings or his identity until he utters “I am Rudolf Lingens” after hearing a biography of himself, provides a compelling illustration of the role of context in belief formation. Lingens’ belief about “I” is inextricably tied to the context in which he is situated. Without access to the correct context, his belief about his identity remains unclear. This scenario parallels the situation of an individual lost in the wilderness. Just as Lingens’ belief about his identity requires him to access the proper context, someone lost in the wilderness must be able to identify their location in relation to the context of their environment. In both cases, the belief about “I” and “here” requires the individual to regain access to the correct context to revise their belief.

This brings us to the concept of *contextual models*, which are formal representations of the context in which an utterance occurs. As noted by Radulescu [87], a context is typically represented as a tuple of parameters, such as (S, D, W), where S is the speaker, D is the day, and W is the world. In these models, the context plays a crucial role in determining the truth-value of indexical expressions. As Kaplan's theory asserts, for indexicals to function properly, they must be evaluated within a context that provides the necessary parameters for reference.

In formal systems like Hybrid Indexical Logic [43], which extends Kaplan's work by introducing the concept of a *point of reference*, the focus is on cases where the context shifts within the same argument. This hybrid approach is especially useful for handling temporal indexicals like "now," "yesterday," and "tomorrow." Hybrid Indexical Logic allows us to handle situations where the context changes dynamically, using a more flexible model to accommodate shifts in time and space. This approach draws on Prior's [81] Tense Logic, which provides the foundational operators for handling time-related indexicals.

A further consideration in this chapter is the role of Belief Revision in contexts where indexicals are involved. As agents' contexts change, so too must their beliefs, and this is particularly true for beliefs involving temporal indexicals. Radulescu [87] has proposed a formal system for dealing with conversational dynamics, where beliefs are revised based on the shifting context of the conversation. This dynamic nature of Belief Revision, especially in relation to time and space, highlights the challenges agents face when trying to update their beliefs in response to changes in context.

Finally, the concept of *indexical beliefs* has important implications for Belief Revision. While the formal systems we discussed, including Kaplan's and Radulescu's, provide mechanisms for handling indexicals, they do not fully address the issue of how beliefs about shifting contexts are formed and revised. Schulz [96] proposed a probabilistic approach to this problem, where the distance between contexts is used to determine how beliefs should be updated. This approach emphasizes the importance of understanding how contexts relate to one another, particularly in conversational situations where contexts shift dynamically.

Therefore, the study of indexical beliefs and their role in Belief Revision provides valuable insights into how context influences our reasoning processes. As we move forward, we will investigate different definitions of *contexts* and explore how centered possible worlds can be used to formalize the relationship between contexts and beliefs. By incorporating insights from Hybrid Indexical Logic and other formal systems, we can better understand how beliefs involving indexicals evolve as contexts change.

Chapter 4

Contexts and Centers

Contexts are a central concept for *Indexical* theories. However, there are many possible interpretations of what a *context* is and what it stands for. In this chapter, we will explore various formal interpretations of *context* and how contextual parameters can be represented in terms of *centers*. Moreover, we will discuss the shifting nature of contexts and the issue of context change in conversational environments, as raised by Radulescu [87].

The role of context in indexical theories cannot be overstated. The understanding of indexicals expressions like “I”, “here”, or “now” is inseparable from the context in which they are used. Context defines how these indexicals pick out their referents. For example, the sentence “I am here now” changes its meaning depending on who is speaking, where they are, and when the statement is made. The truth of such a sentence is always relative to the context of utterance. Thus, the role of context in the interpretation of indexicals is foundational, and understanding how context shifts and changes is essential for a robust theory of indexicality.

A key challenge in the philosophical literature on context is that there is no single, universally accepted definition of what a context is. Different philosophers have proposed various definitions and perspectives, with some focusing on the speaker's intentions, others on the external environment, and still others on the relationship between the agent's beliefs and the world. One influential view is Kaplan's [45] notion of context, where context is defined as a "possible occasion of use" of a sentence, meaning the particular situation in which a sentence is uttered. For Kaplan, the context includes all the parameters necessary to determine the reference of indexicals. However, Kaplan's approach is based on the idea of *proper contexts*, situations in which the parameters for reference are clearly and unambiguously defined. This perspective works well when the context is fixed and stable, but it does not fully account for situations where context shifts over time or across different conversational scenarios.

Stalnaker [99], on the other hand, offers a different perspective on context. For

Stalnaker, context is not a static, fixed entity, but rather a dynamic and evolving set of assumptions that can change as the conversation progresses. This view of context as fluid and dependent on the background knowledge of the speaker and addressee allows for a more flexible interpretation of indexicals. In Stalnaker's framework, the context of utterance is shaped by the beliefs, intentions, and assumptions of the participants in the conversation, which are constantly evolving as new information is introduced. This makes context not only an external feature of the conversation but also a product of the participants' mental states and shared knowledge.

The *Lingens Example* provides a useful illustration of the role context plays in Belief Revision. In this example, Rudolf Lingens, an amnesiac, is lost in the Stanford Library. He reads about himself in a biography, but his knowledge of who he is and where he is located is limited. He is unable to locate himself until he uses the indexical expressions "I" and "here" to refer to himself and his surroundings. His belief about his situation depends entirely on the context in which he uses these indexicals. The shifting context of his understanding from being confused to finally locating himself highlights how beliefs about one's situation can change depending on the context in which indexicals are used.

Similarly, the *Alone in the Wilderness* scenario demonstrates how an agent's belief about their situation evolves as the context shifts. When an individual is lost in the wilderness, they may initially believe they are near safety, but as they move and gather more information, their belief about their location may change. The shifting context both internal (beliefs about the world) and external (location and time) leads to a revision of the agent's beliefs. This example parallels the challenges faced when trying to maintain a consistent belief set in a dynamic, context-sensitive environment.

Both the *Lingens Example* and the *Alone in the Wilderness* scenario underscore the importance of context in belief formation and revision. The key philosophical debate surrounding context is how it can shift and what role this shifting context plays in the interpretation of indexicals. Context is not just a passive background against which statements are made; it actively shapes the meaning of sentences and the truth conditions of propositions. The shift in context can change not only the meaning of indexicals but also the beliefs and knowledge of the agent using those indexicals.

This is where the debate between Kaplan and Stalnaker becomes particularly relevant. Kaplan's *Logic of Demonstratives* (LD) relies on the idea of *proper contexts*, where all the parameters necessary for determining the reference of indexicals are fixed. This works well for simple, stable contexts but becomes problematic when dealing with contexts that shift or are not fully defined. Stalnaker, by contrast, provides a more dynamic view of context, where context is a product of the interaction between the agents involved in the conversation and the evolving background knowledge they share. This allows for a more flexible approach to the interpretation of indexicals, especially in situations where context is fluid and subject to change.

4.1 K-Contexts and S-Contexts: A Proposal

Understanding the role of context in indexical reasoning is fundamental to the study of semantics, Belief Revision, and decision-making. However, the notion of context is not easily defined, and different approaches have been proposed to account for the dynamics of context-sensitive expressions. In this work, I introduce two formal definitions of context: *K-contexts* and *S-contexts*. These definitions provide a structured way to model the interaction between an agent's cognitive state and external situational parameters, facilitating a nuanced understanding of how indexical reasoning operates within different epistemic and environmental conditions.

The complexity of defining context arises from its dual role in communication and cognition. Context serves both as a parameter for interpreting utterances and as a structure that evolves dynamically during discourse. Philosophers such as Kaplan [46] and Stalnaker [102] have proposed different models for understanding context. Kaplan conceptualizes context as a structured tuple of parameters, including the agent, time, and location, which determines the content of indexicals. Stalnaker, in contrast, views context as a *common ground*—a set of propositions mutually accepted by conversational participants.

While these perspectives offer valuable insights, they fail to capture the full complexity of indexical reasoning. Kaplan's model does not require an utterance to take place, making it insufficient for modeling dynamic discourse contexts. Stalnaker's model, on the other hand, does not fully account for the formal mechanisms by which an agent's self-locating knowledge evolves in response to external changes. To bridge these gaps, I propose *K-contexts* and *S-contexts*, two complementary but distinct notions that integrate both epistemic and environmental dimensions of context.

A *K-context* is a formal representation of an agent's cognitive perspective at a given moment. Formally, a K-context is defined as:

$$K = \langle w, t, p, i, B \rangle, \tag{4.1}$$

where:

- w is a possible world,

- t is the time of evaluation,
- p is the agent's spatial position,
- i is the agent,
- B is the agent's belief set at $\langle w, t, p, i \rangle$.

K-contexts are crucial for modeling how agents process indexical expressions such as "I," "here," and "now." The agent's beliefs play a central role in interpreting these expressions, as the truth conditions of self-locating statements depend on the information accessible to the agent at a given moment. For example, in the *Lingens Example*, an amnesiac named Rudolf Lingens navigates the world without knowledge of his own identity. His utterance "I am here" is evaluated relative to his K-context, which initially lacks the crucial self-identification component. As he acquires new information, his belief set B is updated, leading to a transformation in the content of his indexical statements. This dynamic shift underscores the necessity of K-contexts in modeling Belief Revision for self-locating propositions.

While K-contexts represent the agent's cognitive state, *S-contexts* capture the broader situational environment in which an utterance occurs. Formally, an S-context is defined as:

$$S = \langle W, T, P, I, C \rangle, \tag{4.2}$$

where:

- W is the set of possible worlds compatible with the conversational background,
- T is the temporal frame of discourse,
- P is the spatial domain of reference,
- I is the set of participants involved in the discourse,
- C is the common ground—a set of propositions shared by conversational participants.

S-contexts extend beyond the agent's individual perspective to incorporate the shared background knowledge that governs conversational coherence. Unlike K-contexts, which focus on the internal cognitive state of an agent, S-contexts model the external parameters that shape the interpretation of indexical expressions within a social and environmental framework.

Consider the *Alone in the Wilderness* scenario, where an agent utters "I am lost." In this case, the agent's belief system (modeled by their K-context) interacts with the broader situational constraints (modeled by the S-context). The agent's perception of being lost depends not only on their internal state but also on external conditions such as geographical markers, weather conditions, and the availability of navigational tools. The S-context provides a structured representation of these external influences, facilitating a more comprehensive analysis of indexical reasoning. The distinction between K-contexts and S-contexts extends beyond formal semantics into areas such as artificial intelligence, epistemic logic, and philosophy of action. For instance, in epistemic modal logic, Belief Revision mechanisms must account for both the agent's self-locating knowledge (modeled by K-contexts) and the dynamic changes in conversational settings (modeled by S-contexts). In artificial intelligence, context-aware systems must integrate structured contextual parameters to refine decision-making processes, aligning with the dual roles captured by K- and S-contexts.

Moreover, a significant challenge in formal models of context is the temporal evolution of belief states. Since context-sensitive beliefs change dynamically, a model integrating K-contexts and S-contexts must account for the continuous updating of both individual knowledge and shared common ground. A promising direction for further research involves developing a context-sensitive dynamic logic, where context updates are formally characterized as structured operations on both K-contexts and S-contexts. The definitions of K-contexts and S-contexts provide a formal foundation for addressing key challenges in indexical reasoning and Belief Revision. By explicitly modeling both cognitive and environmental dimensions of context, this framework offers a systematic approach to analyzing context-sensitive expressions.

From a formal semantics perspective, K-contexts provide a rigorous structure for evaluating the truth conditions of indexical statements. Meanwhile, S-contexts facilitate a dynamic representation of conversational progress, ensuring that Belief Revision processes align with the evolving informational landscape. The integration of these two perspectives allows for a more comprehensive understanding of indexicality, bridging gaps between traditional semantic models and contemporary theories of context dynamics. As we proceed, we will explore how these formal definitions interact with existing frameworks, including Kaplan's *Logic of Demonstratives* and Stalnaker's *common ground* approach. By synthesizing these insights, we aim to develop a unified theory of context that captures both the individual and collective dimensions of indexical reasoning.

4.2 K-Contexts and S-Contexts in Indexical Reasoning[1].

The distinction between *K-contexts* and *S-contexts* offers several significant advantages when it comes to formalizing the role of context in indexical reasoning. By separating the two notions of context, we can more precisely capture the complexities and nuances of context-sensitive expressions and better account for the dynamic, sometimes uncertain nature of the contexts in which communication occurs.

One of the primary benefits of distinguishing between K-contexts and S-contexts is the ability to address the issue of unclear or uncertain contexts. In natural language, it is common for speakers and listeners to operate within contexts that are ambiguous or not fully specified. This is particularly evident in situations where time, space, or the identity of the speaker may be unclear or shifting. For instance, in the *Lingens Example*, an agent's understanding of where they are and when they are located may change over time, leading to uncertainty about the context. Similarly, in the *Alone in the Wilderness* scenario, the agent's belief that they are lost may evolve as new information about their surroundings becomes available. By distinguishing between K-contexts (which are centered around the agent's knowledge and beliefs) and S-contexts (which are tied to external, environmental factors), we can better capture the dynamic nature of these uncertainties.

K-contexts allow us to formalize how an agent's beliefs about their situation evolve over time, while S-contexts provide a framework for understanding how external factors (such as time, place, and environmental cues) influence the interpretation of context-sensitive expressions. This separation clarifies whether uncertainty arises from the agent's internal belief system (e.g., their knowledge or lack thereof) or from the external situation (e.g., ambiguous temporal or spatial factors). K-contexts and S-contexts also allows us to model how context can shift over the course of a conversation or thought process. As Grice [36] pointed out in his

[1]The distinction between K-contexts and S-contexts is used throughout this work to articulate two structurally distinct roles that context can play in the modeling of belief and meaning. A K-context, following Kaplan, refers to the referential tuple $\langle$agent, time, place, world$\rangle$ that determines the character-to-content mapping for indexical expressions. It serves to fix the circumstances of utterance and the parameters relevant for referential interpretation. In contrast, a S-context, building on Stalnaker's conception of context as common ground, denotes the epistemic space within which propositions are evaluated and beliefs are revised. It consists of the set of possible worlds compatible with what is presupposed or mutually assumed in a reasoning or discursive situation. While these notions are conceptually distinct, their interaction becomes salient in cases involving perspective-sensitive belief revision, particularly in the presence of indexical uncertainty or self-locating information. The terminology is employed with a degree of flexibility across chapters, but the guiding distinction remains: K-contexts fix referential parameters, whereas S-contexts provide the evaluative backdrop for belief dynamics. This layered structure also suggests potential for further developments involving multi-agent reasoning, conversational scenarios, and multi-centered belief models, where epistemic agents may coordinate, revise, or negotiate their positions across intersecting informational contexts.

work on conversational implicature, much of communication depends on the shared context between speaker and listener. However, as contexts shift—whether due to changes in the agent's knowledge or in the external environment—the meaning of indexical expressions can change as well. For example, the phrase "I am here" may have different meanings depending on whether the agent is speaking in a familiar location (a known K-context) or in an unfamiliar, shifting environment (an S-context). By using K-contexts to model the agent's knowledge and S-contexts to account for the external situation, we can more accurately track these shifts and understand how they affect the interpretation of indexical expressions.

Moreover, the separation of these two types of context helps resolve instances of ambiguity that arise in natural language. In a conversational scenario, an indexical expression may be ambiguous due to shifts in both the agent's beliefs and the external situation. By distinguishing between K-contexts and S-contexts, we can specify which aspects of the context contribute to the ambiguity and how the meaning of an utterance might change as the context evolves. For instance, in the *Alone in the Wilderness* scenario, the agent's belief about their location may be uncertain (a K-context issue), but the shifting natural environment (an S-context issue) also contributes to the uncertainty of the agent's situation. By separately modeling these two sources of uncertainty, we can gain a clearer understanding of the complexities of context-sensitive reasoning.

A further advantage of distinguishing between K-contexts and S-contexts is the ability to formalize Belief Revision in uncertain contexts. When agents encounter new information that challenges their existing beliefs, they must revise their belief sets accordingly. However, when the context itself is uncertain or unclear, as in the *Lingens Example*, it becomes more difficult to determine how to update beliefs and which information should be incorporated into the agent's belief set.

By formalizing the distinction between K-contexts (the belief system) and S-contexts (the external situation), we can better model how agents revise their beliefs when confronted with shifting contexts. Stalnaker [102] has argued that successful communication depends on the shared context between speaker and listener, and that context change is central to how we revise our beliefs. By separating K-contexts from S-contexts, we allow for a more nuanced approach to context change, one that accounts for both internal shifts in belief and external changes in the situation. This enables us to formalize Belief Revision in a way that respects the complexity of context-sensitive reasoning. As Stalnaker [102] notes, "The common ground is a set of propositions that both parties to the conversation take to be true, and this set can evolve as the conversation progresses." By distinguishing between the agent's internal knowledge (K-context) and the external circumstances (S-context), we can model how these changes affect the truth of indexical expressions and the revision of beliefs.

The philosophical debate surrounding context is rich and complex, with numerous views on how context shapes meaning and how it should be represented in formal systems. Some philosophers, such as Stalnaker [102], argue that context is primarily a matter of the shared common ground between participants in a conversation. This perspective emphasizes the importance of mutual understanding and shared knowledge in communication. Others, such as Kaplan [45], suggest that context is more dynamic and is tied to the parameters of the conversation, such as the time, place, and identity of the speaker.

The distinction between K-contexts and S-contexts allows us to reconcile these differing views by formalizing both the shared knowledge between agents (K-contexts) and the external situational factors that influence interpretation (S-contexts). This dual approach provides a more flexible and comprehensive model of context, one that can accommodate the complexities of real-world communication, including the uncertainties and ambiguities that arise in everyday conversations. By introducing this distinction, we are able to better address issues of contextual uncertainty and ambiguity. We gain a more robust framework for understanding how agents navigate shifting contexts, revise their beliefs, and interpret indexical expressions in dynamic situations. This framework is crucial for advancing our understanding of indexical reasoning and Belief Revision, as it allows us to account for both the internal and external factors that shape how we understand and engage with the world.

4.3 Intended Contexts

Intended Contexts introduce an important layer of complexity in the discussion of contexts, particularly when it comes to understanding how context-sensitive expressions should be interpreted. Typically, when someone utters a sentence like "I am here now," the proposition is considered true within the context of the utterance—meaning that the truth of the sentence depends on the location and time of the speaker's utterance. According to Kaplan's [45] restriction to proper contexts, the sentence "I am here now" would be true when uttered in the speaker's actual location and time, while "I am not here now" would be false in the same context. However, this understanding changes when we encounter written sentences or recordings, as in the case of a note left on a door or a recorded message. Here, the context of utterance and the context of evaluation may diverge.

As Predelli [79] points out, there is an *intended context of interpretation*, which refers to the context in which the speaker intends the message to be evaluated by the addressee. For instance, if a professor leaves a note on her office door saying, "I am not here now," she is not intending to inform the students that she is not in her office at the time the note is written. Instead, she expects the students to read the note at a later time, and thus the context of interpretation is different from the context of utterance. Predelli [79] highlights this important distinction:

> When I wrote my note, I intended that the sentence 'I am away today' be interpreted with respect to a plurality of contexts of interpretation, thereby managing to express a manifold of contents. In particular, I intended that the sentence on my note be interpreted with respect to contexts differing from each other (and from the context of encoding) with respect to their temporal coordinates appropriately. [79]

This case illustrates how the context of utterance and the context of evaluation are not necessarily the same. This introduces the need for a clearer distinction between the context in which the sentence is produced (the context of utterance) and the context in which it is interpreted (the intended context of interpretation). The key question then arises: how should we formalize this notion of intended context?

In terms of formalizing intended contexts, there are two possible approaches, depending on whether we focus on the external context (S-contexts) or the internal context (K-contexts). On the one hand, *S-contexts* would require that the addressee is familiar with certain propositions that help define the intended context, such as knowledge about the professor's office hours or other background information. This approach emphasizes the social and environmental aspects of the context, where external knowledge helps determine how an utterance should be interpreted. On the other hand, *K-contexts* focus on the agent's internal belief system and the context in which they are operating. In the case of the professor's note, a K-context approach would focus on the agent's knowledge of the content of the note, as well as the temporal and spatial information that the speaker intends to convey, such as when and where the note is meant to be read. K-contexts allow for a more subjective approach, where the context of interpretation depends on the mental states of the participants and how they understand the message within their belief systems.

Gauker [33] argues for a more Kaplanian approach to contexts, suggesting that the context of an utterance is not just defined by the speaker's intentions but also by objective parameters that shape the interpretation of indexical expressions. According to this view, the context of interpretation is not determined solely by the speaker's intentions but is grounded in a more objective understanding of the world in which the communication takes place. Predelli [80] further extends this discussion by emphasizing that intended contexts should be viewed as formal sets of parameters. These sets allow for a more structured approach to understanding how the context of interpretation is determined. According to Predelli, the intended context of interpretation involves a set of parameters that can vary across different situations, allowing for a more flexible approach to indexical expressions.

One key idea in the formalization of intended contexts is the notion that contexts can be represented as sets of parameters. Bianchi [13] builds on this idea by extending the debate to pragmatic concerns, focusing on the common ground between the speaker and the addressee. In her view, the context of utterance is a set of parameters, while the intended context of interpretation is a set of all possible contexts compatible with the speaker's intentions. This distinction is important because it allows us to understand how the meaning of a sentence can vary depending on the context in which it is interpreted, even if the context of utterance remains fixed.

Intended contexts can also be thought of as a series of similar contexts, which can be ordered in a way similar to counterfactual reasoning. Lewis [58] discusses the idea of counterfactuals in terms of possible worlds, and this approach can be applied to the concept of intended contexts. In this framework, the set of intended contexts can be viewed as a series of worlds or situations that share certain characteristics, with the understanding that some worlds may be more relevant or closer to the actual world than others. For instance, in the case of the professor's note, the set of intended contexts could include a range of possible situations where the note is read at different times, from moments shortly after it is written to moments far in the future. The relevant context for interpreting the note is determined by the time at which the addressee reads it, but the set of possible contexts includes all those where the note is read at some point in the future. Thus, intended contexts allow for a more nuanced understanding of how time and place influence the interpretation of indexical expressions.

The next section will explore how counterfactual reasoning can be further applied to intended contexts, particularly in cases where the speaker's intended meaning interacts with hypothetical alternatives to the actual context. By analyzing intended contexts through the lens of counterfactual structures, we can develop a richer model of how agents navigate shifting contextual parameters and how meaning is reconstructed in different interpretative frameworks. By distinguishing between the context in which something is uttered and the context in which it is interpreted, we can capture the complexity of situations like the *Lingens Example* or the *Alone in the Wilderness* scenario. These examples show how agents may revise their beliefs depending on shifts in context, whether through changes in time, location, or perspective. Ultimately, recognizing the role of intended contexts allows for a more nuanced, flexible approach to interpreting meaning in language, one that accounts for the varying contexts in which words are understood.

4.4 Lessons from Counterfactuals

Lewis's theory of counterfactuals [58] provides a valuable framework for understanding how context-sensitive expressions, such as indexicals, can be evaluated across different possible worlds. Counterfactuals are statements that take the form of "If it were the case that ..., then it would be the case that ...", or "If it were the case that ..., then it might be the case that ...". These types of statements are useful for reasoning about hypothetical scenarios and evaluating the truth of propositions under different conditions. Lewis's approach to counterfactuals involves the use of *accessibility relations* between possible worlds, where the truth of a counterfactual depends on how similar the world in question is to the world of evaluation [58].

A key feature of Lewis's system is the concept of a *sphere of accessibility* around a world. In this framework, each possible world w has an associated set S_w of worlds that are considered accessible from w. These accessible worlds are similar to w to varying degrees, and the degree of similarity determines their position in the accessibility hierarchy. A counterfactual is true in world w if it holds true in all the worlds in S_w that are accessible from w [58]. This allows us to reason about the truth of counterfactuals based on the similarity between worlds, providing a formal structure for evaluating hypothetical situations.

The intuitive idea of *similarity* between worlds is central to Lewis's theory [58]. Roughly, possible worlds that are more similar to w are placed in smaller spheres, while worlds that are less similar are placed in larger spheres. Smaller spheres are thus more tightly connected to w, representing worlds that resemble w more closely in terms of both content and context. For example, in the case of a counterfactual like "If it were the case that I was in New York, I would be in the city that never sleeps," the worlds in smaller spheres would represent scenarios in which the agent is in New York and the situation closely resembles the world being evaluated. Larger spheres, in contrast, would represent worlds in which the agent is in a different location, or in which the world differs in more significant ways from w.

In Lewis's model [58], each world w is associated with a set of accessible worlds S_w, and this accessibility is governed by a set of structural conditions. For example, the spheres must be *nested*, meaning that if one set of worlds Γ is included in another set Δ, then either $\Gamma \subseteq \Delta$ or $\Delta \subseteq \Gamma$. Additionally, the spheres are *closed under unions*, meaning that if Γ is a subset of S_w, the union of Γ should also be a valid set of accessible worlds. Similarly, the spheres must be *closed under nonempty intersections*, ensuring that the intersection of two sets of accessible worlds is also a valid set within S_w [58]. This framework has significant implications for the formalization of counterfactuals. By defining a system of spheres around a given world, Lewis provides a way to structure the relationship between different possible worlds and evaluate counterfactual statements based on their relative similarity to the world in question. The model also offers insights into how we can reason about the past, present, and future, as well as about hypothetical scenarios that are not part of the actual world.

One of the main contributions of Lewis's theory is the concept of *centered worlds* [58]. A centered world is not just a world with a specific set of facts but also includes a particular perspective, such as the agent's location, time, and position within the world. In other words, a centered world is a pair $\langle c, w \rangle$, where c represents the agent's context (including their beliefs, knowledge, and perspective), and w represents the possible world itself. This idea of centered worlds is crucial for modeling context-sensitive expressions like indexicals, which depend not just on the content of a proposition but also on the agent's perspective. The concept of *centered worlds* extends Lewis's notion of similarity to include the agent's perspective [58]. In the case of indexical expressions such as "I am here now," the meaning of the sentence depends not just on the world of evaluation (w) but also on the center (c), which includes the agent's location, time, and beliefs. This dual dependence on both the content of the world and the agent's perspective makes it possible to capture the dynamics of context-sensitive expressions, allowing us to reason about how the meaning of indexicals changes depending on the context.

A crucial aspect of this framework is its application to the theory of *intended contexts*, which was developed in the previous section. Counterfactual reasoning and the structure of accessibility relations between worlds provide a way to formalize how intended contexts operate. Just as counterfactuals allow us to evaluate alternative hypothetical scenarios based on world similarity, intended contexts function as structured interpretations that an agent expects their utterance to be understood within. This is particularly relevant when dealing with indexical expressions that are detached from their original utterance context, such as in the case of the professor's note stating, "I am not here now.". The next section will develop this connection further by exploring how counterfactual reasoning can provide a foundation for modeling intended contexts. Specifically, we will analyze how accessibility relations between possible worlds parallel the way an agent constructs an intended context of interpretation. This transition will allow us to bridge the discussion between counterfactual structures and the broader implications of belief revision, self-locating beliefs, and context-sensitive reasoning.

4.5 Weighted Contexts

It is intuitive to think about similarity in terms of ranking, degrees, or levels. It is possible also to think in terms of *weighted contexts*, where the degree of relevance or plausibility of a given context is considered. Just as counterfactuals involve evaluating possible worlds based on their similarity to the world of evaluation, weighted contexts assign a “weight" or degree of relevance to different contexts, depending on how closely they align with the circumstances under consideration. This approach allows us to handle situations where contexts may not be equally relevant or may involve varying degrees of uncertainty, offering a flexible framework for reasoning in contexts where vagueness or imprecision plays a role. In formal terms, the idea of ordering contexts is not entirely new. Klawonn, Gebbhardt, and Kruse [51] proposed a model to handle contexts within a possibilistic and fuzzy logic framework. Their goal was to provide a system capable of integrating uncertainty or vagueness into a knowledge base. The concept of *Context Logic* aims to assign numerical values to contexts, which can then be interpreted to express uncertainty and degrees of truth [51]. In their model, the authors introduce the idea of a set of *weighted contexts* C_d, where each context $c_m \in C_m$ includes observations about the domain under consideration [51]. In this framework, the *weights* of a context are interpreted as a quantification of the relevance of that context to the overall reasoning process. This perspective aligns with the idea that some contexts are more informative or applicable than others, depending on the situation at hand.

Klawonn, Gebbhardt, and Kruse [51] take contexts in our sense of *S-contexts*, as they define $\mu(c)$ to represent the set of all formulas that are *known* in a given context, with the context evaluation reflecting an expert’s knowledge about a specific domain. This mirrors the idea that context is not just a set of parameters, but also a set of beliefs, assumptions, and knowledge that shape the interpretation of propositions. Instead of specifying exact numerical values, the authors suggest that we order contexts in terms of relevance. This ordering is similar to Lewis’s [58] idea of *nested worlds*, where possible worlds are ranked based on their similarity to the actual world. In the case of weighted contexts, the worlds (or contexts) are ranked based on their specificity and relevance, with more specific or relevant contexts given higher weight. As Klawonn, Gebbhardt, and Kruse [51] explain:

> Only those formulae known for sure are taken into account on a very strict level, but on a more speculative level, more formulae are thought to be possible. In this way, the nested worlds can be weighted accordingly. [51, 1376]

This approach provides a way to account for different degrees of certainty within a context, enabling more nuanced reasoning about context-sensitive propositions. However, it is important to note that the truth-conditions for formulas in Context Logic are not always clearly defined. Instead of using precise truth values, the system assigns a degree of truth or relevance, typically in the range of $[0, 1]$, similar to probabilistic approaches used in dealing with context-sensitivity. This degree of truth reflects the uncertainty and variability inherent in many real-world contexts.

One of the key insights from *Context Logic* is that not all contexts share the same degree of probability or relevance. This idea can be extended to incorporate a hierarchy of contexts, where some contexts are considered more plausible or useful than others, depending on the situation. This hierarchy could be influenced by various factors, such as the proximity of a context to the world of evaluation, the relevance of certain parameters, or the degree of certainty associated with a particular context. In this way, weighted contexts provide a valuable tool for reasoning about indexicals and other context-sensitive expressions, offering a framework for dealing with uncertainty and vagueness in a structured manner.

In formal terms, we can define a *weighted context* as a pair $\langle c, w\rangle$ where c represents the context itself, and w is the weight assigned to that context based on its relevance or degree of certainty.

Definition 4.5.1 (Weighted Context) *A weighted context C is a pair $\langle c, w\rangle$, where:*

- *c is a context (a set of relevant parameters or propositions);*
- $w \in [0,1]$ *is the weight of the context, representing its relevance or degree of certainty.*

The weight w quantifies the relevance of the context in a given situation. A higher weight indicates that the context is more relevant or certain, while a lower weight suggests less relevance or more uncertainty. This model provides a flexible framework for reasoning about contexts where the truth of propositions depends not only on the context itself but also on the certainty associated with it. The *truth* of a formula φ in a given weighted context is evaluated with respect to the weight w assigned to the context. Specifically, the truth of a formula in a weighted context is determined by the degree to which the formula holds in that context, weighted by the relevance or certainty of the context.

Definition 4.5.2 (Truth in Weighted Contexts) *Let $\langle c, w\rangle$ be a weighted context, and let φ be a formula. The truth of φ in the weighted context $\langle c, w\rangle$ is defined as follows:*

$$True(\varphi, \langle c, w\rangle) = \begin{cases} 1, & \text{if } \varphi \text{ holds in } c \text{ with certainty} \\ w, & \text{if } \varphi \text{ holds in } c \text{ with a degree of certainty corresponding to } w \\ 0, & \text{if } \varphi \text{ does not hold in } c. \end{cases}$$

This definition allows us to capture the idea that the truth of a statement in a given context can be modified by the relevance or certainty of that context. In cases where a context is more relevant or certain, the truth of the statement is given a higher value. Conversely, in less certain contexts, the truth of the statement is reduced, reflecting the uncertainty involved. Moreover, the concept of weighted contexts is particularly useful for reasoning about context-sensitive expressions, especially in situations where there is uncertainty, vagueness, or varying degrees of relevance. For example, consider the *Alone in the Wilderness* scenario, where the agent's belief "I am lost" depends not only on the agent's internal belief system but also on the external context of being lost in the wilderness. If the agent's belief is expressed in a high-certainty context (e.g., in an environment where they are familiar with their surroundings), the truth of the statement "I am lost" is evaluated with higher confidence. However, in a more uncertain context (e.g., if the agent is unsure about their location or orientation), the truth of the statement is weighted according to the degree of uncertainty in the context.

Let us take the case of the *Lingens* example, the agent's belief about their identity ("I am Rudolf Lingens") is subject to both the agent's knowledge and the context in which the belief is evaluated. If the context is one where the agent has access to all relevant information (e.g., a situation where they know who they are), the weight of the context is high, and the belief is evaluated with certainty. If, however, the context involves uncertainty (e.g., the agent has no access to the necessary information about their identity), the belief is evaluated with lower certainty. In summary, the framework provided to deal with weighted contexts provides a valuable framework for reasoning about context-sensitive propositions, allowing us to account for varying degrees of relevance and certainty in different situations. By formalizing the idea of context as a weighted set of propositions, we can more effectively model the dynamic nature of context-sensitive beliefs and their evolution over time.

4.6 Centers, Centered Possible Worlds, and Contexts

Having explored the concept of formal contexts, we can now examine the specific characteristics of different types of contexts, particularly in relation to the challenge of handling indexical information within the Belief Revision framework. We have discussed how contexts can be understood in terms of their structural elements, such as the set of possible worlds they involve, and how context-sensitive expressions like indexicals rely on specific contextual parameters. Specifically, we made a distinction between contexts based on their structure and the manner in which they interact with indexical expressions, using examples like those from Kaplan's *LD* and Stalnaker's framework.

In the previous sections, we reviewed various formal treatments of context, including how contexts are typically represented and how they can shift in different conversational scenarios. One notable approach was the distinction between *K-contexts* and *S-contexts*, which helped to clarify how contexts are defined relative to the knowledge of agents versus external factors. We also discussed the role of uncertainty in contexts, and how context-sensitive beliefs can evolve over time as new information becomes available. Given these foundations, we now turn to the notion of centered possible worlds, which provide an essential framework for understanding how contexts can be enriched with specific information regarding an agent's location within a possible world. This distinction is particularly important in the Belief Revision context, where agents must constantly update their beliefs based on changing contextual information. Centered possible worlds, which pair a possible world with an individual's location and time, provide a precise method for representing the agent's self-locating beliefs and how those beliefs might evolve as contexts shift.

Conversational situations involve the dynamic interaction between an agent's beliefs and the context in which they are situated. In the case of indexical expressions, the contextual parameters such as time, location, and the individual uttering the expression are crucial to determining the meaning and truth conditions of statements. For instance, in a sentence like "I am here now," the interpretation of "I" and "here" depends not only on the speaker's knowledge but also on the actual time and place of the utterance. This intertwining of knowledge and context is captured by the concept of *centered possible worlds*. Centered possible worlds extend the classical notion of possible worlds by incorporating a center, which represents the agent's location within the world. A centered possible world is a pair $\langle w, c \rangle$, where w is a possible world and c is a center containing the agent's position, time, and possibly other relevant contextual parameters. This provides a way to represent context-sensitive beliefs that rely not just on the world as a whole but on the specific perspective of the agent within that world.

In Belief Revision, as agents must constantly update their beliefs based on new information about their world and their place within it. The use of centered possible worlds allows us to track changes in beliefs with respect to specific contexts. For example, when an agent revises her belief about her location in the world (e.g., " I am now in Berlin"), this revision involves updating the center of the possible world she is in. The ability to track such changes is essential in modeling Belief Revision processes, especially when the agent's location or time shifts, as is common in real-world scenarios. Furthermore, the flexibility of centered worlds allows us to model scenarios where multiple possible contexts are involved, such as in the case of shifting time or location. This becomes particularly relevant when agents are dealing with indexical expressions whose truth depends on their contextual parameters. By using centered possible worlds, we can better capture the shifting nature of beliefs and how agents must constantly adapt their mental models based on the contexts they encounter.

Centered possible worlds offers a powerful formal tool for modeling the relationship between context and belief in the Belief Revision framework. By distinguishing between contexts and centered possible worlds, we can better represent how agents update their beliefs in response to changes in their environment and the shifting nature of time, location, and other contextual parameters. This approach provides a more nuanced understanding of Belief Revision, especially when dealing with context-sensitive beliefs expressed through indexicals. By incorporating both formal contexts and centered worlds, we have the means to model not only how beliefs are updated but also how agents navigate the complexities of changing contexts. The next section will explore how this framework can be applied to specific examples and further develop the implications for the theory of Belief Revision.

We briefly saw that *possible worlds*, for Lewis [58], are the ways a world can be. Moreover, it is standard to represent propositions as sets of possible worlds. Yet, context-sensitive propositions cannot be represented employing *ordinary possible worlds*. Consider an agent with enough information to rule out all possible worlds except the actual one. If the world has several distinct individuals and/or times, she might still be uncertain of which individual she is or which instant of time it is. That means that in the same world, it is possible to have different states of information: in the same world, she could be at an instant t_1 or an instant t_2, for instance. The agent would need an extra piece of information to locate herself within the world, and a center can give that piece of information. A center could be interpreted as contextual parameters that allow one to locate herself within a possible world. Thus, the content of sentences containing indexical expressions cannot be represented as a set of possible worlds. Instead, we need to resort to the notion of *centered possible worlds*, i.e., pairs $\langle w, c\rangle$ of a center c and a possible world w.

The representation of context-sensitive propositions as sets of centered possible worlds is in line with the traditional approach towards indexicals provided by Kaplan [45]: the definition of $Index$ provided by Kaplan [46] states that contextual factors are represented as an index In of coordinates such as $i = \langle w, t, p, i\rangle$, where w is a possible world, t is an instant of time, p is a spatial position, and i is an individual or agent. As we can see, the structure of an index is similar to our centered world $\langle w, c\rangle$ where c would be interpreted as the contextual parameters such as those in the index. Centered worlds are often used to model self-locating beliefs, and the existence of centered content is motivated to adequately characterize belief states [108]. We can represent contents of self-locating beliefs precisely as centered worlds because they provide the contextual information necessary to locate the individual-time-position in a possible world. Also, the contents of self-locating beliefs are often expressed using indexical expressions and, therefore, by sets of centered possible worlds. Some might argue that it is possible to substitute indexical expressions in sentences representing beliefs with co-referential expressions without any harm. However, that substitution does not always represent the same belief to an agent. Part of that issue is related to the agent's beliefs often being connected to her actions. Suppose two individuals i and j are having dinner at 23:58 (the instant t) on a Saturday night. The individual i states the following sentence:

"It is late *now*."

The belief expressed by (1) locates both individuals at the instant t and, thus, can provoke an immediate action: they could just get up and leave to go home, for instance. Now suppose that the same individual i looks at his clock and states the following sentence:

"23:58 is late."

Another individual j present in the context and who is unsure of which time it is probably will not respond the same way to both statements. For instance, in the case of (2), she might not react at all as she might find the utterance strange and out of context. She will need more information; she will need to locate herself at the instant 23:58 to grasp the belief that is being expressed fully. Intuitively, sentences (1) and (2) do not express the same mental states the individuals might have, although the content of (1) can be viewed as an eternal sentence like (2).

In Perry's terms [78], sentences containing indexicals and sentences containing non-indexical substitutes differ in cognitive significance. But what exactly is the difference between (1) and (2)? On the one hand, the first sentence contains an indexical expression; thus, (1) can only be represented through centered possible worlds. Knowing a given center $\langle i, t \rangle$ such as i being an individual and t an instant is crucial to understanding the sentence's content and evaluating the sentences and attributing them a truth-value. On the other hand, (2) is an eternal sentence and can be expressed by possible worlds. There is no need to know the center. To go from a sentence containing an indexical expression like (1) to a sentence with only non-indexical components like (2), we must show that no matter which center of the world we pick, we still have the same possible world w. Perry's idea of *cognitive significance* is central to his analysis of indexicals and self-locating beliefs [73]. Perry argues that traditional accounts of meaning and reference fail to explain why indexical expressions, such as "I" or "here," provide essential information to the speaker in a way that non-indexical descriptions cannot. His famous *Messy Shopper* example illustrates this point: a person may believe “someone is making a mess in aisle five” without realizing that *they themselves* are that person. The shift from an objective belief (“someone is making a mess”) to a self-locating belief (“I am making a mess”) demonstrates the *cognitive significance* of indexicals. This distinction is crucial for understanding *practical reasoning*, as it shows that an agent’s actions often depend on their self-locating knowledge. Perry’s insights challenge traditional views of propositional attitudes, arguing that self-locating beliefs cannot be fully captured by standard possible-world semantics, since they require a *centered world* approach that accounts for the agent’s perspective within a given world.

Regardless of being standard to represent an ordinary proposition as a set of possible worlds, we must consider that not all possible pieces of information can be identified with this framework, as we can see. However, *centered worlds* are not an easy concept to define: roughly, it is taken to be a pair of a center and a possible world. Then, we need to understand what possible worlds are and what a center represents. Since there is no consensus about what a possible world is in the literature, Liao [61] states that most philosophers who deal with centered worlds leave open the question about what possible worlds are since the notion of centered worlds can be adjusted accordingly to whatever notion of possible world one opts to adopt. We will also use Liao’s view in this work and assume that centers contain the necessary information to locate the individual in a time location in a possible world. Moreover, we will say that a piece of information is *indexical* if it represents centered information, that is, a centered world or the location represented by the center. In contrast, an *ordinary* information is taken as any piece of information or evidence about the world itself, without the centered component.

Centered worlds offer a more detailed and precise way of modeling context-sensitive propositions. By representing context-sensitive information as pairs of centered worlds $\langle w, c \rangle$, we can capture the nuances of self-locating beliefs and indexical expressions, which are otherwise difficult to model using ordinary possible worlds. This framework allows us to better understand how agents locate themselves within a possible world and how their beliefs and actions are influenced by their temporal, spatial, and personal contexts. The distinction between centered and ordinary information is crucial for formalizing indexical reasoning, particularly in the Belief Revision framework, where the agent’s context plays a pivotal role in updating and evaluating beliefs.

Another feature that adds an important layer to the understanding of context-sensitive expressions and Belief Revision is the mechanisms of centered assertions. In the framework of centered worlds, assertions are made not only with respect to the content of a proposition but also from a particular center, which includes the agent's position within a possible world, the time at which the assertion is made, and the agent's knowledge or beliefs. This perspective is essential because it recognizes that the truth of a proposition involving indexical expressions depends not only on the world in which it occurs but also on the specific context in which it is asserted. For example, a statement like "I am here now" is true at a specific time and place relative to the speaker's context. When an agent asserts such a proposition, they commit to the truth of the proposition from their own center, which may be different from the centers of others involved in the conversation. This commitment to a proposition, grounded in the agent's particular center, plays a crucial role in Belief Revision, as agents continuously update their belief systems based on the shifting context of the world around them. Assertions from centered worlds allow us to formalize how agents modify their beliefs in response to new information, taking into account the dynamic nature of both the world and the context in which beliefs are expressed and revised.

4.7 Centered Worlds and Self-Locating Beliefs

Quine's [82] account of centered possible worlds is a fascinating and somewhat surprising contribution to the field of Possible Worlds Semantics. Although Quine is not typically associated with the theory of centered worlds, his work offers a perspective that aligns closely with the ideas of self-location and the way agents experience the world from a specific position within it. Quine's notion of centered worlds arises in his discussion of knowledge and meaning, where he emphasizes the importance of an individual's point of view in interpreting their world. A centered possible world, in Quine's terms, is essentially a possible world that is viewed from a particular individual's perspective at a specific time. This is akin to what we now recognize as a centered world in contemporary modal logic, where the center is defined by a specific individual and a particular instant in time. The surprise in finding that Quine used the expression "centered possible worlds" is due to his traditionally more austere, empiricist stance, which often downplays the need for metaphysical constructs like possible worlds. However, in his work, Quine implicitly acknowledges that understanding any world even a possible one requires grounding it in the perspective of a particular observer.

In Quine's own words, he explores the idea of centered worlds through a concrete example, often referred to as the Cat Example. In this example, Quine describes a scenario in which he observes a cat. The cat, in this case, is an entity within a possible world, but to understand the cat fully, one must be situated within a particular perspective: that of the observer. The observer's perspective is crucial for interpreting the cat's existence, as the cat can only be fully understood when considered from the viewpoint of the individual who perceives it. This example emphasizes the importance of context or the "center" of the world in the interpretation of any given object or event. The cat example illustrates how, in Quine's framework, an individual's position within a world (a center) shapes the interpretation of everything within that world. Thus, knowledge about any object is never purely objective or detached from the observer; rather, it is always situated within a specific context, marked by the individual's location and perspective.

Quine's [82, 43] use of centered worlds challenges us to consider the limitations and biases inherent in any knowledge claim. It also sets the stage for a broader philosophical exploration of self-location and the subjectivity of experience. While Quine did not explicitly formalize centered worlds in the same way as Lewis or Kaplan, his insights provide a foundational understanding of the role of perspective in interpreting possible worlds. This emphasis on the individual's center of reference contributes significantly to later developments in modal and Belief Revision theories, where the center plays a crucial role in understanding how agents navigate and revise their beliefs about the world. As he [82, 43] points out, "A cat may be at a given place, but it will not do to say that it exists only at that place unless we are willing to specify the time at which it is there." This highlights the importance of specifying both the location and the time in any given belief or statement, aligning with the idea of centered worlds where both an individual and a time are necessary to fully understand any proposition.

Moreover, as Liao [61] highlights, Lewis argued in favor of a centered world as a way to capture self-locating contents of attitudes such as belief. The Quinean account towards centered worlds inspired Lewis's approach. To Quine, a center is a set of space-time coordinates, while to Lewis, a center is a set of an individual and a time. Both assume that the center is an ordered set and suppose that even if we rule out all other possible ways a possible world can be, there is still a need to capture a specific location within a possible world. That is why Stalnaker's approach might not be sufficient to represent self-locating beliefs. We could borrow Milano's [68] interpretation of both accounts and define them as follows:

Definition 4.7.1 (Quinean Centered Worlds) *Centered worlds are a pair $\langle c, w \rangle$, where w is a traditional possible world and c is a point within a system of coordinates R.*

Definition 4.7.2 (Lewisian Centered Worlds) *Centered worlds are a pair $\langle c, w \rangle$, where w is a traditional possible world and c is a pair $\langle i, t \rangle$ of an individual i and an instant t.*

Since the Lewisian account identifies centers considering the individuals on them, we will focus our approach on Lewis's interpretation of what a center is. Like

Lewis, we will also adopt the notion that a belief state of an agent can be identified as the set of all propositions (sets of centered possible worlds) that the agent believes in. This shift of approach is connected to our topic because of some intuitive assumptions. First, we take agents as humans and, thus, imperfect reasoners. We are often unaware or uncertain about pieces of information about what the world is like. Sometimes, we are unaware or uncertain about our location within that world. As Milano [68] states, "we often lack complete information about what the world is like."

We call *self-locating uncertainty* when that lack of complete information is about our position within a possible world or what role we play in it. Even when we have all relevant information about what the world is like, that does not necessarily imply that we will be certain about our center. Indeed, there are lots of examples in which that occurs. A famous one in the literature is Perry's [74] *Messy Shopper*:

> I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally, it dawned on me. I was the shopper I was trying to catch. [74, 125]

What puzzles us with the messy shopper example is that the agent does not learn something about the world. He actually learns his own position on it by realizing who he is. His beliefs, then, changed not because he faced a new content of a standard uncentered proposition, but instead, he acquired information about a centered component and the fact that he was at the center. Beliefs concerning centered propositions might be challenging, but we must also consider the role that centers play in successful communication. In the case of the *Messy Shopper*, for instance, someone could have said to him, "You are making a mess." How he processes that information and how that changes his behavior is something that models of communication concerning indexicals are trying to explain. This case illustrates how the agent's lack of awareness about their own position in time and space can cause them to misinterpret or fail to act on crucial information until they locate themselves.

The shift to recognizing the center in communication is crucial for the success of the conversation. In cases of self-locating beliefs, where indexicals such as "I" and "here" come into play, the center is integral for understanding how the agent is situated within the world, both in terms of their self-location and how that relates to their communication with others. Models of communication that incorporate centered worlds help us better understand how agents adjust their beliefs, especially when those beliefs are influenced by the agent's own positioning in time and space. This understanding also provides valuable insights into how agents revise their beliefs. Just like the *Messy Shopper*, an agent's beliefs may not change due to new external information but because they have come to understand their own position within the world. In communication, this involves navigating not only the external facts of the world but also the internal, centered aspects of the agent's own perspective. This dynamic interaction between the self and the world is central to both Belief Revision and effective communication, particularly in the realm of indexical information.

In that sense, centered assertions and the understanding of self-location are indispensable when modeling communication and Belief Revision. The example of the *Messy Shopper* serves to illustrate how centered worlds provide the crucial information needed to understand an agent's beliefs and actions, especially in the context of indexical expressions. When agents recognize their position within the world, they can adjust their beliefs accordingly, leading to more effective communication and a clearer understanding of how Belief Revision works in real-world situations. This dynamic interaction between the agent's center and the world is essential for building a comprehensive theory of communication and Belief Revision in the context of indexicals.

4.8 Centered Assertions and Communication

Radulescu [87] emphasized the crucial role of indexicals in communication, particularly in conversational situations. Since indexicals are represented by means of centered worlds, a deeper understanding of centered assertions becomes essential. After all, the challenges related to revising indexical information are inherently connected to how that information is obtained, processed, and communicated. Centered assertions offer a framework to explore how agents navigate the dynamics of communication where their position in space and time is constantly influencing the way they convey and interpret meaning. This framework is directly related to Belief Revision since, in order to revise indexical beliefs, we first need to understand how those beliefs are formed and shared. Centered assertions not only clarify what centers are, but they also highlight their importance in the revision of beliefs, especially those that are context-dependent.

The importance of understanding centered assertions stems from the fact that agents do not simply revise beliefs based on the content of information alone; their revision is heavily influenced by their location in a possible world, their individual perspective, and the way these parameters interact with the information being communicated. For example, when agents communicate using indexical expressions, they are conveying not just propositions but their own position within those propositions. Indexicals like "I am here" are inherently tied to the speaker's self-location. Consequently, the meaning of an assertion containing such expressions cannot be fully understood without considering the agent's perspective.

A practical way of understanding how centered assertions function can be drawn from the work of Perry [78], who explored the difference between sentences containing indexicals and those without. Perry pointed out that sentences involving indexicals have a distinct cognitive significance for the agent. Consider the example from Perry's *Messy Shopper*:

> I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally, it dawned on me. I was the shopper I was trying to catch. [74, 125]

This example vividly illustrates the role of centered assertions in shaping beliefs. The agent's beliefs did not change because they gained new factual information about the world. Instead, the change occurred because the agent became aware of their own position within the world they realized they were the shopper they had been trying to find. This realization demonstrates a profound shift in their belief state, which was driven by self-locating information. This shift highlights the distinct nature of beliefs related to self-location: agents revise their beliefs not simply based on new objective information, but because they update their understanding of their own position in the world.

The Messy Shopper example emphasizes that, when revising beliefs, agents must update their centered world where they are located in time and space because these centered beliefs influence how they interpret and react to new information. This realization directly connects with the idea of *self-locating uncertainty*. As Milano points out, “we often lack complete information about what the world is like." [68] Even if we possess full knowledge about the external world, our uncertainty about our position within that world can lead to revised beliefs. This is precisely the role that centered assertions play in communication and Belief Revision. Further, we can think about the Messy Shopper scenario in the context of communication. Suppose a friend tells the shopper, "You are making a mess!" How does the shopper process this information? They might be surprised initially, but upon reflection, they realize that the mess they are referring to is their own. This self-realization shifts the shopper's belief state and might prompt an immediate change in behavior. This example underscores the idea that Belief Revision is not simply about acquiring new knowledge about the world, but about acquiring new knowledge about one's own position within it.

This brings us back to the role of centered assertions in communication. In many communication situations, indexical expressions like "I," "here," "now," and "this" are crucial in determining the meaning of a message. However, for the speaker and the addressee to successfully communicate, they must align their understanding of the centered world their shared knowledge of location, time, and perspective. This is where centered worlds and centered assertions come into play. The speaker's assertion is not simply about the content of the world, but about how they relate to the world, and it is this relational aspect that the addressee must grasp to understand the message fully.

The concept of centered assertions helps bridge the gap between indexical expressions and Belief Revision by providing a way to formally model how agents update their beliefs. This is particularly important in conversational settings, where the agents involved often need to revise their understanding of the world and their own position in it as the conversation progresses. The act of communication is, in a sense, an act of Belief Revision, as agents adjust their beliefs based on the information they receive, which is always colored by their self-location within the conversation. To deepen our understanding, we turn to how centered worlds influence Belief Revision in systems involving indexical expressions. The shift in belief, such as in the *Messy Shopper* scenario, is not just about adding new propositions to one's belief system but involves recognizing the role of one's position in the world and how that influences the interpretation of new information. The agent's Belief Revision occurs because they adjust their understanding of their own location within a set of possible worlds.

To further illustrate this point, consider the following scenario, the *Annoying Bob Example*: Alice and Bob are at a party, and Alice says, "I had a great day today." Bob, noticing the time on the clock, sees that it is 12:05 AM, and the date has changed. He immediately interrupts and corrects her, saying, "You mean yesterday, not today." This simple misalignment reveals a deeper issue in how context-sensitive information is interpreted and underscores the importance of self-location and context-tracking in communication. Alice's use of the word "today" is intrinsically tied to her self-location in time. When she says "today," she is referring to the day she has just experienced, and for her, it holds a specific temporal significance. However, Bob, by observing the clock and noting that it is already past midnight, interprets "today" based on the present moment from his own perspective. For him, it is now "yesterday," and thus he corrects Alice.

The issue arises not from the content of the utterance but from the failure to properly track the context in which it was uttered. Alice's intended context, rooted in her experience of the day, is not aligned with Bob's interpretation, which is based on the clock time. This example highlights the challenge of interpreting indexicals, such as "today," in communication. The meaning of such expressions is inherently context-sensitive, and their interpretation can vary dramatically depending on the perspective of the speaker and the hearer. Alice's intended context is grounded in her personal experience of the day, whereas Bob, by relying solely on his external observation of time, misinterprets her statement. This is a classic case of context failure, where the context of utterance is not adequately tracked, leading to miscommunication.

In a standard Kaplanian K-context, Alice's utterance occurs within a set of parameters, including the speaker, time, place, and the agent's epistemic state. If Bob is in the same conversational environment, the statement is subject to immediate interpretation, potentially influencing how Bob reacts. However, if Alice whispers to her friend in a distant corner of the party, the statement exists in a different K-context, where the relevant parameters exclude Bob's awareness. From an S-context perspective, the broader situational dynamics influence interpretation. If Bob later hears a third party recount Alice's statement, the S-context shifts, as Bob's new epistemic state changes how he processes the utterance. Unlike K-contexts, which are anchored in the agent's belief system at the moment of utterance, S-contexts track evolving social factors and potential shifts in the shared common ground. This distinction is crucial for understanding how information propagates in a dynamic conversational setting.

The example of Alice and Bob highlights how context-sensitive expressions, such as indexicals, require precise tracking of the conversational and epistemic states of the participants. When Alice says, "I had a great day today," she is referencing a temporal period grounded in her subjective experience. However, Bob, observing that it is past midnight, reinterprets her utterance in terms of the objective calendar date, leading to misalignment between their respective understandings. This misalignment underscores the concept of *re-centering*, which occurs when different agents occupy different informational positions and must reconcile their perspectives to maintain effective communication. In a standard Kaplanian K-context, Alice's utterance is evaluated based on the parameters of her own epistemic state, whereas Bob's interpretation arises from a shifted perspective influenced by external cues, such as the passage of time [46].

From an S-context perspective, Bob's reaction is shaped by the broader social and environmental factors that inform his interpretation. If Bob later hears Alice's statement from a third party, his understanding shifts yet again, as the S-context accommodates additional information and the evolving common ground [102]. Moreover, this example also illustrates the role of *centered assertions*—assertions that incorporate an agent's self-location within the world. Alice's statement was not merely a neutral claim about an event but was grounded in her immediate temporal experience. The failure to align her self-locating belief with Bob's external interpretation resulted in miscommunication, demonstrating how belief revision occurs when agents update their understanding in response to context-sensitive expressions. Ultimately, the Annoying Bob scenario exemplifies the dynamic interplay between K-contexts, S-contexts, and belief revision. The ability to track context shifts and re-center interpretations is crucial for effective communication, particularly in cases involving indexicals. By refining our understanding of how context operates in such exchanges, we gain deeper insights into the mechanisms underlying belief formation, revision, and shared knowledge in discourse.

A central issue in the Theory of Assertion is how an utterance's content is communicated and understood within a given conversational context. An assertion typically involves the speaker providing information or making a claim that the addressee must interpret and, if successful, incorporate into their belief system. The aim of assertion is to achieve successful communication, where the addressee not only understands the content of the proposition but also integrates it into their own cognitive framework. Traditional accounts, such as those advanced by Stalnaker [102], posit that an assertion introduces a proposition that should be true in the context shared by both the speaker and the addressee. However, these models face significant challenges when dealing with context-sensitive expressions, particularly indexicals, which are words or phrases whose meanings shift depending on the context in which they are used.

The core of the issue lies in the fact that for an assertion to be successful, the addressee must not only understand the content of the proposition but also correctly interpret its context. This becomes particularly problematic in the case of indexicals, where the speaker's utterance depends on their self-location in the world. For example, when the speaker says, "I am here now," the meaning of the statement is inherently tied to the speaker's present time and location. For the addressee to properly interpret this statement, they must adjust their beliefs to align not only with the proposition expressed but also with the speaker's context. This requires more than just transferring information it demands a shift in perspective or, as Weber [110] argues, a process of recentering the belief system of the addressee to account for the speaker's position.

As Stalnaker [102] explains, the success of an assertion depends on its ability to be incorporated into the conversational context. However, when the assertion involves indexicals or self-locating beliefs, this simple transfer of information becomes insufficient. To address this, models like Weber's [110] Recentering Model suggest that the addressee must not only understand the content of the assertion but also adjust their beliefs to account for the speaker's context. The Recentering Model posits that communication is not just about transferring uncentered content but also involves adapting one's belief system to a centered world that reflects the speaker's self-location. This highlights that successful communication requires more than just transmitting content it requires alignment of both the speaker's and addressee's contexts for mutual understanding. This issue becomes especially evident in scenarios where the speaker and the addressee are not in the same context. For instance, in Perry's [74] "Messy Shopper" example, the agent fails to recognize their own role in a situation, leading to confusion and misinterpretation. This example highlights how self-locating beliefs and the correct interpretation of indexicals depend on more than just understanding the propositional content they require an awareness of the agent's position within the world. Torre [108] emphasizes that assertions involving centered worlds are more complex than those that only involve uncentered propositions. The inclusion of centered worlds in the conversational context set is necessary to properly account for the speaker's self-location, which is vital for the accurate interpretation of context-sensitive expressions and for ensuring the success of communication.

4.8.1 Modelling Communication

Understanding communication requires a framework that accounts for both the successful transmission of information and the challenges posed by context-sensitive expressions. Traditional models of communication, such as the *Transmission Model*, focus on the idea that communication succeeds when a proposition is accurately transferred from a speaker to an addressee. This approach works well for uncentered propositions, where truth conditions remain stable across different perspectives. However, as Weber [110] argues, centered propositions—statements whose content depends on the speaker's self-location, such as "I am sleepy"—introduce complexities that the Transmission Model fails to address. Unlike uncentered statements, centered propositions require the addressee to *recenter* the information relative to their own perspective. That is, the addressee must reconstruct the speaker's meaning in a way that aligns with their own epistemic state and location within the world. This need for perspectival adjustment reveals the limitations of a purely transmissive account of communication and motivates the adoption of a more dynamic framework.

In response, Weber [110] proposes the *Recentering Model*, which acknowledges that the *expressed content* by the speaker is not necessarily identical to the *acquired content* by the addressee. Instead, communication involves a process of contextual alignment, where meaning is adjusted based on shifts in perspective. Before detailing the steps of the *Recentering Model*, it is essential to explore why standard models of communication fail to account for centered propositions and how these failures highlight the necessity of a more refined, context-sensitive approach. . The Recentering Model involves several steps [110], which are briefly summarized as follows:

1 **Perceiving** The addressee perceives (or encounters) the utterance u;

2 **Centering** The addressee believes that the content expressed by u is true of the speaker's centered world;

3 **Locating** The addressee believes that her own centered world is R-related to the speaker's centered world;

4 **Recentering** The addressee infers information about their own centered world from steps 2 and 3.

A key feature of this model is that the addressee does not need to share the speaker's center; rather, the addressee must simply have a relationship to the speaker's centered world. To make sense of an utterance like "I am sleepy ", it is sufficient that the addressee understands her relative position to the speaker, without requiring further information about the properties associated with the speaker's location. This is captured in the following definition [110]:

Definition 4.8.1 *Given two centered worlds* $\langle w, c_a \rangle$ *and* $\langle w, c_s \rangle$ *where the first represents the addressee's centered world and the latter the speaker's,* $\langle w, c_a \rangle$ *and* $\langle w, c_s \rangle$ *are in an R-relation if and only if* $\langle w, c_s \rangle$ *is the centered world where the utterance was produced and that the individual of* $\langle w, c_a \rangle$ *perceives at the time of* $\langle w, c_a \rangle$*.*

Weber's definition suggests that the addressee's understanding of the utterance depends on their ability to relate to the speaker's center. This concept can be extended to consider *intended contexts*, where the addressee can identify the context in which the utterance was produced and adjust their own interpretation accordingly.

The crucial insight of the Recentering model is that the acquired content may not exactly match the expressed content. To clarify, according to the Recentering Model proposed by Weber [109], the expressed content reflects the beliefs of the speaker, while the acquired content is not simply what is expressed, but rather is content relative to the addressee's center. This distinction becomes important when misunderstandings arise, as the addressee may interpret the speaker's centered content in relation to their own centered world, leading to potential revisions in beliefs. Consider the following principles presented by Weber [109]:

1 *Mind-to-Speech Principle*: The content of the utterance and the content of the belief expressed by the speaker are identical.

2 *Speech-to-Mind Principle*: The content of the utterance and the content of the belief the addressee acquires is identical.

In the Recentering Model, the first principle, the mind-to-speech principle, holds, but the second principle, the speech-to-mind principle, does not [109]. Let us examine the steps of the Recentering Model to better understand how this distinction works. Consider a p-individual, which refers to an individual related to the addressee. The *Centering* step can be broken down as follows [109]:

1 *Understanding*: The utterance u is true for the restricted set of p-individuals.

2 *Trusting*: The addressee accepts that u is a true statement from the speaker.

From the Centering step, the conclusion is that the speaker is a p-individual. The Recentering step can then be decomposed as follows:

1 *Centering*: The speaker is a p-individual.

2 *Locating*: The addressee is R-related to the speaker.

Thus, the conclusion is that "the addressee is R-related to a p-individual ". This result is particularly insightful because we often assume that both speaker and addressee share the same world and spatio-temporal coordinates. However, as demonstrated by the *Answering Machine Paradox*, this is not always the case [2].

Moreover, in everyday communication, we often encounter *disagreement* between agents. Those disagreements can take different forms such as believing in different contradictory propositions or honest mistakes and misunderstandings. Specifically, we are concerned with *Centered Disagreement* [109], which is defined as follows:

Definition 4.8.2 (Centered Disagreement) *Two utterances u and v are in disagreement if and only if the belief an agent holds about u's and v's contextual information makes both utterances incompatible.*

Weber [109] highlights that in most communication situations, the content expressed is uncentered. Some argue that it is only possible to express uncentered beliefs, and indexical beliefs should be reduced or restricted to uncentered counterparts. However, this reduction may not capture all the relevant

[2]For instance, we could use the concept of *multicentered worlds* to indicate that individuals are situated in the same context. We will discuss multicentered worlds further below.

context-sensitive features. Thus, we follow Weber [109] and consider *Centered Utterance Content*. We assume that communication occurs between a competent speaker and a competent hearer or addressee. Furthermore, we assume that competent addressees are able to acquire centered information from utterances. The acquisition of centered information is only possible if the content of the utterance is itself centered. Weber [109] concludes that there must be utterances whose content is centered [3].

The treatment we give to self-locating beliefs is deeply tied to the notion of context we assume. Torre [108] recalls that the content of an assertion in Stalnaker's account consists of uncentered propositions, which do not always capture the full range of contextual information necessary for interpreting assertions involving self-location. To address this, Torre proposes a model based on *multicentered worlds*:

Multicentered Worlds Multicentered worlds are centered worlds that contain more than one designated individual. A multicentered world can be represented by an ordered pair consisting of a world w and an n-tuple of individuals within w, allowing for a richer representation of conversational dynamics where multiple agents share a joint perspective.

Building on this framework, Kindermann [49] extends the notion of *multicentering*, where an assertion communicates a multicentered content, representing a property of a group of individuals. This approach helps explain how participants in a conversation jointly locate themselves in the world, incorporating both common ground and individual perspectives.

These insights are relevant to probabilistic approaches to self-locating beliefs. Probabilistic models of communication provide a structured way to analyze the likelihood of different contextual alignments and misunderstandings. In probabilistic frameworks, communication success is often evaluated based on expected information gain, with Bayesian models tracking how agents update their beliefs about their location in the world given new contextual cues. Furthermore, probability-based models allow for an analysis of uncertainty in communication. In particular, situations where misunderstandings arise due to conflicting K-contexts or S-contexts can be captured by assessing the probability distributions over possible interpretations. This probabilistic dimension enriches our understanding of belief revision, particularly in cases where agents face ambiguity regarding their self-location.

The representation of communication using probabilistic approaches provides a

[3] Some might argue that all uncentered content can be associated with a center, thereby turning it into centered content.

bridge between traditional modal logic and contemporary formal epistemology. In the next section, we explore how probabilistic methods can formalize self-locating beliefs and address cases such as the Sleeping Beauty Problem, where agents must update their degrees of belief based on evolving context-sensitive information. These connections allow us to see how self-location is not only relevant in belief revision but also plays a crucial role in how agents distribute and process uncertain information.

Part III

Probabilistic Insights

Chapter 5

Probabilistic Approaches and Self-Locating Beliefs

Centered Possible Worlds approaches towards changing one's mind with context-sensitive information are already present in probabilistic frameworks. These frameworks attempt to model how agents revise their beliefs in response to new information, and much of the existing work in this area is based on *updating by conditionalization*. Conditionalization is a process by which an agent revises their probability distribution over possible worlds (or outcomes) upon receiving new evidence, ensuring that their beliefs remain coherent and consistent with the evidence at hand. This process has become a standard tool in the formalization of Belief Revision, particularly in probabilistic and epistemic contexts.

The basic idea behind conditionalization is that when an agent receives new information, they adjust their belief distribution in such a way that the new evidence is incorporated as a constraint, while maintaining the consistency of their previous beliefs. Formally, the conditional probability is updated based on Bayes' theorem, which allows the agent to adjust their belief state given new observations. In the case of Belief Revision using centered possible worlds, the agent's belief about their location within a possible world (i.e., their center) is updated according to the newly acquired evidence, much like how the probability distribution over worlds is updated in probabilistic frameworks.

We will review what these probabilistic proposals suggest in the context of centered possible worlds and Belief Revision, particularly focusing on whether they can be considered as a quantitative counterpart to our investigation into Belief Revision with context-sensitive information. For instance, in the probabilistic framework, the Belief Revision process is often modeled as a probabilistic update that respects the structure of conditional dependencies among different pieces of information. In contrast, centered possible worlds provide a more qualitative perspective, where the agent's location in time, space, and their self-awareness are crucial for determining how beliefs should be updated.

Furthermore, while the probabilistic approach works well for formalizing how agents update their beliefs based on external observations or changes in context, it may not fully capture the nuances of how beliefs are revised when agents are confronted with changes in their own self-location or the shifting perspectives in communication. We will explore the connection between these two frameworks, analyzing how the mathematical rigor of conditionalization can complement the more conceptual models offered by centered possible worlds, and whether this connection helps refine our understanding of Belief Revision when context-sensitive information is involved. In the present work, we aim to evaluate whether probabilistic Belief Revision, particularly via conditionalization, offers a useful quantitative counterpart to the centered possible worlds framework for Belief Revision in the context of dynamic, context-sensitive environments. This investigation will shed light on whether both approaches can be integrated or if they need to remain distinct, based on the nature of the information being updated and the role of context in shaping that information.

5.1 Preliminary Remarks

There is no single formal system that is *Probability*. Standard definitions usually take into consideration *Kolmogorov's Axiomatization* to define *probability theory* [38]. However, probabilistic notions are pervasive across various fields, including sciences, decision-making processes, and other areas where uncertainty must be managed. In contexts where agents must revise their beliefs or make decisions, probability plays a key role. In particular, when it comes to changing one's mind, the concept of *updating* via probability provides a coherent framework for understanding how belief systems adapt in light of new information.

Let $\Omega \neq \emptyset$ be the universal set. An *algebra* is a set $\mathcal{F}$ of subsets of Ω with Ω as a member, closed under complementation and union [38]. Let A and B be sentences, and let P be a function from $\mathcal{F}$ to R, obeying the following conditions:

I. **Non-negativity**: $P(A) \geq 0$, for all $A \in \mathcal{F}$

II. **Normalization**: $P(\Omega) = 1$

III. **Finite Additivity**: $P(A \cup B) = P(A) + P(B)$ for all $A, B \in \mathcal{F}$ such that $A \cap B = \emptyset$

We will call P a *probability function*, and $\langle \Omega, \mathcal{F}, P \rangle$ a *probability space*. The concept of conditional probability will be presented in more detail in the next section. In simple terms, the bearers of probabilities are often referred to as "events" or "outcomes". The conditional probability of event A given event B can be expressed as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

In this framework, probabilities are often interpreted as *degrees of belief* under the umbrella of *Subjective Bayesianism*. According to this interpretation, probabilities represent the degree of confidence (or *credence*) that a rational agent assigns to a particular belief or proposition at a given time. More formally, this relationship links a number (the probability) with an agent at a specific time, alongside the object of belief [38]. In contrast to the epistemic entrenchment involved in Belief Revision, the notion of *expected utilities* is different. Expected utilities are not the degree of belief an agent assigns to a proposition; rather, they represent the expected value of an outcome from the agent's perspective. This idea ties closely to decision theory, where utilities of outcomes, probabilities, and rational preferences are interconnected and directly influence the decision-making process. Ramsey, for example, in his analysis, suggests:

> It is possible to derive both probabilities and utilities from rational preferences alone. First, he defines a proposition as ethically neutral - relative to an agent - if the agent is indifferent between having that outcome when the proposition is true and when it is false. The idea is that the agent does not care about the ethically neutral proposition as such - it is a means to an end that he might care about, but it has no intrinsic value. There is a simple test for determining whether, for a given agent, an ethically neutral proposition N has probability 1/2. Suppose that the agent prefers A to B. Then N has probability 1/2 iff the agent is indifferent between the gambles: A if N, B if not; and B if N, A if not. Ramsey assumes that it does not matter what the candidates for A and B are. We may assign arbitrarily to A and B any two real numbers $u(A)$ and $u(B)$ such that $u(A) > u(B)$", thought of as the desirabilities of A and B respectively. Having done this for the one arbitrary chosen pair A and B, the utilities of all other propositions are determined. [38]

This interpretation aligns with the *Classical Interpretation of Probability*, where the probability of an event (or outcome) is the ratio of favorable cases to the total number of equally possible cases, given the absence of any logical reason to favor one over another [53]. In this classical interpretation, the probability function is treated as a function mapping events to real numbers within the interval [0,1].

The *Principle of Indifference* comes into play when we assume that, in the absence of any logical reason to prefer one outcome over another, the probabilities of outcomes are equally distributed. This principle is fundamental in the context of decision theory and probability theory, where it is applied to ensure consistency in probabilistic reasoning when we have no prior information.

5.1.1 Axioms and Theorems in Probability Theory

In probability theory, a set of axioms is essential for defining the properties of probability functions and ensuring consistency across various probabilistic models. These axioms form the foundation of the formal system and allow us to derive useful theorems about the relationships between events and their associated probabilities. The following axioms are commonly accepted in probability theory and provide the structure for defining events, calculating probabilities, and understanding how probabilities relate to each other in different situations.

Considering a field (or algebra) $\mathcal{F}$ of subsets of the sample space $\mathcal{S}$ and sentences A and B, we have the following axioms of probability theory [53]. These axioms form the foundational structure that governs the behavior of probabilities in formal systems.

Axiom 1. (Empty Set) $\emptyset \in \mathcal{F}$

This axiom asserts that the empty set is an element of the field $\mathcal{F}$. In probabilistic terms, it indicates that the event with no outcomes (the empty set) has a probability of 0. The empty set represents the situation where no event occurs, and thus, its probability is 0.

Axiom 2. (Complementation) $A \in \mathcal{F} \rightarrow \overline{A} \in \mathcal{F}$

The second axiom states that if a set A is in the field $\mathcal{F}$, then its complement, $\overline{A}$, is also in $\mathcal{F}$. This means that the probability of an event can always be paired with the probability of its complement. For example, if A represents the event "it rains tomorrow," then $\overline{A}$ represents "it does not rain tomorrow," and both are valid events in the field.

Axiom 3. (Union) $A \in \mathcal{F} \wedge B \in \mathcal{F} \rightarrow A \cup B \in \mathcal{F}$

This axiom asserts that if A and B are events in $\mathcal{F}$, then their union, $A \cup B$, is also in $\mathcal{F}$. The union of two events represents the situation where either event A or event B (or both) occurs. This axiom allows for the combination of different events within the same probability space.

Axiom 4. (Total Probability) $P(S) = 1$

The fourth axiom states that the probability of the entire sample space S is equal to 1. This is because the sample space represents all possible outcomes, and since one of these outcomes must occur, the probability of the sample space is 1. This axiom is fundamental in ensuring that probabilities are normalized.

Axiom 5. (Nonnegativity) $A \in \mathcal{F} \rightarrow P(A) \geq 0$

This axiom ensures that the probability of any event A is always greater than or equal to zero. Probabilities cannot be negative, as they represent the likelihood of an event occurring, which cannot be less than zero.

Axiom 6. (Additivity) $A, B \in \mathcal{F} \wedge A \cap B = \emptyset \to P(A \cup B) = P(A) + P(B)$

The additivity axiom states that if A and B are mutually exclusive events (i.e., $A \cap B = \emptyset$), then the probability of their union is equal to the sum of their individual probabilities. This is a crucial property in probability theory, as it allows for the calculation of the probability of a union of disjoint events by simply adding their individual probabilities.

Several important theorems are derived from these axioms and are useful in various applications of probability theory. These theorems provide deeper insights into the relationships between events and their probabilities.

Theorem 1 $A \in \mathcal{F} \wedge B \in \mathcal{F} \to A \cap B \in \mathcal{F}$

This theorem asserts that the intersection of two events A and B is also an element of $\mathcal{F}$, meaning that the probability of both A and B occurring simultaneously is also a valid event. This follows from the axioms, particularly the complement and union axioms.

Theorem 2 $P(\overline{A}) = 1 - P(A)$

This theorem expresses the relationship between the probability of an event and the probability of its complement. If $P(A)$ is the probability of event A, then $P(\overline{A})$ is simply the complement of that probability, reflecting the fact that either A occurs or it does not occur, and the total probability must add up to 1.

Theorem 3 $P(A) = P(A \cap B) + P(A \cap \overline{B})$

This theorem decomposes the probability of event A into two mutually exclusive parts: the probability that A occurs and B also occurs, and the probability that A occurs but B does not. This is an application of the additivity axiom and is useful in understanding the relationship between events.

Theorem 4 (General Addition Theorem) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

The general addition theorem corrects for the overcounting of the probability of the intersection of A and B. If A and B are not mutually exclusive, the probability of their union is the sum of the probabilities of A and B, minus the probability of their intersection, to avoid double-counting the overlap.

Theorem 5 (Theorem of Total Probability) $(\forall_i)(\forall_j)(i \neq j \to B_i \cap B_j = \emptyset \wedge \bigcup_{i=1}^{n} B_i = S) \to P(A) = \sum_{i=1}^{n} P(A \cap B_i)$

The theorem of total probability expresses how the probability of an event A can be computed by considering all possible partitions of the sample space S. The event A is decomposed into mutually exclusive events B_i, and the total probability of A is the sum of the probabilities of A occurring with each of these partitions.

These axioms and theorems form the basis of probability theory and are essential for understanding how probability functions operate. By defining the relationships between events, their complements, and unions, the axioms provide the structure needed to analyze various probabilistic scenarios. These foundational rules ensure that probability functions are consistent and can be applied across different contexts and domains. Furthermore, the axioms allow us to derive important results, such as theorems on additivity, the total probability of events, and the relationship between an event and its complement, all of which are crucial for reasoning about uncertainty. In the following section, we will explore how these how conditional probability is essential for understanding how the probability of an event changes when we know that another event has occurred. This concept serves as a bridge between the foundational axioms of probability and the more complex processes of belief updating, where the probability of certain outcomes depends on the context and prior beliefs of the agent.

5.1.2 Conditional Probability

In Belief Revision, updating an agent's beliefs in light of new evidence is a key process. This process of updating is deeply connected to the concept of conditional probability, which provides a formal framework for how new information influences the agent's beliefs. The core idea behind updating is that an agent revises their beliefs when presented with new evidence, adjusting their degree of belief in various propositions accordingly. This is precisely the role that conditional probability plays in probabilistic frameworks: it models how the probability of a given event or hypothesis changes when new evidence is introduced.

Conditional probability, denoted by $P(A|B)$, is the probability of event A occurring given that event B has already occurred. This concept is central to belief updating, as it allows an agent to adjust their beliefs based on the relationship between the prior beliefs (represented by $P(A)$) and the new evidence (represented by B). The formula for conditional probability, $P(A|B) = \frac{P(A \cap B)}{P(B)}$, shows that the updated probability of A depends on both the likelihood of A and the occurrence of B. In Belief Revision, this aligns with the idea that an agent updates their belief in A based on the new information B. For instance, suppose an agent holds a belief that a certain event, say A, has a certain probability. When the agent encounters new information B, they revise their belief about A based on the conditional probability $P(A|B)$. This updated belief takes into account not just the original belief in

A but also how A and B are related. Thus, conditional probability provides the mathematical tool for formalizing this process of Belief Revision, ensuring that the updated belief is consistent with the new evidence.

Moreover, conditional probability allows for a coherent and systematic way of adjusting beliefs over time, especially in scenarios where an agent may receive multiple pieces of evidence. As more information becomes available, the agent continues to revise their beliefs by applying conditionalization, ensuring that the revised beliefs reflect the accumulated evidence. This iterative process of updating is essential for modeling how agents handle uncertainty and revise their beliefs in dynamic environments, making conditional probability an integral part of the Belief Revision framework.

Updating, for instance, generally relies on *Conditional Probability*, a concept that reflects how the occurrence of one event affects the probability of another. Formally, the conditional probability function $P(B|A)$ is defined as follows: if $P(A) > 0$, then

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

However, since this definition is not logically adequate in all cases, many scholars argue that $P(B|A)$ should be treated as a primitive expression in probability theory, rather than being defined in terms of other probabilities [53]. To capture this relationship, the following axiom is introduced:

Axiom 7. $P(A \cap B) = P(A)P(B|A)$.

From conditional probability, we can also define the symmetrical notion of *Independence*:

Definition 5.1.1 $Ind_p(A,B)$ *iff* $P(A \cap B) = P(A) \cdot P(B)$.

Several elementary theorems about conditional probability are as follows:

Theorem 6 *If A is independent of B, then B is independent of A,* $\overline{A}$ *is independent of* $\overline{B}$*, and B is independent of* $\overline{A}$.

Theorem 7 *If* $P(B) > 0$*, then A is independent if and only if* $P(A|B) = P(A)$.

Theorem 8 *If* $P(A) > 0$*, then* $P(B|A)$ *is a probability function defined on the field of sets in* $S - \overline{A}$*, where S is the universal set of the original field of sets.*

One of the most important results in conditional probability is *Bayes' Theorem*, which provides a way to update beliefs based on new evidence. It is defined as:

Theorem 9 (Bayes' Theorem)

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(E)} = \frac{P(H) \cdot P(E|H)}{P(H) \cdot P(E|H) + P(\overline{H}) \cdot P(E|\overline{H})}.$$

In this context, the letters H and E typically represent a hypothesis and evidence, respectively. However, this notation is not always accurate in every case. If $P(E|H) = 1$, the theorem simplifies to:

$$P(H|E) = \frac{P(H)}{P(H) + (1 - P(H)) \cdot P(E|\overline{H})}.$$

This formulation allows us to update our beliefs about a hypothesis H given some new evidence E. Bayes' Theorem captures the interplay between prior belief (represented by $P(H)$), new evidence ($P(E|H)$), and the likelihood of that evidence ($P(E)$), enabling more accurate and updated Belief Revision.

Now, we will define the concept of *Expectation*, which plays a key role in probabilistic reasoning, especially when dealing with uncertain outcomes. To formalize this, we need to introduce some additional concepts. A *random variable* is a function that maps from the sample space $\mathcal{S}$ to R, yielding a partition of the sample space [53]. A *frequency function* f_N assigns the probability that a member $y \in \mathcal{S}$ has a value x. A *distribution function* $F_N(x)$ assigns the probability that a member of the sample space has a value less than or equal to x [53].

Definition 5.1.2 (Expectation) *If X is a discrete random quantity, then the expectation of X is defined as*

$$E(X) = \sum_i x_i \cdot f_x(x_i),$$

where the summation extends over all values of x_i such that $P(X = x_i) > 0$.

It is important to note that a discrete random quantity can only take on certain values, which simplifies the cases we need to consider here. We will limit the discussion to non-continuous random variables, as they are more relevant for our current purposes.

The variance of a random variable X measures the extent to which the values of X deviate from the expected value, and is defined as:

Definition 5.1.3 (Variance)

$$D^2(X) = E[(X - E(X))^2].$$

Finally, we can define the concept of *Independence* between random variables. Two random variables X and Y are independent if and only if for every real x and y, the joint distribution function $F_{x,y}(x, y)$ factors as:

Definition 5.1.4 (Independence) *X and Y are independent if and only if for every real x and y,*

$$F_{x,y}(x,y) = F_x(x) \cdot F_y(y).$$

In summary, conditional probability provides a formal mechanism for updating beliefs when new information becomes available. By applying Bayes' Theorem, we can revise our probability assessments of hypotheses based on the observed evidence. Additionally, expectations and variances help quantify the uncertainty surrounding probabilistic outcomes, further enhancing the Belief Revision process. These concepts are foundational for probabilistic models of Belief Revision and provide a bridge between formal probability theory and the practical task of reasoning about uncertain or context-sensitive information.

5.1.3 Logical Interpretation of Probability

The *Logical Interpretation of Probability* takes probability as a degree of validity of a given argument [38]. In that sense, logical probability legislates over rational belief [53]. Moreover, Carnap conceives probability as a framework for rational belief, where the degree of belief assigned to a hypothesis H is determined purely on logical grounds, given the total available evidence. Within this perspective, probability serves as a normative guide for rational belief formation. He assumes a logical language capable of expressing both deductive and inductive relations. In deductive logic, when a statement H entails another statement H', the truth of H necessitates the truth of H'. Analogously, in inductive logic, if H serves as evidence for a hypothesis H', Carnap characterizes their relationship as one of partial entailment. This notion captures how the evidence supports a hypothesis without necessitating its truth, reflecting a logical or necessary connection between them. [53, p.81]

Carnap's [19] conception of logical probability considers a degree of confirmation of a hypothesis H relative to an evidence E as conditional probability, as we saw above. However, there is a requirement that E is taken to be *total evidence*. With that approach, Carnap [19] suggests we define a *measure function m* over the sentences of a First Order Language used to formalize the argument. Moreover, with that, we can describe everything about an individual or world and call it *state descriptions* [53]. We assume that each expression has a unique referential in the possible worlds, and the domain of state descriptions are propositions on the standard sense, that is, sets of possible worlds. The set of state descriptions making the sentence *true* in a disjunction, for instance, is called the *range* of the statement [53]. If all the state descriptions turned out to be negated, the range is equal to an empty set. What happens when we have inconsistent statements (or contradictions). But how do we find a measure m for state descriptions? Carnap's [19] suggestion is that we assign the same measure to every state description in a given language, inspired by the Principle of Indifference. This approach, however, does not give room to conditional probabilities when learning new evidences and, *prima facie*, seems very limited. Carnap's [19] suggests, then, we take a measure function $m^{\dagger}$ to assign equal weights to each structure description, and then divides that weight equally among the state descriptions in that structure description. [...] The weight of a structure description is divided among the state descriptions that belong to it to obtain the measure assigned to each state description [53].

There is no unique way of distributing weights among state descriptions, but the sum of all weights should be 1 considering the interval [0,1] of probability functions presented above. That makes every state description in a given language possible at first unless we have tautological or contradictory statements. In that framework, we do not need to restrict the Language used to describe state descriptions, but it has a limitation: it does not provide a proper way to deal with uncertainty. Recall that, according to the logical interpretation of probability the aim is to legislate over rational belief. We would take degrees of belief as credences and interpret those degrees as levels of disposition to act [53]. With that in mind, the notion of *total evidence* is taken to be all the evidence, explicit or not, logical or not, that leads us to assign those specific degrees of belief to each credence we hold.

Imagine you are willing to bet. Your degree of confidence in credence will directly impact your choices of betting or not and, in case you decide to bet, which alternative you will choose. However, sometimes it is probabilistic impossible to win a bet, and we choose to run the odds anyway. In that sense, we consider the odds of a statement S to turn out to be *true* and take the probability measure we assign to a given statement S to be $1 - P(S)$. In other words, the willingness to bet in favor of a statement is to think that the odds of that statement turning out to be true are greater than the odds of turning out to be false.

When talking about betting, it is standard to say about the *Dutch Book* as a set of odds on which the better will inevitably lose. Sometimes, the odds and our degrees of confidence are not exactly what we think they are. Our preferences may be contaminated with a Dutch Book without us realizing it! The Dutch Books are widely studied in Probability, and we can start with the Dutch Book Theorem as stated below:

Theorem 10 (Dutch Book Theorem) *If $\mathcal{F}$ is a finite field of propositions and P is a probability function for the propositions in that field, then I will be protected against certain loss if and only if P is an additive probability function.*

Even when we are not talking about bets, the logical interpretation of probability shows some apparent limitations. First, there is no extension of the language to include indexical expressions and their particularities. Second, the attribution of measures over the statements on a field is often regarded as too arbitrary. Finally, there is the issue of uncertain inference:

> [...] we can measure the degree of confirmation of one sentence by another sentence or a set of sentences representing our total evidence. We infer the degree of confirmation, if you will, but that inference (given that we have accepted a particular logical measure function) is not at all uncertain. Even if the evidence e gives the hypothesis h a very high degree of confirmation, that does not entitle us to make the inductive leap of accepting H. [53, p.94]

Those reasons lead some authors to prefer a more subjective interpretation of probability when dealing with uncertain inferences and Belief Revision. In particular, the *Subjective Bayesian* approach offers a powerful framework for understanding how agents form and revise beliefs under uncertainty. Unlike classical interpretations of probability, which view probabilities as frequencies or objective measures of uncertainty, the subjective interpretation treats probabilities as a reflection of an agent's personal degree of belief about a particular event or hypothesis. This degree of belief is influenced by the agent's available evidence, prior experiences, and internal reasoning, making the probability function inherently context-dependent and dynamic.

In the context of Belief Revision, the subjective interpretation of probability allows for a more flexible and personalized account of how agents update their beliefs in light of new information. When an agent receives new evidence, they do not simply adjust their belief according to fixed, objective rules. Instead, they update their subjective beliefs based on how the new evidence aligns with their existing knowledge and how confident they are in that knowledge. This process of updating is often formalized through *conditionalization*, where an agent revises their probability assignments in response to new information, adjusting their beliefs in a way that is coherent with the new evidence.

Moreover, the subjective interpretation of probability accommodates the complexities of everyday decision-making, where agents often face uncertain, incomplete, or conflicting information. In situations where the available evidence is vague or contradictory, subjective probabilities provide a way for agents to reason about their uncertainty and make decisions based on their own judgment and risk tolerance. This personalization of Belief Revision processes makes the subjective Bayesian approach especially useful for modeling decision-making under uncertainty, as it emphasizes the importance of individual perspectives and the iterative nature of belief updates. As such, it offers a valuable tool for understanding how agents navigate the complexities of uncertain reasoning and make informed choices in dynamic environments.

5.2 Conditionalization and Belief Revision

Probability approaches of Belief Revision often resort to Bayes's Conditionalization when revising degrees of belief. That approach is known as *Bayesianism*, and it is taken to be a probabilistic or quantitative counterpart for Belief Revision. However, one of the main issues that view encounters is the *Non-Uniqueness Problem*. We know that given a probabilistic function P and a proposition A and B, we assign a value between 0 and 1 to each proposition. We will say that the set of all propositions assigned the value 1 is the *top set of* P. Because of the possibility of existing a top set of P, equivalent probability functions can give rise to distinct and non-comparable revision operators [62]. Two agents that are certain of a given proposition, for instance, might not agree on the certainty degrees after being confronted with a new piece of information. That usually happens because they might not share the exact ordering of probability measures assignments on set.

To avoid those cases, Lindstom and Rabinowicz [62] suggest that we interpret belief sets as the common top set of P of the *revised probability functions*. Instead of choosing between different probability functions, we take what is common to all of them to determine our resulting belief set. Moreover, when facing the Non-Uniqueness Problem, we might need to choose a particular function to behave as a dictator (more about it on the Arrow's chapter of this Thesis). They [62] also suggest we create a *base function* $b : K \to P$ from the belief set K to the set of probabilistic functions P such that for every sentence $A \in K$, $b(A)$ is one of the probabilistic functions that has A as their top. To revise A with a proposition B, we first take the resulting $b(A)$, revise it by B, and select the top of the resulting revision. Selecting the correct base function, however, is not an easy task. If the assignments are too arbitrary, they might not be able to generate adequate Belief Revision functions [62]. A requirement that we can make to choose a perfectly appropriated base function is to select one that satisfies the commutativity condition; that is, changes on the orders of the elements do not alter the result.

With that in mind, we consider that an agent's policies for Belief Revision may not result in a unique belief set. Suppose, for instance, that an agent has a policy $*$ for revising her degrees of belief. We might have a set R of equally acceptable results of revising her belief set K with a sentence B. R, then, has as members all possible results by only revising K with B. When we add more pieces of information and find ourselves in a *iterated Belief Revision* situation, things can get even more complicated. Lindstom and Rabinowicz [62] then suggest another approach, namely, to identify belief states not as sets but as primitive notions associated with a belief set as its top.

As we can see, probabilistic beliefs appear to be more complicated to deal with than propositional beliefs. Moreover, there are some doubts about how meaningful they are when talking about belief changes since the tools were established for certainty, that is, those beliefs that are assigned with probability measure 1 [48]. We might assign probability 1 to those beliefs we are certain of - although some might argue that we should only assign 1 to *knowledge*.

It is worth briefly saying that there is a debate whether probabilistic revision with statements assigned probability 1 could be considered cases of revision or focusing. At the probabilistic approach, *Focusing* is to resource to a reference class that is able to describe a *context of interest* and usually is done by Conditionalization [48]. Since Conditionalization also does probabilistic revision, we might think they are the same. However, Revision is a different concept that aims to refine the belief set and consistently incorporate new information.

Besides the standard Bayes conditionalization formulation, *Jeffrey's Conditionalization* is thought to allow modifications on the probability distribution in order to deal with changes on the belief sets [48]. Jeffrey's Rule can be defined as follows:

Definition 5.2.1 (Jeffrey's Rule) *Let P be a probability measure and let A be a proposition such that $P(A) \neq 0$. Suppose, now, that $P(A)$ has changed to x in the interval [0,1]. The new probability measure: $P'(A) = x$ is given by Jeffrey's Rule:*

$$P'(B) = x.P(B|A) + (1-x).P(B|\neg A)$$

The main advantage of Jeffrey's conditionalization is that it does not assume certainty as standard Bayesian conditionalization. An interesting application of the rule is applying simultaneous Jeffrey's rules to successive belief updating [72]:

> In these belief changes, a course of experience leads an agent's credence in each member of two or more partitions to be directly changed simultaneously. I will call this kind of belief update simultaneous belief updates. In particular, when and only when a course of experience leads an agent's credence in each member of two partitions E and F to be directly changed simultaneously and nothing else is directly changed. I will call this the simultaneous belief update relative to E and F. [72, p.2]

5.3 Simultaneous Belief Updates and Self-Locating Uncertainty

The concept of *simultaneous belief updates* introduces an important layer to the study of Belief Revision, particularly in cases where agents must adjust their credences regarding multiple aspects of their epistemic state at once. Park [72, p.2] characterizes simultaneous belief updates as cases in which an agent's credence in multiple partitions changes concurrently as a result of a single course of experience. This notion is especially relevant in contexts where agents navigate *self-locating beliefs*, requiring them to incorporate both new external information and adjustments to their own perspective.

Park's framework aligns with the broader discussion of *context-sensitive Belief Revision*, particularly in situations where the *agent's self-location* plays a critical role in determining how information should be updated. In standard Bayesian Belief Revision, updates occur via conditionalization: an agent receives evidence E and updates their prior probability distribution accordingly. However, in cases involving *self-locating uncertainty*, such as in *indexical beliefs* or *self-locating thought experiments*, this process is less straightforward. The agent must track not only the *factual content of a proposition* but also the *context in which it is evaluated*.

A relevant example of simultaneous belief updates in self-locating scenarios arises in the *Sleeping Beauty problem*. In this well-known paradox, Sleeping Beauty is put to sleep on Sunday and wakes up on Monday and possibly on Tuesday, depending on a fair coin toss. Upon waking, she must revise her credences about what day it is and whether the coin landed heads or tails. This case exemplifies how agents undergoing belief updates concerning their *self-location* (e.g., "What day is it?") must simultaneously update their beliefs about external facts (e.g., "What was the coin's outcome?"). The challenge here is that the agent's *self-locating uncertainty* directly interacts with their probability assignments, leading to competing interpretations of the correct Belief Revision rule (e.g., the *thirder position vs. the halfer position* in the debate).

The structure of *simultaneous belief updates* also has implications for how agents process information in *multicentered worlds* (Torre [108], Kindermann [49]). In a *multicentered world*, multiple agents share a conversational space, but each agent has their own epistemic position (or *center*) within that space. When new information is introduced, different participants in the conversation may undergo belief updates that are *not necessarily uniform*. For example, in a scenario where one agent gains access to privileged information while another does not, their *epistemic divergence* must be modeled probabilistically. The notion of *multicentering*, introduced by Kindermann [49], suggests that conversational participants update their beliefs *not in isolation*, but in a way that reflects the *joint self-location of the group*.

The probabilistic dimension of *self-locating beliefs* becomes even more significant when dealing with *incomplete or noisy information*. Consider a situation in which two individuals are placed in separate rooms and given different probabilities about an event occurring. If both individuals undergo a *simultaneous belief update* upon receiving a new clue about their situation, their individual belief changes may be interdependent. In such cases, Bayesian updating alone does not fully capture the structure of the epistemic shifts occurring, and a more sophisticated framework—one that allows for probabilistic interactions between different *centers* in a given world—is needed.

Thus, *simultaneous belief updates* provide an essential connection between *probabilistic epistemology and context-sensitive reasoning*. They demonstrate how Belief Revision is not always a linear process but can involve multiple dimensions of change, particularly when agents must track both external facts and their self-locating status in an evolving informational environment. This interplay between *Belief Revision, self-location, and probabilistic inference* will be further explored in the next section, where we examine the quantitative formalization of *context-sensitive updates* and their implications for broader epistemic models.

5.4 Self-Locating Credences

There are some situations in which Bayesian's and Jeffrey's conditionalization might not be able to deal. Self-locating credences, for instance, pose a problem to the probabilistic approach. We take self-locating beliefs to be interpreted as *centered credences* since centered possible worlds can be used to represent self-locating credences. According to Milano [68], we can define probabilities also for centered events. Suppose, for instance, that an agent holds credence at an instant t_0 and at a later instant t_1 she learns a shred of new evidence. Then, for every centered event A the credence $P(A)$ at t_1, the credences satisfy standard conditionalization if they change over time considering the following rule [68]:

Definition 5.4.1 (Conditionalization Over Time) $P_{t_1}(A) = P_{t_0}(A|E)$, *where* E *is the total evidence learned between* t_0 *and* t_1.

That definition corresponds to standard conditionalization, making clear what was learned at a later time. However, beliefs about one's own temporal location are self-locating beliefs, and learning new evidence might lead to learning something about their own location (and vice-versa). About the relation between learning self-locating and non-self-locating information, Titelbaum [106] presented the following thesis:

Relevance-Limiting Thesis It is never rational for an agent who learns only self-locating information to respond by altering a non-self-locating degree of belief.

This thesis turns out to be false, as Titelbaum [106] states, as learning self-locating information may alter non-self-locating beliefs and/or credences. He then suggests we should develop a new framework to deal with rational changes in the degrees of belief over time. For that, we will present some definitions. Think about the changes over a period of time as a *story*:

Definition 5.4.2 (Story) *A story describes an agent who starts with a particular set of certainties and then becomes certain of other things at various times.*

Recall that the probabilistic approach towards degrees-of-belief aims to assign a valued function, the *credence function*, to represent the agent's degree of belief in a given statement. In a story, those degrees change over a period of time. For that reason, Titelbaum [106] suggests we separate credence functions for each instant t, as Milano's definition above. Because we have different instants of time in a story, we will need to define a *Time Set*:

Definition 5.4.3 (Time Set $\mathcal{T}$) *The time set of a model is a non-empty, finite set of instants* $\{t_1, t_2, ..., t_n\}$ *during the story. The subscripts reflect their temporal order, at which we will model the changes of the agent's degrees of belief.*

An interesting linguistic remark is that Titelbaum calls *claims* those sentences that take a truth-value according to a *context*, paving the way to context-sensitivity within that framework. With that in mind, we have the following definitions:

Definition 5.4.4 (Unconditional Credence Function) *An unconditional credence function is a function from sentences in the modeling language to the reals. A value of an unconditional credence function represents the agent's degree of belief at a given time in a particular claim. A higher credence value represents greater confidence in a particular claim, and unconditional credence of 1 represents the certainty of that claim.*

and

Definition 5.4.5 (Conditional Credence Function) *A conditional credence function is a partial function from ordered pairs of sentences in the modeling language to the reals. The conditional credence* $P_1(A|B)$ *represents the agent's degree of belief at time* t_1 *in the claim represented by A conditional on the supposition of the claim represented by B.*

Although a story describes an agent's changes on the credences over time, a *History* is like a screen of how those credences changed over time:

Definition 5.4.6 (History) *A history is a set of credence functions containing exactly one conditional credence function and one unconditional credence function indexed to each time in the time set.*

Some constraints, however, must be imposed on the credences functions in order to guarantee consistency - and, therefore, that the credences are rational. *Systematic constraints* and *Extrasystematic constraints* legislate over the certainty of the claims in a story. Moreover, for each A at a time t_i, there will either be an extrasystematic constraint that $P_i(A) = 1$ or an extrasystematic constraint that $P_i(A) < 1$ [106]. That intuitive constraint takes certainty to be assigned probability 1, and that will be the case of tautological statements, for instance. In Titelbaum's [106] framework, a *model* is a set of all possible Histories that meet both kinds of constraints that evaluates the evolution of the epistemic states of an agent. Moreover, the positive results of this evaluation represent the necessary conditions for rationality [106]. Those results are referred to as *veredicts* and, sometimes, limit the permissible degree of belief of a claim to one. Titelbaum's approach is aligned with AGM Postulates for rational revision, but the verdicts are not enough to determine precise probability measures to all claims.

Another kind of constraint, the so-called *Synchronic Constraints* for Titelbaum's proposal, take into consideration the Kolmogorov Axioms of the standard interpretation of Probability as follows:

Definition 5.4.7 (Kolmogorov Axioms)

Let $\mathcal{M}$ be a model, $\mathcal{L}$ the set of all sentences in a given language, and $\mathcal{T}$ represents a time set:

(1) For any $t_i \in \mathcal{T}$ and any sentence $A \in \mathcal{L}$, $P_i(A) > 0$.

(2) For any $t_i \in \mathcal{T}$ and any tautological sentence $\top \in \mathcal{L}$, $P_i(\top) = 1$.

(3)For any $t_i \in \mathcal{T}$ and any mutually exclusive sentences $A, B \in \mathcal{L}$, $P_i(A \vee B) = P_i(A) + P_i(B)$.

Note that those axioms are analogous to those presented on the standard interpretation of probability, but have the instants in a time set to specify changes over time. They also take the credence's values to be a number in the interval [0,1] on which assigning value 1 to a claim A represents the certainty of A while assigning the value 0 to A represents the certainty of $\neg A$. Titelbaum [106] also presents an additional synchronic constraint as follows:

Definition 5.4.8 (Additional Synchronic Constraint)

(4) For any $A, B \in \mathcal{L}$ and any $t_i \in \mathcal{T}$, if $P_i(A) > 0$, then $P_i(A|B) = \frac{P_i(A \wedge B)}{P_i(B)}$. If $P_i(B) = 0$, then $P_i(A|B)$ is undefined.

That additional constraint offers a rationality criterion: When an agent assigns a credence value 0 to a claim, assigning a degree of belief greater than 0 to any of the conditional probabilities involving that same claim would be irrational. Indeed, that is also valid when we have *updating by conditionalization*, that is, when an agent assigns a credence to a claim A at an instant $t_i \in \mathcal{T}$, then her conditional credence on A in a later time t_j takes into consideration everything she learned between t_i and t_j. With that in mind, we can define a *Certainty Set* as follows [106]:

Definition 5.4.9 (Certainty Set)

Given a model $\mathcal{M}$ and an instant $t_i \in \mathcal{T}$, we define the agents's certainty set Cer at t_i as $Cer = \{A \in \mathcal{L} : P_i(A) = 1\}$. Given two times t_i and t_j in $\mathcal{T}$ with $t_i < t_j$, the set $Cer_j - Cer_i$ will represented the gained certainty between t_i and t_j.

Conditionalization seems to be a great tool to update the certainty set. However, it might fail when applied to context-sensitive statements since it was thought to model situations where credence values vary from 1 to 0. Thus, the uncertainty cases in the interval lack adequate treatment in this approach. Those cases are very interesting for pragmatic changes in an agent's epistemic states and decision-making processes. When context-sensitive claims appear in a story, we might lose some degree of confidence in a claim without necessarily negating it. It is possible, as Titelbaum [106] states, for an agent's degree of belief in a claim goes from 1 to < 1 in different contexts. Suppose, for instance, I say Today is Monday and my degree of confidence in that claim is 1 (because I checked the calendar, for instance); but then I sleep, and the next day my degree of confidence in that claim goes to 0 (since I know it is Tuesday). Intuitively, I do not violate any rationality requirement when making that change due to the passage of time. To make things a bit more complicated, we can also have situations in which I am not sure about what day is, and my degrees of confidence vary between the interval [0,1] as I learn more about the world and my temporal location in it. Thus, in the case of context-sensitive claims, we should restrict the application of Conditionalization to those cases on which we have certainty (that is, the extremes).

Analogous to AGM Postulates, Titelbaum [106] presents the notions of *Expansion* and *Reduction* of a model. We should, however, restrict our language set. To find a proper reduction, we can create a language $\mathcal{L}^-$ by removing all the sentences representing context-sensitive claims from $\mathcal{L}$. Thus, $\mathcal{L}^-$ does not include any context-sensitive sentences, only *eternal claims*. Although that solution shows an important restriction to the application of updating with conditionalization, it does not provide us with a way to deal with indexical information in this probabilistic framework.

In some cases, we can substitute an indexical expression with a co-referential eternal expression, as debated before. In that case, we will say that if $P_i(A \equiv B) = 1$, then B can be replaced with A. For this substitution to be possible, they have to share the same truth-value, but what is the impact of that substitution on the agent's doxastic state? Furthermore, what about those cases in which the substitution is not possible?

To Titelbaum [106], *every* context-sensitive claim could be substituted by a context-insensitive counterpart. That is a very strong statement that does not seem to consider the *essentiality* of some indexical expressions [73]. However, what Titelbaum seems to have in mind as context-sensitive claims are those containing temporal indexical expressions such as *Now*. For that, if the contextual information is known, it is possible to find context-insensitive equivalents to substitute the context-sensitive claim and avoid most confusion.

We want to represent the agent's degrees of confidence in claims about the world and the self-locating information about where she is located in that world. Self-locating credences can find treatment on Lewis's *centered worlds* approach and their values distributed over centered propositions. According to Titelbaum [107], the probability measures can be distributed over sets of centered possible worlds the same way they would be distributed over standard uncentered propositions. In that sense, *updating by conditionalization* would work the same way for centered and for uncentered claims. He argues that the updating rule would capture the evolution of an agent's degree of confidence over a period of time and yield plausible results no matter what kind of claim we are considering. To capture the distinctions, Titelbaum [107] presents the different schemes that can incorporate context-sensitive claims.

Definition 5.4.10 (Shifting Schemes) *On a shifting scheme, we shift the indexical expression to a co-referential one changing tenses appropriately.*

Thus, to keep the level of certainty of the claim Today is Monday at 1 when it has passed a day, we shift the indexical expression to *Yesterday*. That way, Yesterday was Monday would continue to have the degree of confidence 1 assigned earlier on. To make that more precise, we can include an operator *at* and say that, given a claim A and a context c, A is true if it A *at* c is true.

Consider, for instance, the case in which the agent learns nothing but the day's passage. If that is the case, her credences that A is true at t_i should be reflected by her credences at t_j that A is true at t_i. Thus, the agent should update their belief set by conditionalization as follows:

Definition 5.4.11 (Temporal-Sensitive Conditionalization) *Given a claim A and an instants $t_i, t_j \in \mathcal{T}$,*

$$P_j(A) = P_i(A \ at \ t_j)$$

You might consider that learning that time has passed means that the agent learned the centered proposition Now is t_j . Learning that time has passed is learning something, in the sense of updating the credence on where she is located within in a possible world. Titelbaum [107] suggests the following definition for *Shifted Conditionalization*:

Definition 5.4.12 (Shifted Conditionalization)

If an agent learns a proposition B and nothing else between t_i and t_j, and it is certain at t_j that is t_j, then for any proposition A

$$P_j(A) = P_i(A \ at \ t_j | B \ at \ t_j)$$

However, shifted conditionalization does not contemplate the cases in which the agents might lose track of the contextual parameters. Indeed, to shift correctly, the agent needs to be certain about the contextual features, and her own location within a world, and that is captured by the requirement that the agent is certain at t_j that is t_j. If the agent is uncertain about his self-locating information, the shifting scheme will not be able to deal with a new conditional probability measure.

Another type of scheme presented by Titelbaum [107] is the *Stable Base Schemes*. Stable Base Schemes consider the agent's credences in uncentered propositions rely on the fact that centered propositions pose a problem for standard conditionalization.

Definition 5.4.13 (Stable Base Schemes) *Their core idea of Stable Base Schemes is to focus on the set of uncentered beliefs an agent holds and apply updating by conditionalization only to the uncentered members of that stable base. Credences in centered propositions are then set by coordinating them with the uncentered distribution at a given time.*

To make stable base schemes possible, it is necessary to apply a *Compartmentalized Conditionalization* [107] and follow the following steps:

1. Temporarily consider your $P_[i$ distribution only over all possible worlds, centered and uncentered;

2. Assign any worlds incompatible with your t_i evidence credence of 0, then redistribute the remaining (non-zero) credences over worlds;

3. Assign centered worlds incompatible with your t_i evidence credence of 0. Then take the credence assigned to each world and distribute it among the centered worlds indexed to that world compatible with your t_i evidence.

The compartmentalized conditionalization result is a new credence function over only the uncentered claims by conditionalizing the strongest uncentered proposition, ruling out possible worlds, and leaving that credence aside while conditionalizing only over the centered claims. Once you generate a distribution function over the centered propositions, you apply the axioms to generate unique credence for all propositions (centered and uncentered). Titelbaum [107] states that

> Because relevant centered evidence may go unreflected in uncentered propositions, we cannot follow Compartmentalized Conditionalization in letting only uncentered propositions learned to affect one's uncentered distribution. [107, p.674]

A possible solution, once more, is to restrict the language to only uncentered claims, and we get a less complicated approach. Nevertheless, since our purpose is to incorporate context-sensitivity claims into our language, we still search for new approaches to deal with those cases. The updating scheme *Certainty-Loss Framework* proposed by Titelbaum [107] aims to deal with possible flaws in the application of compartmentalized conditionalization and allows an agent to update by compartmentalized conditionalization every time a centered claim can be substituted by an equivalent uncentered counterpart with the same truth-value. Stable base schemes might be a good option when talking about uncentered credences. However, when an agent lacks contextual information to find an equivalent uncentered proposition to a centered claim, those schemes do not work. But do not worry: we might have another kind of scheme, namely *Demonstrative Schemes*, that might do the trick.

Demonstrative Schemes, as proposed by Titelbaum [107], assume Stalnaker's view that centered information always has an uncentered equivalent counterpart, and that substitution between the two is always possible. This approach relies on the traditional understanding of possible worlds, where demonstrative information is incorporated into the construction of centered worlds. In this framework, the shift from uncentered to centered propositions can be accomplished through a straightforward substitution of terms, which implies that centered information can be reduced to uncentered forms. According to this perspective, one would only need to apply standard conditionalization to update an agent's beliefs when new information is introduced, as the Belief Revision process is fundamentally rooted in the manipulation of these uncentered counterparts.

However, as discussed in previous sections, we know that this assumption does not hold universally. The simplification that all centered information can be substituted with uncentered counterparts overlooks important complexities in Belief Revision processes, particularly when it comes to context-sensitive and self-locating beliefs. For instance, in scenarios where agents are dealing with indexical information or self-locating propositions, the Belief Revision process cannot be adequately captured by traditional conditionalization alone. This is due to the inherent difference between how agents process uncentered information—where conditionalization can indeed be applied—and how they handle centered propositions, which involve the agent's specific location within the possible world. The addition of context-sensitive elements into the Belief Revision framework introduces nuances that standard conditionalization fails to account for.

Moreover, the notion of substitution between centered and uncentered information, while useful in some cases, fails to account for the dynamic and interactive nature of belief updates in everyday contexts. In situations where agents are uncertain about their position in a given context—such as in the case of indexicals or self-locating beliefs—this substitution does not capture the complexity of how beliefs are formed or revised. In these cases, the shift in beliefs is not simply a matter of updating an uncentered proposition with a new set of uncentered beliefs, but rather involves a process of recentering, where the agent must update both their knowledge and their understanding of their position within the world. Thus, while demonstrative schemes and traditional conditionalization can work in some contexts, they are insufficient for handling the more intricate aspects of Belief Revision that arise when dealing with self-locating and context-sensitive information. This is why we argue that a more nuanced approach is needed—one that incorporates the complexities of centered worlds and the unique challenges posed by context-sensitive beliefs. Only by extending the framework to account for the interaction between centered and uncentered information can we develop a more comprehensive model of Belief Revision that truly reflects the way agents process and update their beliefs in the face of new, often context-dependent, information.

5.5 Centered Worlds and Probabilistic Approaches

As we can see by the issues posed by the Sleeping Beauty Problem, dealing with self-locating credences within a probabilistic approach is a challenge. Milano [68] suggests we take *centered probability* to be the different applications of probability functions to objects depending on the context. If the context is said to be a set, different contexts will provide different ranges for the probability functions. Let us look at two proposals to deal with the dynamics of indexical information from a probabilistic point of view.

5.5.1 Milano's Centered Probability

Standard interpretations of probability usually range over the set of context-insensitive formulas. However, the set of all possible (or plausible) outcomes might be different if we change the context of the application. Those contexts can be interpreted as the set of ordinary possible worlds on with the probability function will range over, and it will be the set of all possible worlds W. Thus, an event $E \subseteq W$ will be a subset of possible worlds in which that specific claim turns out to be true. Centered worlds are an extension of ordinary possible worlds as they provide additional information about a specific location within that possible world. Thus, Milano [68] defines a *centered event* as follows:

Definition 5.5.1 (Centered Event) *A centered event E_c is defined as a set of centered worlds W_{ec} in the set of all centered worlds W_c. Since W_{ec} forms a field, a probability function P_c can be defined on W_{ec}.*

According to Milano [68], the main advantage of centered events is to provide a more fine-grained probability measure to self-locating uncertainty since ordinary possible worlds might not be able to capture all contextual features of indexical beliefs. The relationship between the two kinds of events is given by the following correspondence:

Definition 5.5.2 (Correspondence between E and E_c) *For any ordinary event $E \subseteq W$ and a set of contexts C, we can define the corresponding centered event $E_c = E \times C$, that is, the set of centered worlds $\langle w, c\rangle$ in the field such that $w \in E$ and $c \in C$.*

The intuitive interpretation of the correspondence principle above is to ensure that for every uncentered event E it is possible to identify a centered event E_c such that all possible ordinary worlds $w \in E$ are also in E_c. Notice that the converse does not hold, since not every centered event will have a corresponding uncentered event. This result implies that every treatment provided for centered events will be a generalization able to deal with ordinary events. That is a good argument on behalf of developing a formal treatment for centered probabilities. Moreover, Milano [68] distinguishes centered events from *indexical events*, defining the latter as follows:

Definition 5.5.3 (Indexical Event) *E_c is indexical if for all $w, w' \in W$ and for all $c \in C$ such that $\langle w, c\rangle, \langle w', c\rangle$ in the field, $\langle w, c\rangle \in E_c$ iff $\langle w', c\rangle \in E_c$.*

The interesting aspect of that distinction is that it allows us to formally separate two kinds of uncertainty concerning centered events. On the one hand, uncentered events can represent uncertainty with respect to the possible world $w \in W$ while indexical events represent uncertainty with respect to the contextual feature $c \in C$. Being able to distinguish those is important to provide a more concise framework to deal with uncertainty in self-locating beliefs. Being uncertain about the world is not necessarily the same as being uncertain about our own location in the world, as the Messy Shopper Example showed us.

If the standard interpretations of probability can be extended to centered possible worlds, we might have our way out for the Sleeping Beauty Problem. Milano's [68] solution to Sleeping Beauty Problem relies on how we model the problem:

> 1. Bayesian reasoning can be naturally applied to self-locating uncertainty. We do not have to reform probability theory or design new updating schemes to deal with this sort of case.
>
> 2. To avoid puzzling conclusions, it is important to model evidence correctly. In particular, we should be careful to model what is the prior probability of receiving different pieces of evidence.
>
> 3. self-locating uncertainty does not cause a rift between fair betting odds and credences, and the case of Sleeping Beauty does not undermine Dutch Book arguments for probabilism and conditionalization. [68, p.220]

In the case of our Sleeping Beauty Experiment, we take Beauty's uncertainty to be one of the two forms: she might be uncertain about the world she is in, or she might be uncertain about what day is. In the second case, she fails to keep track of the contextual changes (namely, the temporal ones). To deal with that, Schulz [96] suggested another approach.

5.5.2 Schulz's Dynamics of Indexical Belief

The reader should realize that sometimes the evidence we acquire is an indexical belief. When we cannot keep track of the relevant contextual information, we cannot find equivalent uncentered claims to make a substitution and apply standard updating by conditionalization. However, it is intuitive that our beliefs may evolve with time, primarily our beliefs about temporal location within a possible world. Although *Sleeping Beauty Experiment* raises some of the questions, the dynamics of indexical beliefs are more general.

We know that the indexical expressions "Today" and "Yesterday´´ might be coreferential with respect to successive contexts (if we consider instants to be days).

However, that does not hold on the Sleeping Beauty case because Beauty cannot keep track of the contextual features and how two distinct contexts relate to each other. Because of that, Beauty does not retain indexical beliefs, as Schulz [96] states. Even though she knows that Monday and Tuesday are instants later than Sunday, she does not know how distant she is from Sunday on a temporal line of successive instants.

Suppose we learn nothing new about the world between t_i and t_j. In theory, we should keep believing the same things at t_j we did at t_i. But that cannot hold because agents often have self-locating beliefs that might change with the passage of time. Beauty, for instance, might hold beliefs about which day it is will change as the days go by. In order to incorporate the context-sensitivity of indexical beliefs, Schulz [96] suggests we assign credences to sentences within specific contexts. In that sense, $P_c(A)$ will be the probability measure assigned to a sentence *A attached* to a context c. Contexts, for him, will mean pairs of $\langle t, s \langle$ on which t is an instant and s a spatial position. A *distance between contexts* will formalize the successive line of contextual features we need to keep track of and it will be defined as follows:

Definition 5.5.4 (Distance between Contexts)

The distance $c_1|c_2$ between two contexts $c_1 = \langle t_1, s_1 \rangle$ and $c_2 = \langle t_2, s_2 \rangle$ as $\langle t_2 - t_1, s_2 - s_1 \rangle$.

Following Schulz [96] nomenclature, given an indexical expression α and a distance x between contexts c_1 and c_2 (that is, $c_1|c_2 = x$), $\alpha[x]$ will represent a corresponding indexical expression $\alpha*$ such that the truth value of $\alpha*$ is the same as the truth value of α. That correspondence aligns with the idea that we can substitute the indexical expression $Today$ at c_1 for $Yesterday$ at c_2 when we know the distance between both contexts is one day. Formally, we would have the following 'yesterday[1]' = 'tomorrow'. We can extend that definition for sentences containing occurrences of indexical expressions by substituting $\alpha[x]$ for all indexical elements α in the sentence [96]. Thus, distances are operations on expressions or sentences resulting in corresponding expressions or sentences.

Schulz's [96] approach is very interesting to deal with time and space coordinate, although it was not yet extended to other contextual features. Knowing the distance between two contexts will often allow me to identify an appropriate expression to refer to indexical content in later contexts. That new expression (or sentence) allows us to express the same proposition in two distinct contexts, making the proper changes on tenseness.

> When we move from a context c_i to a context c_j and consider an indexical sentence φ in the present context c_j, the sentence $\varphi[c_i|c_j]$ is

> a sentence which expresses the same proposition in the earlier context c_i as the sentence φ we are currently considering expresses in c_j. Now, when we aim at continuity of belief between a past context c_i and a present context c_j and ask ourselves what our present credences in an indexical sentence φ should be, it therefore seems that we should assign to it the subjective probability we assigned to $\varphi[c_i|c_j]$ in c_i. [96, p.340]

That notion is aligned with Kaplan's [46] Logic of Demonstratives and it is not very controversial. Let's consider, then, the *continuity principle*:

Definition 5.5.5 (Continuity Principle)

Let c_1 and c_2 be two successive contexts such that the agent gains no new evidence after c_1. Then,

$$P_{c_2}(A) = P_{c_1}(A[c_1|c_2]).$$

A good thing about the continuity principle is that it does not necessarily characterize new evidence: it can also represent changes in the context. When moving from a past instant to a present one, we consider our present beliefs related to our previous ones. Continuity principle describes this evolution of indexical beliefs when no more information about the world is gained. To Schulz [96], we get to our present credences by conditionalizing our past beliefs. In the case of indexical beliefs, we need to conditionalize our past beliefs with respect to the resulting $\alpha[x]$. With that in mind, we can define *continuos conditionalization* as follows:

Definition 5.5.6 (Continuous Conditionalization)

Let c_1 and c_2 be two successive contexts and E_c the evidence available in c_2. Then, we need to consider the following steps:

1. First, we let $P'_{c_2}(A) := P_{c_1}(A[c_1|c_2])$.

2. Second, let $P_{c_2}(A) := P'_{c_2}(A|E_c)$.

Combined, the two steps result in

$$P_{c_2}(A) = P_C(A[c_1|c_2]|E_c[c_1|c_2]).$$

In the first step, we accommodate the changes of the contextual features, such as in the passage of time; and the second step accommodates the new evidence gained between t_1 and t_2. Continuous Conditionalization, however, does not provide a way to deal with situations in which we cannot keep track of the contextual changes, such as Beauty in the experiment. We cannot fulfill the first step if we lose track of which day it is. But although we might lose track of the changes of contextual features, it is quite rare that we are completely off on our guesses about which day is. Beauty might not know if today is Monday or Tuesday, but she probably knows that it is not Friday (because of how the experiment was presented to her). If she was asleep for a hundred years, then that might be the case - which is not. Thus, our contextual information credences play a role in the probability measures we assign to centered events.

When misunderstandings arise in conversational contexts—such as Alice and Bob's exchange around midnight—they highlight the importance of adapting to newly acquired information as the certainty about contextual parameters increases. In such cases, individuals instinctively revise their beliefs to align with updated contextual information. This process reflects a broader distinction in how we handle contextual shifts: we may either operate under conditions of certainty, where we have a stable understanding of the context, or under uncertainty, where key contextual elements remain ambiguous or unknown.

To formalize this intuition, we can appeal to the framework of indexical belief dynamics. As argued in [96, p.344], when an agent gains new evidence regarding contextual changes, this evidence directly influences their credences in sentences that partition possible contexts. The *Acceptance Principle* (ACC) states that any newly acquired evidence should be assigned credence 1, meaning that an agent must fully accept the contextual update. Additionally, an agent's prior credence function may already encode expectations about contextual changes, further constraining how beliefs should be revised in light of new information.

> [...] if information about the contextual changes is part of the evidence gained by the subject, this will bear directly on the credences in the partition forming sentences. (ACC) requires that newly found evidence is assigned credence 1. Also, there will be cases in which the past credence function contains non-indexical information relevant to the expected contextual changes. Thus, (ACC) imposes some constraints on what our credences in the partition forming sentences should be. [96, p.344]

Formally, we can define ACC as follows:

Definition 5.5.7 (Approximated Continuous Conditionalization (ACC))

Let c_1 and c_2 be two successive contexts and $x_1,...,x_n$ a set of distances with respect to c_1. Let P_{c_2} be a probability function such that the sentences of the form 'the distance between the present context c_2 and the past context c_1 I am updating from is x_i'. Let E_{c_2} be the evidence available in c_2. Then

$$P_{c_2}(A) = \sum_{i=1}^{n}[P_{c_2}(c_1|c_2 = x_i \times P_{c_1}(A[x_i]|E_{c_2}[x_i])].$$

A paradigmatic case of this phenomenon is the Sleeping Beauty problem, where an agent's self-locating belief is influenced by the passage of time and the conditional structure of the experimental setup. On Monday, Beauty is put to sleep, and upon waking, she must determine whether it is Monday or Tuesday. If she possesses knowledge of the experimental protocol, she can estimate the likelihood of each possibility based on probabilistic reasoning. However, at any given waking moment, she lacks direct epistemic access to certain crucial contextual details—such as whether she has been woken once already—leading to an apparent loss of self-locating knowledge.

This raises an important question: when Beauty awakens, is she truly *forgetting* information, or is she merely in a situation where she lacks direct epistemic access to her temporal location? Some might argue that Beauty undergoes a process of memory loss when she "loses" self-locating belief upon waking. However, it is more accurate to say that she experiences a shift in the available contextual information. Her memory remains intact, but her position in time is not accessible to her in a way that allows her to distinguish between Monday and Tuesday. In this sense, her epistemic state is altered by external contextual changes, rather than by the loss of previously held knowledge.

The distinction between true forgetting and contextual ignorance is crucial for understanding how agents update beliefs in indexical scenarios. If an agent undergoes a contextual shift while unconscious or unaware—such as Beauty while she is asleep—the change in her epistemic state does not constitute genuine forgetting but rather a temporary epistemic limitation. In such cases, conditionalizing on newly available evidence allows agents to regain their self-locating knowledge and revise their credences appropriately. This dynamic process mirrors the way we intuitively adapt to contextual changes in everyday scenarios, such as Alice and Bob's midnight conversation: as new information about the context becomes available, agents adjust their understanding of the situation without necessarily discarding previously held knowledge.

Thus, understanding self-locating Belief Revision requires a nuanced approach that distinguishes between true memory loss and shifts in contextual awareness. In the Sleeping Beauty problem, Beauty is not *forgetting* information in the traditional sense; rather, she is awakening into an epistemic situation where certain crucial details—such as whether she has already experienced a waking event—are temporarily inaccessible to her. The ability to estimate contextual changes, either probabilistically or through rational constraints such as (ACC), provides a way to bridge this epistemic gap.

5.6 Self-Location and Distance in Belief Revision

In the Alone in the Wilderness scenario, John is the lost ranger who finds himself separated from his group in the vast wilderness. Initially, John's belief about his location is uncertain, as he cannot be sure of his exact position in relation to the rest of the world or even the immediate surroundings. His beliefs about the world are deeply affected by his self-location, and in particular, by his sense of time and space. As he navigates through the wilderness, John is continually faced with new clues about his environment—changes in his surroundings, the position of the sun, or landmarks he encounters. These clues provide him with new information, which he updates into his belief system. However, unlike Jane the detective, John's Belief Revision is complicated by the fact that he cannot always trust his initial self-location.

In probabilistic terms, John's belief about his location is time-sensitive, just like the detective's investigation, but his self-locating beliefs must also contend with the uncertainty about which “center" he occupies within the wilderness. His self-locating uncertainty, or the inability to precisely pinpoint his position within a possible world, leads to a series of belief updates that are guided by the distance between the current belief (his current location) and other possible locations. The concept of distance in this case refers not only to the physical distance in space but also to the conceptual distance between John's current belief and the ideal belief about his location.

This distance plays a key role in Belief Revision. If we apply Schulz's [96] idea of the “distance between contexts,” we can think of each potential self-location as a "centered world" where John could be located. The farther John's current belief is from the true location, the greater the “distance” he needs to traverse through the revision process. When new information is obtained—perhaps a new landmark or a sudden change in the environment—John's belief about his position is updated accordingly. However, the farther the new information moves him from his previous belief, the larger the shift in his self-locating beliefs becomes.

For John, every new piece of information reduces the distance between his current belief and the true belief about his position. However, unlike standard probabilistic updating, which would simply adjust beliefs based on available evidence, the process of Belief Revision in John's case also involves an ongoing negotiation between various possible worlds (his possible positions in the wilderness) and the distance he must cover to align his belief with reality. In this case, the distance between centers is not just a metaphor for the spatial separation of John from the “truth,” but also a measure of how much new Belief Revision must occur to close that gap. Thus, when we talk about Belief Revision in John's context, we are dealing with both the physical distance (the real, tangible changes in his surroundings) and the conceptual distance (the changing belief that moves him closer to an accurate understanding of his self-location). This dual aspect of distance is crucial in addressing the complexities of Belief Revision, particularly in the context of self-locating beliefs.

By applying these insights to Belief Revision models, we can see that probabilistic models that deal with conditional probability and belief updating provide a quantitative framework for understanding these shifts. However, just as with the Sleeping Beauty Experiment, the nuances of context-sensitive Belief Revision require careful attention to how we represent and track these “distances” in our model. In the next chapter, we will explore how this distance-based approach can help improve the way we handle Belief Revision in scenarios involving self-locating uncertainty, such as the one faced by John the lost ranger.

Chapter 6

Belief Revision and Updating: Probabilistic Approaches

As previously discussed, *Belief Revision* is a formal process that modifies a knowledge base when new information is acquired. This process ensures that an agent's belief system is updated to maintain consistency in light of new evidence. Another approach to handle new information, particularly in a probabilistic framework, is *Updating*. In contrast to Belief Revision, which typically focuses on logical changes to a knowledge base, updating involves the adjustment of an agent's credence distribution based on new information. The standard model for updating belief states is through *Bayesian Conditionalization*, a probabilistic process that adjusts the probability of an event based on the acquisition of new, context-sensitive information. The standard formulation for Bayesian Conditionalization is given by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

This equation shows how the probability of event A, given B, is updated by considering the joint probability of A and B, normalized by the probability of B. In simple terms, conditionalization involves updating the agent's belief about an event (or proposition) given the occurrence of another event. This process relies on the assumption that prior credences (probabilities assigned to various events) are available, and the new evidence is integrated by adjusting the conditional probability. However, there are several scenarios where the application of Belief Revision and Updating leads to different results, particularly when the available evidence is context-sensitive. A well-known example that highlights this difference is the *Sleeping Beauty Experiment*.

6.1 Revising and Updating: Sleeping Beauty Experiment

The Sleeping Beauty Experiment has multiple versions, but the basic setup remains the same. In this thought experiment, Beauty is part of an experiment where the objective is to understand how updating beliefs works in light of new information that causes ambiguity regarding the true state of affairs. The setup is as follows:

- Beauty is put to sleep on a Sunday night.
- An experimenter tosses a coin:
  - If the coin comes up heads, Beauty will be awakened for a short period on Monday.
  - If the coin comes up tails, Beauty will be awakened both on Monday *and* Tuesday.
- Each time she wakes up, Beauty will forget the previous awakening and drink a potion that erases her memory of it.

In this scenario, Beauty's belief about the probability of the coin toss outcome changes as she wakes up and learns whether it is Monday or Tuesday. The core issue with this experiment is how Beauty should update her belief about the fairness of the coin given the new evidence each time she wakes up. Should she assume that the coin toss was more likely to land on heads because she was only awakened on Monday, or should she give more credence to the idea that the coin landed on tails, given that she woke up twice?

The challenge here lies in how to interpret the update process. When Beauty updates her belief each time she wakes up, she is incorporating new evidence based on the temporal context of her awakenings. This makes the problem of belief updating more complex, as it requires understanding how the passage of time and the context of each awakening affect the probability distributions involved. Furthermore, the distinction between Belief Revision and belief updating becomes crucial. While Belief Revision deals with incorporating new beliefs into an existing framework, potentially revising previous beliefs to accommodate the new ones, updating through Bayesian Conditionalization requires calculating the likelihood of new evidence given the previous probabilities. These two processes may lead to different conclusions, especially when we consider scenarios with time-sensitive and context-dependent information, like the Sleeping Beauty Experiment. We will explore the different treatments of the Sleeping Beauty Experiment in the next sections and argue that, by incorporating a more nuanced approach to context-sensitive Belief Revision, we might arrive at a better understanding of how agents should revise their beliefs when faced with scenarios like this.

6.2 Sleeping Beauty Experiment: A Probabilistic Approach

In the probabilistic versions of the Sleeping Beauty Experiment, the central question is: What is Beauty's degree of belief "that the coin toss outcome is Heads?" [27]. Traditionally, there are two possible answers to that question: 1/2 and 1/3. The argument for the 1/3 view goes as follows:

> If we imagine this scenario repeated many times, we would expect approximately half the coin tosses result in heads, and half result in tails. For each toss that results in heads, Beauty will experience one awakening; for each toss that results in tails, she will experience two awakenings. Therefore, she will experience two awakenings in which the most recent toss was tails, for every awakening in which the most recent toss was heads. In a repeated sequence of two-day slumbers, the relative frequency of head-awakenings will be one-third. Since the heads awakenings and tails-awakenings are qualitatively identical, her degree of belief in heads on a given awakening ought to be 1/3. [40]

On the other hand, the argument for 1/2 is as follows:

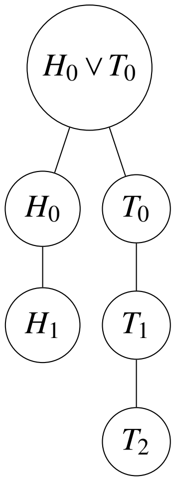
> Since Beauty knows that the coin is fair, her prior probability for H is 1/2. [...] She knows that she will be awakened from her sleep at least once. Therefore, upon awakening, she has gained no new information. So her degree of belief in H should remain at 1/2. [40]

Consider the following centered worlds: On Sunday, before tossing the coin, we have two possibilities, H_0 (Heads) and T_0 (Tails), each with an associated state. Let us consider Adam Elga's analysis of the Sleeping Beauty Problem [27]. According to Elga, we have the following "predicaments":

- H_0: Heads and it is Sunday ($\langle W_H, s\rangle$);
- T_0: Tails and it is Sunday ($\langle W_T, s\rangle$);
- H_1: Heads and it is Monday ($\langle W_H, m\rangle$);

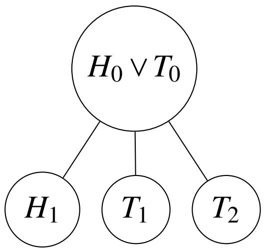
- T_1: Tails and it is Monday ($\langle W_T, m\rangle$);
- T_2: Tails and it is Tuesday ($\langle W_T, t\rangle$).

On Sunday, before tossing the coin, those defending the 1/2 point of view have the following idea in mind: tossing the coin will lead to two possibilities, H_0 or T_0; if it lands Heads, the H_0 "path" is "selected" and Beauty will be awakened only on Monday (H_1); if it lands Tails, the T_0 "path" is "selected" and Beauty will be awakened on Monday (T_1) and on Tuesday (T_2).



$1/2$ *argument*

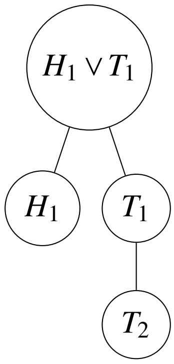
On the other hand, the 1/3 point of view argues that, given the possible frequency of the experiment being repeated several times, Beauty will be awake 1/3 of the times when the coin lands Heads, and so on. That is, on any particular waking, she should have credence 1/3 that it is a Heads-waking, and hence have credence 1/3 in the coin landing Heads on that trial [27].



1/3 argument

Elga highlights an important difference between the arguments: [1] the distinction between being in the state T_1 and being in the state T_2 is not a difference in which possible world is actual, but rather a difference in your temporal location within the world [27]. It is, then, a matter of self-location. However, there is another possibility: since Beauty will be awakened on Monday, the experimenter might choose to toss the coin on Monday, not Sunday. In that case, Beauty is awakened, and the result of tossing the coin is unknown; Sunday is irrelevant to the matter, so we can ignore H_0 and T_0 in the analysis. If the coin lands on Tails, she will awaken one more time.

[1] In parallel, this difference also appears in the distinction between *revisions* and *updates*.



1/3 argument

If the action-guiding reasoning is "Beauty wakes up" and not "the coin is tossed" as above, things are different and relate to the issue of self-locating beliefs presented by Elga [27]. Additionally, if after waking, Beauty can *learn* the result of the coin toss (Heads or Tails), the reasoning process would change. Perhaps, an analysis inspired by the *Revision/Update* approach provided by Katsuno and Mendelzon [47] would help to understand the reasoning involved in our Sleeping Beauty Experiment.

6.3 Updating and the Possible Models Approach

The *Possible Models Approach* (PMA) was introduced by Winslett and, as per Katsuno and Mendelzon, is a suitable way for "reasoning about action in certain applications" [47]. Its operator $\diamond_{pma}$ is considered by Katsuno and Mendelzon as an update operator. For each model I of a knowledge base ψ and a new sentence μ, "the PMA selects from the model of μ those which are 'closest' to I" [47]. Formally, we can define the operation as follows:

$$Mod(\psi \diamond_{pma} \mu) = \bigcup_{I \in Mod(\psi)} \text{Incorporate}(Mod(\mu), I),$$

where Incorporate$(Mod(\mu), I)$ represents the set of models that are "closest" to I in $Mod(\mu)$. The notion of "closeness" between models is inspired by Dalal's idea, where "the distance between two models is simply the number of propositional letters on which they differ" [47]. This provides a formalized means of reasoning about how new information can be incorporated into existing models while maintaining consistency with the model's existing structure.

Let us apply this concept to the Sleeping Beauty Experiment. Let α represent "Beauty is sleeping" and β represent "Beauty is awake"; we can consider the knowledge base $\psi \leftrightarrow ((\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta))$, which informs us that either Beauty is sleeping or Beauty is awake, but not both at the same time. Now, when we wake up Beauty (the action), we update ψ with β. This action represents the knowledge that Beauty is awake, but it's important to note that this action differs from simple Belief Revision, which could involve additional inferences or re-evaluations of the information involved.

Katsuno and Mendelzon highlight an interesting distinction between *Belief Revision* and *Belief Update* [47]. In the case of the PMA, we only know β that is, we only know that "Beauty is awake" without further clarification. However, when applying the Revision Postulates (specifically, **(R2)**), the result of revising ψ with β would be $\psi \wedge \beta$, leading us to the conclusion that Beauty is not sleeping and she is awake, formally represented as $\neg\alpha \wedge \beta$ [2].

This difference becomes particularly relevant when considering the generalization of these concepts. In the Sleeping Beauty example, it may seem intuitive to equate "Beauty is not sleeping" (or "it is not the case that Beauty is sleeping") with "Beauty is awake," but this analogy becomes problematic when substituting α for other propositions, such as χ where "Beauty wears black." In this case, inferring $\neg\alpha \wedge \beta$ after revising ψ with β becomes counterintuitive. Suppose A represents "Today is Monday" and B represents "Today is Tuesday." Beauty is awake and unaware of whether the coin landed Heads or Tails, so upon waking, she knows that "Today is Monday or Today is Tuesday" ($A \vee B$). The action of waking up does not provide much information about the coin toss because she is uncertain of her temporal location. If she learns A, she knows something new, but it doesn't significantly affect her belief about the coin toss, since she could have woken up in either H_1 or T_1. On the other hand, learning B will confirm that the coin landed Tails since she will have learned T_2. Thus, before waking up, Beauty's belief is that $H_1 \vee T_1 \vee T_2$. Upon learning A, it becomes $H_1 \vee T_1$. After learning B, the only consistent possibility is T_2.

If we revisit the action of tossing the coin, two possible scenarios emerge: the coin could have landed Heads or Tails. Beauty's subsequent learning, such as discovering A or B, will lead her to revise her beliefs accordingly. This is particularly important when considering the *self-locating* nature of these beliefs. Learning an indexical expression like "Today is Monday" makes Beauty aware of the context of the utterance and her own temporal location.

PMA works well for updating information in our simple Sleeping Beauty Example, but updates can be defined differently. Let us now consider the *Update* approach, where $\psi \star \mu$ denotes updating ψ with μ. Then:

- **(U1)** $\psi \star \mu$ implies μ.
- **(U2)** If ψ implies μ, then $\psi \star \mu$ is equivalent to ψ.
- **(U3)** If both ψ and μ are satisfiable, then $\psi \star \mu$ is also satisfiable.
- **(U4)** If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$, then $\psi_1 \star \mu_1 \leftrightarrow \psi_2 \star \mu_2$.

[2] Since we have $((\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)) \wedge \beta$

- **(U5)** $(\psi \star \mu) \wedge \varphi$ implies $\psi \star (\mu \wedge \varphi)$.
- **(U6)** If $\psi \star \mu_1$ implies μ_2 and $\psi \star \mu_2$ implies μ_1, then $\psi \star \mu_1 \leftrightarrow \psi \star \mu_2$.
- **(U7)** If ψ is complete, then $(\psi \star \mu_1) \wedge (\psi \star \mu_2)$ implies $\psi \star (\mu_1 \vee \mu_2)$.
- **(U8)** $((\psi_1 \vee \psi_2) \star \mu) \leftrightarrow ((\psi_1 \star \mu) \vee (\psi_2 \star \mu))$.
- **(U9)** If ψ is complete and $(\psi \star \mu) \wedge \varphi$ is satisfiable, then $\psi \star (\mu \wedge \varphi)$ implies $(\psi \star \mu) \wedge \varphi$.

As mentioned earlier, there are significant differences between *Revisions* and *Updates*. One major difference, as Katsuno and Mendelzon [47] point out, is that updates cannot eliminate inconsistencies, whereas revisions can. Additionally, the primary distinction between revision and update is that updates induce a different ordering for each model of ψ, whereas revisions induce a single ordering for the entire knowledge base ψ. This distinction leads Katsuno and Mendelzon to highlight the *local* behavior of updates versus the *global* behavior of revisions [47].

In our Sleeping Beauty example, waking up Beauty can be interpreted as an action. Katsuno and Mendelzon present a way of reasoning about actions using *preconditions* and *postconditions*, where an action is represented as a pair: a precondition encoding the world required for the action and a postcondition describing the effect of the action on the world [47]. They argue that an update operator, as described above, could be used to reason about actions by considering postconditions as new knowledge and preconditions as assumptions.

Moreover, in our version of the Sleeping Beauty Experiment, we assume some kind of ordering. Recall that Ordering is a common feature in Belief Revision. In Katsuno and Mendelzon's approach, $\psi \bullet \mu$ represents "a new knowledge base obtained from an old knowledge base ψ by contracting μ" [47]. The contraction postulates are presented as follows:

- **(C1)** ψ implies $\psi \bullet \mu$.
- **(C2)** If ψ does not imply μ, then $\psi \bullet \mu$ is equivalent to ψ.
- **(C3)** If μ is not a tautology, then $\psi \bullet \mu$ does not imply μ.
- **(C4)** If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$, then $\psi_1 \bullet \mu_1 \leftrightarrow \psi_2 \bullet \mu_2$.
- **(C5)** $(\psi \bullet \mu) \wedge \mu$ implies ψ.

Temporal features are the main difference between revisions and updates, according to Katsuno and Mendelzon. Revisions deal with changes in beliefs about a static world (one that hasn't changed), while updates incorporate changes in the world into the set of beliefs. Katsuno and Mendelzon suggest making the temporal parameter explicit in the knowledge base, where instead of just a theory, the knowledge base is now a pair $< \psi, t >$, with ψ representing a theory and t denoting a time instant [47]. Moreover,

> Instead of two distinct change operations, update and revision, let us introduce a single one called Tell(μ, t), where μ is the new formula to be incorporated, and t is a time instant. The effect of applying Tell(μ, t) to a knowledge base is to *replace* the knowledge base with a new one that incorporates the sentence μ and has time parameter t, unless t is earlier than the KB's time. More precisely, we define the result of applying Tell(μ, t) to $< \psi, t >$ as $< \psi \circ \mu, t >$ if $t = t'$ and $\psi \star \mu, t'$ if $t' > t'$ [3] [47, 7].

This proposal ties together the concepts of update and revision, allowing us to treat both operations under the same umbrella when temporal features are made explicit. Katsuno and Mendelzon's [47] work on Belief Revision and update provides a crucial framework for understanding how agents handle changes in their beliefs when confronted with new, context-sensitive information, which directly relates to the core themes of this thesis. Their distinction between *Belief Revision* and *belief update* helps clarify the processes by which agents either adjust their beliefs to accommodate new information (revision) or integrate new evidence without changing the structure of their existing beliefs (update). This distinction is particularly relevant when considering how agents navigate complex situations, such as the Sleeping Beauty problem, where temporal and contextual factors influence belief formation. Moreover, their work on incorporating temporal parameters into knowledge bases aligns with the thesis's exploration of context-sensitive Belief Revision, providing a foundation for modeling how agents update their beliefs in a dynamic world. By introducing a unified operator for both revision and update, Katsuno and Mendelzon's approach offers a more flexible framework that can be adapted to model the interplay between static and dynamic beliefs, a key aspect of the thesis's investigation into context-sensitive reasoning and self-locating beliefs [47].

6.4 Revising and Betting on Beauty's Beliefs

The Sleeping Beauty Experiment brings some possible solutions but still needs to provide us with a proper way of applying Belief Revision to its case. Because we are left with centered possible worlds, we cannot guarantee that we will be able

[3] Undefined when $t' < t$.

to apply Standard Belief Revision operators. However, it brings a very interesting result: we might not have a unique solution to the problem of revising with indexical information. Considering different centers, we might have different orderings of ordinary possible worlds. Thus, when Beauty is uncertain about her self-locating beliefs, we must find a way of generating a plausible ordering of beliefs. Then, we can apply ordinary Belief Revision operators that actually can give us back consistent belief sets for Beauty. Next, we will see how that can be possible and if there is no unique solution to this issue.

Let us consider the *Sleeping Beauty Experiment* presented once more. In an alternative treatment, we will have three possible centers on the Sleeping Beauty Example:

- c_s: It is Sunday;
- c_m: It is Monday;
- c_t: It is Tuesday.

We will also have two possible worlds w:

- w_h: The coined tossed *Heads*;
- w_t: The coined tossed *Tails*.

We will then construct the following plausible centered possible worlds:

	c_s	c_m	c_t
w_h	$\langle w_h, c_s \rangle$	$\langle w_h, c_m \rangle$	$\langle w_h, c_t \rangle$
w_t	$\langle w_t, c_s \rangle$	$\langle w_t, c_m \rangle$	$\langle w_t, c_t \rangle$

However, the centered world $\langle w_h, c_t \rangle$ is not a possibility for our example in particular, since if tossed Heads the Beauty would not be awakened on Tuesday. Let us take a look at plausible cases of Beauty's uncertainties. We will separate Beauty's uncertainties in three ways: she is uncertain of the world w or the center c or uncertain of both w and c.

Case 1 Beauty is uncertain of the world w, but is certain of the center c

The first case we will consider is when Beauty is uncertain of the world w but is certain of c. In this case, she could be:

(I) Certain of c_s or c_m; thus, her ordering of W could be $w_h \geq_{W'} w_t$ or $w_t \geq_{W''} w_h$. When revising, we need to know which $\geq_W$ Beauty holds before applying standard AGM.

(II) Certain of c_t is the trivial case because there is only one rational possibility: w_t.

However, this case does not illustrate the actual Sleeping Beauty Experiment if we consider that the question asked to her is "Which day is *today*?". Here, Beauty knows which day it is, but it is asked "Did the coin toss Heads or Tails?"

Case 2 Beauty is uncertain of the center c and of the world w

Let us now consider the case when Beauty is uncertain of the center c but is certain of w. In this case, she would be uncertain about ordering $\langle w, c\rangle$. Beauty can have the following orderings of beliefs depending on the center she is located in, where $w \geq_c w'$ indicates that she believes *"the world w is more plausible than the world w'"*:

$\geq_{c_{s'}}$: $w_h \geq_{c_{s'}} w_t$;

$\geq_{c_{s''}}$: $w_t \geq_{c_{s''}} w_h$;

$\geq_{c_{m'}}$: $w_h \geq_{c_{m'}} w_t$;

$\geq_{c_{m''}}$: $w_t \geq_{c_{m''}} w_h$;

$\geq_{c_{t'}}$: $w_t \geq_{c_{t'}} w_h$.

Recall that there is only one possibility for $\geq_{c_{t'}}$. Now, we will combine those orderings into possible sequences of Beauty's belief hierarchy through the experiment:

p_i: $\langle \geq c_{s'}, \geq c_{m'}, \geq c_{t'}\rangle$;

$p_{i'}$: $\langle \geq c_{s'}, \geq c_{m''}, \geq c_{t'}\rangle$;

$p_{i''}$: $\langle \geq c_{s''}, \geq c_{m'}, \geq c_{t'}\rangle$;

$_{i'''}$: $\langle \geq c_{s''}, \geq c_{m''}, \geq c_{t'}\rangle$.

That means, in the first sequence p_i, Beauty will believe on Sunday that the world on which the coin tossed Heads is more plausible than the one in which the coin tossed Tails; on Monday and Tuesday, she will believe the same thing. Thus, her beliefs would have stayed the same over time, and she would answer accordingly. However, in all the other possibilities, her belief orderings will change depending on the day. That disagreement will make Beauty give different answers depending on the day she is awakened.

Suppose, for instance, we want to increase our chances of knowing what Beauty will answer by applying a *Simple Majority Rule*. Then, for all the possibilities above, we will have the $_t \geq_{c^*} {}_h$. This happens because $\geq_{c_{t'}}$ only admits $_t \geq_{c_{t'}} {}_h$ and ends up giving the *casting vote*. If we go this way, we will lean toward the 1/3 argument.

Case 3 Beauty is uncertain of the world w but is certain of c

Let us now consider the case when Beauty is uncertain of the center c but is certain of w. If that is the case, Beauty would know if the coin tossed Heads or Tails, but she would not know which day is today. However, because Beauty understands how the days of the week pass and the rules of the experiment, she could have a respective ordering of beliefs about her centered information.

Consider $\geq_{C^*}$ to be Beauty's plausibility ordering of centers, on which $c \geq_{C^*} c'$ means that Beauty believes that the center c is more plausible than the center c'. In this case, she could be:

(I) Certain of w_h; thus, her ordering of plausible centers in C could be $c_s \geq_{C'} c_m$ or $c_m \geq_{C''} c_s$. Notice that if she knows that the coin tossed Heads, she will also know that c_t is impossible.

(II) Certain of w_t; thus she would have the following possibilities for ordering of C:

C^1: $c_m \geq_{C^1} c_s \geq_{C^1} c_t$;

C^2: $c_m \geq_{C^2} c_t \geq_{C^2} c_s$;

C^3: $c_s \geq_{C^3} c_m \geq_{C^3} c_t$;

C^4: $c_s \geq_{C^4} c_t \geq_{C^4} c_m$;

C^5: $c_t \geq_{C^5} c_s \geq_{C^5} c_m$;

C^6: $c_t \geq_{C^6} c_m \geq_{C^6} c_s$.

Things got more complicated in Case 3 since it is more than just knowing Beauty's beliefs on w. If Beauty only needed to know the world w, then we could suppose that she would also know the contextual features represented by the centers. However, she knows about the world and how the centers could be, and she still needs to find out where she is located within that particular world. This information can come across as new information or upon closer reflection on her beliefs. The Sleeping Beauty case shows us that not all kinds of uncertainties are the same. However, an issue in this short analysis is that if we were supposed to bet on Beauty's beliefs, we might need to consider all the equally possible scenarios. In order to do this, we should aggregate all the possibilities to find a global ordering on which we can revise and make our bets. Nonetheless, as we will see next part, this aggregation might not yield a unique global ordering.

6.5 Belief Ordering and Context-Sensitive Reasoning

The Sleeping Beauty case highlights an important distinction between different types of uncertainty. The problem demonstrates that the way we handle uncertainty in Belief Revision can vary depending on the context and the nature of the information available. In the context of Sleeping Beauty, the key challenge is determining Beauty's degree of belief regarding the outcome of the coin toss. The two main perspectives 1/2 and 1/3 emphasize the complexity of dealing with probabilistic beliefs in situations where the agent's state of knowledge is affected by both temporal and contextual factors. However, an important issue in this short analysis is that if we were to bet on Beauty's beliefs, we would need to consider all equally likely possibilities. This leads us to the task of aggregating all possible scenarios in a way that allows us to make an informed judgment about her beliefs. To address this, one would need to establish a global ordering of possibilities that accounts for all the possible outcomes of the coin toss, as well as the different potential awakenings Beauty could experience. This process of aggregation involves evaluating the relative likelihood of each possible state (i.e., whether Beauty is in a Head state or a Tail state) and subsequently updating her beliefs based on the new evidence she acquires when she wakes up. However, as we will see in the next section, this aggregation may not necessarily yield a unique or singular global ordering of beliefs.

The need for a global ordering is grounded in the idea that beliefs, particularly those about the outcomes of probabilistic events, are often formed based on the available set of possibilities and the agent's position within that set. In a probabilistic framework, such as the one used in the Sleeping Beauty problem, agents must update their beliefs by incorporating new evidence into an existing structure. This process typically involves a form of conditionalization or Belief Revision, where the agent's credence in each possible outcome is adjusted based on the available information. However, the aggregation of these possibilities into a single, global ordering can be problematic, especially when the information is context-sensitive. This issue of ordering beliefs is particularly relevant when considering the implications of context-sensitive reasoning. In the Sleeping Beauty case, the context in which Beauty wakes up whether on Monday or Tuesday affects the way she perceives the outcome of the coin toss. If she wakes up on Monday, she knows that the coin toss could have resulted in either Heads or Tails, but she has not yet learned which. However, if she wakes up on Tuesday, the fact that she has already experienced the first awakening changes the context in which she updates her beliefs. Thus, the ordering of beliefs is not fixed but contingent upon the agent's position within the set of possible states. Furthermore, as we will explore in the next part of the thesis, the difficulty of establishing a unique global ordering of beliefs in the Sleeping Beauty case mirrors a broader issue in Belief Revision: the challenge of aggregating context-sensitive information in a way that respects both the temporal ordering of events and the agent's self-location within those events. This reflects the fundamental distinction between updates and revisions, where updates tend to be local and context-dependent, while revisions aim to provide a more global perspective on belief change. As we proceed, we will see that the aggregation of possibilities in Belief Revision does not always lead to a straightforward, globally consistent ordering. Instead, different contexts and the agent's evolving position within those contexts may yield multiple, equally valid orderings of beliefs. Thus, while the Sleeping Beauty problem provides a useful framework for exploring how agents update and revise their beliefs in response to new information, it also illustrates the complexities involved in constructing a coherent global ordering of beliefs when the available information is context-sensitive. This chapter sets the stage for a deeper exploration of how context, temporal factors, and belief ordering interact in Belief Revision, and how these dynamics might inform more general approaches to probabilistic reasoning and decision-making in uncertain environments.

As we have seen, the Sleeping Beauty problem presents significant challenges in Belief Revision, particularly when attempting to establish a unique global ordering of beliefs. This difficulty arises due to the context-sensitive nature of belief updating, where an agent's beliefs are influenced not only by new information but also by their position within the set of possible outcomes. The question of how to aggregate these context-sensitive beliefs into a coherent ordering is central to understanding the complexities of Belief Revision and decision-making under uncertainty. This discussion naturally leads us to the next chapter, where we explore how such belief orderings can inform decision-making processes, particularly in the context of social choice theory. One of the most profound results in social choice theory is Arrow's impossibility theorem, which deals with the aggregation of individual preferences into a collective decision. Just as the Sleeping Beauty problem illustrates the challenges of belief aggregation in uncertain contexts, Arrow's theorem demonstrates the inherent difficulties in aggregating preferences in a group setting.

In Arrow's framework, the goal is to find a voting system that can consistently aggregate the individual preferences of members of a society into a collective decision that respects certain fairness criteria, such as non-dictatorship and independence of irrelevant alternatives. However, as Arrow's theorem shows, no such system can satisfy all these conditions simultaneously, highlighting the tension between individual preferences and collective choice. This problem mirrors the challenges in Belief Revision, where the task of aggregating context-sensitive beliefs into a single, coherent ordering is similarly fraught with complications. The connection between Belief Revision and Arrow's theorem lies in the fundamental issue of how to reconcile multiple, potentially conflicting perspectives or beliefs into a unified decision. Whether we are updating beliefs in the Sleeping Beauty experiment or aggregating preferences in a social choice framework, we face the same underlying problem: how to structure a process that accommodates diverse and context-dependent information, ensuring that the resulting outcome is both consistent and fair.

In the next chapter, we will explore how the principles of Belief Revision and the insights gained from the Sleeping Beauty problem can be applied to the domain of social choice theory. The challenges of aggregating beliefs in uncertain and dynamic contexts provide a compelling framework for analyzing the complexities of collective decision-making. Just as individual agents must revise their beliefs in response to changing information, groups and societies face the challenge of incorporating diverse individual perspectives into a coherent and collectively acceptable decision. This process is fraught with tensions between maintaining internal consistency and adapting to new information, mirroring the structural challenges observed in epistemic updates. The formal tools developed in Belief Revision theory, including the handling of context-sensitive updates and the role of probability in self-locating beliefs, offer a useful lens for addressing issues in collective choice, particularly when considering the role of fairness, rationality, and preference aggregation.

A central theme of our next discussion will be the tension between local updates—where individual agents modify their beliefs based on their own contextual information—and global consistency, where these beliefs must be reconciled into a stable and fair collective outcome. This tension becomes especially salient in the context of Arrow's impossibility theorem, which demonstrates the inherent difficulties in aggregating individual preferences into a collective decision while maintaining certain rationality conditions. The impossibility of designing a flawless aggregation mechanism mirrors the constraints faced in Belief Revision, where no single update mechanism can perfectly balance coherence, new evidence, and contextual adaptation. By examining these parallels, we gain a richer understanding of how beliefs and preferences evolve within groups and how formal constraints shape the structure of collective reasoning. The next chapter will further develop these connections, shedding light on how Belief Revision models can inform social choice theory and how Arrow's theorem helps us understand the structural limitations of rational decision-making at both the individual and collective levels.

Part IV

A New Approach

Chapter 7

Impossibility of Choosing Beliefs

Choosing which beliefs we should give up while revising requires making *choices*. Moreover, since we often hold on to uncertain beliefs, the resulting revised set might not always be unique. Rational choice functions are precisely those that can be construed based on given underlying preference relations, including orderings such as *Epistemic Entrenchment*. When making choices, it is important to consider preference orderings. Therefore, some interesting results can bring clarity to the issue: *Which beliefs should one give up when revising with indexical information?* This question is central to the process of Belief Revision, particularly when it comes to the challenges posed by context-sensitive information.

To fully understand the significance of making choices in Belief Revision, we need to start by recognizing that choices involve more than just the act of selecting one belief to retain or discard. In the context of Belief Revision, a choice is often guided by underlying preferences about which beliefs are more entrenched, more reliable, or more central to the agent's knowledge base. When an agent is faced with new, sometimes conflicting information, the revision process requires the agent to weigh the relevance and robustness of their existing beliefs. The agent must determine which beliefs should be prioritized and which beliefs should be adjusted or discarded. This process involves assessing the *degree of belief* (or *credence*) in existing propositions, which is often influenced by context and self-location. Making choices in this context is not just a matter of logic; it is an exercise in managing epistemic uncertainty and deciding which beliefs are most consistent with the new information while preserving coherence in the belief system.

Recall that the standard AGM Postulates were developed for *uncentered beliefs*. These postulates typically assume that beliefs are fixed and not influenced by the agent's specific context within the world. To revise beliefs with centered information properly, however, one must first select a proper possible world to be the *sphere-centered* one. We know that not all propositions can be represented as sets of ordinary possible worlds. The issue arises from the fact that standard

Belief Revision does not account for *centered worlds*. In other words, there is no formal framework within standard AGM postulates to revise with information that includes self-locating or context-sensitive propositions, such as those that depend on the agent's perspective or the time and place of the belief's formulation.

To bridge this gap, we propose moving from centered information to a global ordering of possible worlds on which we can apply the AGM Postulates. The first step involves selecting a plausible ordering of ordinary worlds from a list of centered worlds. This approach is more nuanced than simply substituting context-sensitive expressions. The key idea here is that an ordinary proposition, in all acceptable alternatives (possible worlds), will be true in one particular world from those. The selected world must be shared or accepted by the most relevant centers and included in the choice set. This ensures that the center most representative of the agent's beliefs is taken into account when revising the knowledge base. After selecting the appropriate possible world, the agent will revise over the global ordering. If it turns out that the *sphere-centered world w* is not the case, the agent will return to the global ordering and select the next closest possible world. It is essential to remember that ordinary possible worlds are treated as those on which we can add any center and remain unchanged. This flexibility allows us to adapt the ordinary possible worlds framework to the needs of Belief Revision involving centered information.

In this way, the agent's policies to revise remain rational and consistent with AGM Postulates for Belief Revision, even if the agent adopts nonmonotonic inferences within her belief set. This is because the revisions will only admit consistent information (and we want that). The process of Belief Revision, then, ensures that new information, even if context-sensitive, is incorporated in a way that preserves the integrity of the agent's belief system. The key challenge is to bring centered information into the framework in a consistent manner, ensuring that the revision process remains rational and coherent. In this chapter, we will explore the role of choice functions in rationally selecting the best alternatives for Belief Revision and the challenges that arise in doing so. The analysis will be informed by Arrow's Impossibility Theorem, which provides important insights into the conditions under which rational choice functions can be defined. Arrow's Theorem will guide our understanding of the constraints that must be satisfied for a rational choice function to exist and offer plausible conditions for selecting beliefs in the revision process.

7.1 Rational Choices

Social Choice Theory is the study of collective decision-making, and it plays a crucial role in understanding how individual preferences can be aggregated to determine a collective decision. One of the key aspects studied in this field is the interaction between *individual* and *collective* preferences or utilities. The *Paretian Property*, a desirable criterion in social choice theory, posits that whenever individual welfare increases, collective welfare should follow suit, and vice versa [6]. This concept emphasizes the importance of ensuring that decisions reflect both individual preferences and the collective welfare.

When making decisions about which beliefs to revise, we often rely on preference orderings, such as degrees of belief, to rank and compare different elements of a set. However, there are instances where comparing elements is not straightforward or even relevant. For example, we might prefer chocolate ice cream over vanilla, and Janet's party over Phillip's, but comparing the preference for ice cream with the preference for a party seems entirely unrelated. This highlights a challenge in collective decision-making, where preferences must be aggregated into a *social welfare function*, a function that characterizes rational ethical belief choices over a set of individual preference orderings. When applying a social welfare function, it is possible that the resulting choice set includes multiple maximal elements, leading to non-uniqueness. This could, in some cases, result in an empty choice set, as some would claim happens when preferences regarding politicians are involved.

In the context of Belief Revision, we face a similar challenge when selecting which beliefs to revise based on context-sensitive information. Just as in collective decision-making, it is important to establish a rational process for revising beliefs, taking into account the preferences and priorities that should guide the selection. We define an *alternative* as any element in the set over which the choice function operates, and in the case of Belief Revision, these alternatives might be potential revisions or contractions of a knowledge base. Let us define a *choice function* in the standard way, as its characteristics will be applied to the revision process when constructing contraction and revision operators:

Definition 7.1.1 (Choice Function) *Let X be a finite set of possible alternatives, and let $\Gamma \subseteq X$ be the set of all non-empty subsets of X. A choice function $Ch : \Gamma \to \Gamma$ assigns to each and every $Y \in \Gamma$ a non-empty subset $Ch(Y)$.*

As Sen [98] points out, the existence of a choice function is a requirement for rational choice. If a choice function is available, it is possible to rationally select between two or more alternatives. Consistency is another key requirement for choice functions. A choice function is considered *rationalizable* if it is both complete and transitive [6]. Additionally, if we have a weak ordering $\leq$ giving rise to a choice function Ch and $\leq$ is reflexive and complete, a necessary and sufficient condition for a choice function to be defined over a finite set X of alternatives is that $\leq$ should be *acyclical* over X. One major challenge when using choice functions is the problem of non-uniqueness: it is possible to have multiple maximal elements in the resulting choice set, meaning that a single, clear choice may not always emerge.

Social Choice principles can be applied to a wide range of fields, with their adaptable informational features enabling the integration of new concepts and approaches. Traditionally, a set of individuals $i \in V$ is considered to be a set of *voters* [70], who each express their preferences over a set of alternatives. When considering two alternatives x and y, a specific set of voters V_d is said to be *decisive* if, whenever $x \leq_{p_i} y$ for every $i \in V_d$, we have $x \leq_p y$ as the aggregated preference ordering.

A qualitative preference ordering can be translated into a *utility function*, which plays a critical role in probabilistic approaches to social choice. One of the most celebrated results in social choice theory is *Arrow's Impossibility Theorem*, also known as the *Theorem of the Impossibility of Democracy* [70]. This theorem, which earned Arrow the Nobel Prize for his contributions to Economics and Welfare Theory, identifies several necessary conditions for rationality when aggregating individual preferences into a global ordering. This result is not only foundational in social choice theory but also has profound implications for other domains, including Belief Revision. We will explore the details of Arrow's Theorem in the next chapter, examining how it informs our understanding of rational decision-making and the aggregation of preferences, particularly in complex contexts like Belief Revision.

7.2 Rational Choices and Belief Revision

Rational choices are central to our study of Belief Revision, as both share key concepts and insights. Rational Choice Theory, in particular, provides a detailed investigation into how choices are made and what constitutes a rational choice. Rott [90] argues that the principles underlying Belief Revision can be derived from the principles of Rational Choice. In both cases, we aim to make decisions that optimize certain criteria, whether those are personal preferences or the coherence of our belief system.

The set of alternatives and preference orderings serve as the foundation for defining a choice function. Traditionally, these preferences and alternatives are assumed to be context-insensitive, meaning they are considered fixed and unchanging in relation to the specific context. Rational Choice, in this view, would deal with preferences over context-insensitive claims. However, the challenge presented in the context of Belief Revision with indexical information is to extend these concepts to accommodate context sensitivity. Our task is to understand how we can update our beliefs rationally in the presence of context-dependent information.

Rational criteria for making choices can be applied at two levels: the semantic level (on the models) and the syntactic level (on the propositions) [90]. Rott [90] suggests that these criteria can be applied interchangeably and yield the same set of plausible consequences for belief change operations. In a scenario where an agent is presented with a set X of alternatives, we say that X constitutes the *menu*. When a choice function Ch is applied to the menu X, it returns a choice set $Ch(X)$ of the elements in X that can be considered the best choices. Choice functions can be applied to multiple menus simultaneously, provided they are all subsets of X [90]. In Belief Revision, a choice function represents the agent's *disposition* to choose between alternatives at a given time. Nonmonotonic inferences and revision operations can be seen as analogous to choice functions:

> An inference operation Inf returns the set of conclusions for a vast field of potential premise sets H, and it is itself independent of the particular premise set which an agent entertains at a certain time. Similarly, a belief operation $*$ returns the revised belief set for a vast field of potential inputs and is independent of the particular proposition that happens to come in as input. [90, 6]

As we noted, choice sets often contain multiple best elements. In some cases, we will encounter *ties* between equally preferred alternatives, as well as *incomparabilities* between alternatives. This can necessitate the use of non-logical mechanisms to resolve which option to select. Rott [90] presents two main strategies for handling such situations:

- The agent can pick a *random* element from $Ch(X)$, as the alternatives are equally preferred.
- The agent can look for a satisfactory combination of the preferred alternatives. However, this approach might come with challenges. If, for example, x represents the proposition A and y represents $\neg A$, combining these two would lead to a contradiction, violating the consistency requirement.

Rott [90] offers a compelling interpretation of choice sets with multiple alternatives. He suggests that such sets represent *potential choices* that an agent can make. If we interpret the choice set $Ch(X)$, the agent's rational best alternatives will be restricted to those that are equally preferred. In many cases, the agent will choose only one best element, as in voting for a new president. In some instances, however, the agent might be able to select a combination of rational alternatives. For instance, in some voting systems, voters can select a political party instead of an individual politician. In this case, the best alternatives within the set would be selected.

If no satisfactory alternatives can be found, the agent might be forced to randomly choose among the best elements. This introduces the *Principle of Satisficing* [90], which suggests that when facing preferred alternatives within a belief set $Ch(X)$, an agent should choose based on fixed acceptability criteria, even if there might be better alternatives:

Definition 7.2.1 (Principle of Satisficing) *When facing preferred alternatives within a belief set $Ch(X)$, an agent should choose based on fixed acceptability criteria, even if there might be better alternatives.*

If all the alternatives in a menu are satisfactory, they will be selected as the best choices by the selection function Ch and considered *good enough alternatives*. To ensure that these alternatives are rational, certain constraints must be satisfied.

Definition 7.2.2 (Faith) $\exists R \neq \emptyset \forall X \in \mathcal{X}$: *if* $X \cap R \neq \emptyset$ *then* $Ch(X) = X \cap R$

Definition 7.2.3 (Indiscrimination Condition) $\exists R \forall X \in \mathcal{X} : Ch(X) = X \cap R$

Finally, if there are no satisfactory alternatives, the agent can rationally refuse to choose and end up with an empty set. However, it is considered a *desideratum* for choice functions $Ch(X)$ to return non-empty choice sets when the original set X is non-empty [6]. Intuitively, if there is a unique alternative, it is the *best element*. Therefore, choice functions should be *successful* or *decisive*, formally defined as follows:

Definition 7.2.4 (Successful Choice Function) *If* $X \neq \emptyset$, *then* $Ch(X) \neq \emptyset$

Rott [90] outlines several *coherence constraints* on choice functions applied to various menus. These constraints are designed to provide additional rationality criteria based on the consistency requirement:

- **(Heritage)** If $Y \subseteq X$, then $Y \cap Ch(X) \subseteq Ch(Y)$.
- **(Concordance)** $Ch(Y) \cap Ch(X) \subseteq Ch(Y \cup X)$.
- **(Aizerman's Axiom)** If $Y \subseteq X$ and $Ch(X) \subseteq Y$, then $Ch(Y) \subseteq Ch(X)$.
- **(Strict Heritage)** If $Y \subseteq X$ and $Ch(X) \cap Y \neq \emptyset$, then $Ch(Y) \subseteq Ch(X)$.

We can also define the *taboo sets* of a choice function Ch:

Definition 7.2.5 (Taboo Set) $\overline{R}$ *is a taboo set of* Ch *if the following two conditions hold for all* X*:* ***(t1)*** $Ch(X) \cap \overline{R} = \emptyset$ ***(t2)*** *If* $Ch(X) = \emptyset$, *then* $X \subseteq \overline{R}$

We will use the taboo set in formalizing the selection of best elements based on a preference relation $\leq$, as presented by [90]. Let Ch be a choice function with a taboo set $\overline{R}$:

$$Ch(X) = smallest_{\leq}(X) = \begin{cases} \{x \in X : x \leq y \text{ for all } y \in X\} & \text{if } S \not\subseteq \overline{R}, \\ \emptyset & \text{if } X \subseteq \overline{R}. \end{cases}$$

If we consider a strict preference relation $<$, we can formalize it as follows:

$$\sigma(X) = min_{<}(X) = \begin{cases} \{x \in X : y < x \text{ for no } y \in X\} & \text{if } S \not\subseteq \overline{R}, \\ \emptyset & \text{if } X \subseteq \overline{R}. \end{cases}$$

Moreover, a preference relation $\leq$ is said to be *connected* if for every $x, y \in X$, either $x \leq y$ or $y \leq x$. The relationship between choice functions and strict preferences can be described as follows. Let $Pow(\overline{R})$ be the *power set* of $\overline{R}$. A choice function Ch with taboo set $\overline{R}$ is called relational with respect to a strict preference relation $<$ over X, formally $Ch = X(<)$, if for every $X - Pow(\overline{R})$:

$$Ch(X) = min_{<}(X).$$

This ensures that, if the menu contains at least one element that is not taboo, the choice function Ch will always pick the minimal elements of the set of alternatives X under the preference relation $<$ [90]. This requirement ensures that, in situations where there are valid alternatives, the choice function is guided by a consistent preference relation and selects the best alternative from the available choices, even when there are competing options.

Moreover, the relationship between choice functions, preference relations, and Belief Revision is quite robust. Choice functions provide a mechanism for rational decision-making that can be applied to Belief Revision processes, especially in situations where nonmonotonic reasoning or context-sensitive information is involved. By adopting the principles and formal structures developed in rational choice theory, Belief Revision can be modeled as a process of selecting the most rational alternatives in response to new information, while maintaining consistency with the agent's existing beliefs. Moreover, the concept of taboo sets and relational choice functions can help refine our understanding of how agents handle situations where contradictions or ties arise, ensuring that revisions are not only logically sound but also practically applicable to complex decision-making scenarios.

Through the use of these formal models, we can better understand the dynamics between belief change and choice, as well as the challenges that arise when dealing with multiple, potentially conflicting beliefs. The framework laid out by choice theory, particularly through the principles of rationality, consistency, and preference ordering, offers valuable insights into how agents navigate the complexities of revising their beliefs, especially when confronted with new and potentially contradictory context-sensitive information. The next section will explore how these concepts are applied within the framework of Arrow's Impossibility Theorem and its implications for Belief Revision in dynamic contexts.

7.3 Choices and Centered Information

Choice functions, as defined above, assume a context-insensitive treatment for choosing context-insensitive alternatives. To properly handle centered information, we must extend these definitions accordingly. Weber [110] argues that deriving centered information from uncentered ones can be counterintuitive if no additional information is learned. Centered information often relies on contextual features or parameters, which allow us to make sense of indexical statements within Belief Revision.

The particularities concerning indexical expressions and the centered worlds approach need further exploration in the literature on Belief Revision. However, some relationships between Belief Revision functions and other functions have already been established. The relationship between Belief Revision and choice functions is closer than one might initially think. For example, according to Rott, choice functions especially those involved in nonmonotonic reasoning approaches are comparable to Belief Revision operations [90]. He suggests that a choice function represents the disposition of an agent to choose, given a set of alternatives [90]. In this sense, a choice function returns a set of the "best choices" from a set of alternatives. However, in Belief Revision, what is needed is a function that provides a ranking of alternatives, which would allow us to construct revision and contraction functions based on this ordering. A ranking closely resembles a welfare function in Arrow's terms [8], which can provide insights into how these rankings influence the revision process.

According to Grove [37], the relation $\leq$ can be interpreted as a ranking of how *compatible* possible worlds are with an agent's current beliefs. With this interpretation, a Belief Revision operator $*$ that respects the AGM rational postulates for Belief Revision can be defined. In our Belief Revision framework for languages containing indexical expressions, we first assume that each center corresponds to an individual ranking of the most preferred possible worlds or those considered more plausible by the agent. This ranking represents the agent's disposition to revise her beliefs. From these individual orderings of possible worlds, the goal is to construct a global ordering of worlds to which a Belief Revision operator can be applied.

Consider the case where an agent is uncertain about the date. She knows either today is March 5th or March 6th, and that a solar eclipse will occur on one of these days. The center provides the agent's location within a possible world: the information about which day it is today. The possible world represents information about what that world is like, including the occurrence of the solar eclipse. The agent can then learn new information that may force her to update her beliefs. When considering indexicals, learning the centered information "Today is March 5th" and "Today a solar eclipse will occur" may not represent the same belief as learning non-centered information, such as "The solar eclipse will happen on March 5th," because the agent does not know if today is actually March 5th. Substituting the indexical expression with a non-indexical co-referential expression could change matters of cognitive significance, and thus the belief one ought to hold after acquiring new evidence [78].

Now, consider the eclipse example again, assuming all the beliefs an agent holds are categorical. Suppose the agent has evidence suggesting it is more plausible that the solar eclipse will occur today than it will not. Since the agent does not know which day it is, she faces a decision about how to aggregate these beliefs into a context-insensitive ordering of possible worlds. This aggregation issue is connected to the problem of judgment and *preference aggregation* [1]. For example, suppose we have three individuals, i_1, i_2, and i_3, each with their preference rankings for alternatives X, Y, and Z as follows:

$i_1 : Y \leq Z \leq X$
$i_2 : Y \leq X \leq Z$
$i_3 : Y \leq X \leq Z$

All individuals rank alternative Y as the least-preferred, so intuitively, we might expect the aggregated preference order to rank world Y as the least-preferred. However, what about the aggregated preferences of X and Z? One might argue that a voting method like the *Majority Rule* could be used, leading to an aggregated ordering of $Y \leq X \leq Z$. But alternative methods might lead to different aggregated orders.

When considering Belief Revision with Centered Worlds, every center c corresponds to an individual ranking of possible worlds w. Aggregating these individual orderings into a global ordering of worlds is similar to preference aggregation, with choices about how to rank worlds in terms of plausibility. However, this aggregation process is not straightforward and may involve selecting a global ordering from competing individual preferences. Moreover, Arrow's Impossibility Theorem from Social Choice Theory offers important insights into the constraints and challenges of such aggregation methods. In our case, Arrow's Impossibility Theorem might provide insights into how we could rationally define a Belief Revision operator for a non-indexical sublanguage, given the context-relative Belief Revision operators for a language containing indexicals. The challenge lies in rationally defining a context-insensitive ordering of possible worlds to which we can apply standard Belief Revision, as no such universal ordering exists when we include context-sensitive, centered information.

Let us now shift to a more vivid example. Consider John , a lost ranger alone in the wilderness, unsure of his exact location but certain that he is somewhere within a defined region. His survival depends on his ability to gather accurate information and revise his beliefs about his position. Initially, John knows that he could either be on Mountain Peak A or Mountain Peak B . Each peak offers a different set of possibilities: if he's on Peak A, he can see the solar eclipse that is due to occur, but if he's on Peak B, his chances of witnessing the eclipse are slim. Let's consider the relevant centered worlds for John's situation:

- c_1: John is on Mountain Peak A and it's March 5th.
- c_2: John is on Mountain Peak B and it's March 5th.
- c_3: John is on Mountain Peak A and it's March 6th.
- c_4: John is on Mountain Peak B and it's March 6th.

From each center, John has a ranking of the possible worlds:

- From c_1 (Peak A, March 5th), John prefers w_1 (the eclipse occurs on March 5th) to w_2 (the eclipse does not occur on March 5th), and w_2 is preferred to w_3 (the eclipse occurs on March 6th).

- From c_2 (Peak B, March 5th), John prefers w_2 (the eclipse does not occur on March 5th) to w_3 (the eclipse occurs on March 6th), and w_3 is preferred to w_1 (the eclipse occurs on March 5th).

- From c_3 (Peak A, March 6th), John prefers w_1 (the eclipse occurs on March 5th) to w_2 (the eclipse does not occur on March 5th), and w_2 is preferred to w_3 (the eclipse occurs on March 6th).

- From c_4 (Peak B, March 6th), John prefers w_3 (the eclipse occurs on March 6th) to w_2 (the eclipse does not occur on March 5th), and w_2 is preferred to w_1 (the eclipse occurs on March 5th).

From these individual rankings of possible worlds, we aim to construct a global ordering of the possible worlds. The challenge here is similar to the aggregation problem in social choice theory. We have a mix of preferences across different centers, and aggregating them into a global ordering requires addressing conflicts between the centers' preferences. For instance, if John learns that the eclipse is imminent, he would update his beliefs based on his current center. Suppose the agent's ranking from c_1 and c_2 suggests that the eclipse on March 5th is more likely to occur, but the agent must also reconcile this with information from other centers, such as c_3 and c_4. This process is where Belief Revision functions come into play, as they guide the agent in reconciling different belief sets and formulating a rational update to her knowledge.

John's challenge mirrors the one faced when revising beliefs in a context-sensitive setting: deciding how to aggregate information across multiple possible worlds and centers. As with preference aggregation in social choice theory, there is no simple, unique global ordering, and choices must be made based on the relevant context whether it's the proximity of the centers to the agent's current belief or the plausibility of the worlds involved. In that sense, the task of Belief Revision in centered worlds involves aggregating contextualized beliefs into a global ordering that is both rational and consistent. By examining the relationship between Belief Revision, centered worlds, and choice functions, we gain insight into how agents revise their beliefs under uncertainty and how these revisions can be formally structured to align with rational decision-making frameworks.

7.3.1 Arrow's Desirable Conditions and Centered Worlds

Arrow provides us with some desirable conditions for rationally constructing a global ordering of possible worlds. These conditions are meant to ensure that the aggregation of individual preferences into a collective, context-insensitive ordering respects the rationality of the decision-making process. However, Arrow's Impossibility Theorem demonstrates that it is impossible to satisfy all these conditions simultaneously in a fair and consistent way when there are multiple agents with potentially conflicting preferences. As Arrow himself states, "No social decision function can simultaneously satisfy all of the following conditions: unrestricted domain, non-dictatorship, Pareto efficiency, and independence of irrelevant alternatives" [8]. This means that while we can strive to satisfy most of the conditions, we must be prepared to make compromises when constructing the global ordering. *Prima facie*, we would have a few options to resolve this problem:

- We could ignore some of the profiles, potentially losing some information but simplifying the process;
- We could accept the resulting ordering as partial, meaning that some alternatives may not be ranked in the global ordering;
- We could restrict our global ordering to two relevant possible worlds, focusing only on the most pertinent alternatives and ignoring the others;

- We could select a center where the context-insensitive ordering would follow, prioritizing a particular perspective or frame of reference in the Belief Revision process.

Each of these approaches comes with trade-offs. For instance, ignoring some of the profiles might make the aggregation process easier, but it could result in a loss of important information about the preferences of some individuals or centers. Similarly, a partial ordering might be more manageable but could leave us with incomplete or unclear conclusions about the relative merits of the available alternatives. Restricting the ordering to only two relevant worlds might help streamline the decision process but could sacrifice the richness of the model. Lastly, selecting a center for the context-insensitive ordering might simplify the aggregation process but could lead to bias if the chosen center doesn't fairly represent all perspectives.

Any option we choose might not allow us to construct a Belief Revision function that satisfies all rationality criteria. Revising beliefs with indexical information complicates the process further, as it introduces the need to manage context-sensitive expressions like "today," which inherently involve both possible worlds and centers. As Weber [110] points out, deriving centered information from uncentered ones can be counterintuitive when no new contextual information is provided, emphasizing the difficulty of defining a globally rational Belief Revision function that handles both world-related and context-related dimensions of belief.

It is important to highlight that we may not arrive at a unique method for revising beliefs when indexicals are involved. The process of selecting which conditions to drop whether it's ignoring some profiles, accepting a partial ordering, or focusing on a specific center will depend on the specific context and goals of the Belief Revision process. These choices will be explored later in this thesis.

When indexical expressions like "today" occur in a sentence, we encounter a two-dimensional feature to consider: the possible worlds and the centers. The use of indexical expressions necessitates a more sophisticated approach to Belief Revision. To avoid getting caught in the relativity of beliefs with respect to different centers, we aim to construct a global ordering of possible worlds that aggregates individual-centered rankings in a way that yields a context-insensitive Belief Revision operator. Ideally, the output of this aggregation should be a preorder that guarantees a corresponding Belief Revision operator, as in Grove's model [37]. This global ordering would allow us to revise beliefs in a way that is context-insensitive and consistent across different centers, but the question remains: Is there a unique rational way of defining such an ordering?

To answer this question, we might turn to Arrow's famous result on Social Choice Theory, which sheds light on how we can rationally aggregate preferences while adhering to desirable conditions. Arrow's traditional framework and the Impossibility Theorem were initially presented in [8], and they are well explored in traditional literature [1]. Arrow's conditions for a fair and rational social decision function have important parallels in our context, particularly when dealing with centers and possible worlds. We will examine Arrow's conditions and their relevance to our approach to Belief Revision, discussing how they help illuminate the challenges of constructing a globally rational ordering in the presence of indexicals and centered information.

In our setting, centers (or contexts) correspond to individuals (or voters) in Arrow's framework; we take our set of possible worlds to be the set of finite alternatives; context-relative orderings (or individual-centered orderings) correspond to Arrow's individual orderings; our intended context-insensitive ordering corresponds to Arrow's social ordering. Then, let the worlds $w_1, ..., w_n$ be the alternatives and $W = \{w_1, ..., w_n\}$ be the set of all possible worlds. Consider that every possible center c_i provide individual preorder $\leq_{c_i}$. Let the profile be the list $\langle \leq_{c_1}, ..., \leq_{c_m} \rangle$ of total context-relative preorders of the set W of worlds, one for each centre $c_1, ..., c_m$. We will write $\langle \leq_i \rangle$ instead of $\langle \leq_{c_1}, ..., \leq_{c_m} \rangle$. Now let F be the function that assigns for each profile $\langle \leq_i \rangle$ in some domain a binary relation $F\langle \leq_i \rangle$ on W. The function F in our setting corresponds to Arrow's *social welfare function* that ranks the alternatives in the social ordering. We will write $\leq_W$ instead of $F\langle \leq_i \rangle$. Arrow sets some rational constraints (conditions) to the social welfare functions. The conditions represent some desirable features to guide the construction of a social ordering and are open to debate. Arrow's conditions intend to provide a scheme for aggregating preferences from distinct profiles. First, we will consider Arrow's constraints corresponding to our setting and, later, debate whether they are desirable.

7.4 Arrow's Impossibility Theorem for Centered Worlds

The application of Arrow's impossibility theorem to centered worlds introduces a novel framework for understanding the aggregation of preferences and beliefs in context-sensitive settings. Traditional formulations of Arrow's theorem address the difficulties of constructing a fair and rational collective decision-making system when individual preferences must be aggregated into a single social ordering. The impossibility result demonstrates that, under reasonable conditions of rationality, independence, and non-dictatorship, no voting system can perfectly

[1] See [8], and [70] for a concise introduction to the topic.

aggregate individual preferences while satisfying all fairness criteria simultaneously. However, standard formulations of Arrow's theorem operate under the assumption that preferences are uncentered—meaning they do not inherently depend on the self-locating perspective of individual agents within a given informational or epistemic context. In contrast, in a centered world framework, preferences and beliefs are not merely abstract rankings but are anchored in agents' self-location within a possible world, influencing their epistemic access to information and their perception of the choice space. This contextual dependence creates new challenges in defining social choice functions that respect both individual self-locating beliefs and collective rationality constraints. The adaptation of Arrow's theorem to centered worlds thus raises key questions: How does the structure of self-locating beliefs influence the aggregation process? Can traditional fairness and consistency criteria be maintained when individual beliefs about their own position in the world play a role in their preferences? And, crucially, does the impossibility result extend to contexts where agents' preferences depend on their shifting epistemic standpoint? To explore these issues, we now state and analyze the conditions required for an *Arrow's Impossibility Theorem for Centered Worlds*, formalizing how preference aggregation operates in scenarios where self-location plays a fundamental role in decision-making. Consider now the following conditions stated for an *Arrow's Impossibility Theorem for Centered Worlds*:

Unrestricted Domain (U): The domain of F includes every list $\langle \leq_i \rangle$ of context-relative total preorders of possible worlds.

Intuitively, (U) seems desirable since there is no reason to exclude a particular list of individual orderings. Moreover, it guarantees that every profile will be considered and that any kind of context-relative ordering for each center c_i is possible. Since there is no logical way of selecting which individual orderings would be included on the profile list, only considering some possible list is to ignore relevant centers.

Global Ordering (GO): For any profile $\langle \leq_i \rangle$ in the domain, $\leq_W$ is a context-insensitive total preorder of possible worlds.

This condition guarantees that there will be a corresponding context-insensitive ordering that gives rise to a Belief Revision operator for the possible worlds. If we did not guarantee that, then $\leq_W$ might not be a total preorder, and, therefore, we could not construct a Belief Revision function for the set of possible worlds.

Weak Pareto (WP): For any profile $\langle \leq_i \rangle$ in the domain of F, and any alternatives

w_j and w_k, if $w_j \leq_i w_k$ then $w_j \leq_W w_k$.

Weak Pareto is an intuitive constraint that says given any alternatives w_j and w_k, if all $\leq_{c_i}$ on the profile agree that $w_j \leq_{c_i} w_k$, then the global ordering will also agree $w_j \leq_W w_k$. For a finite number of alternatives and a finite number of contexts, if all the individual orderings rank the possible worlds the same way, it is rational to suppose that the context-insensitive ordering follows.

Arrow's conditions consider the idea that in a democracy a *dictator* should not exist, that is, a voter whose individual ordering determines the social global ordering no matter what the other individual orderings are. We say the center c_d is a *dictator* of F if for any alternatives w_j and w_k, and for any profile $\langle ..., \leq_{c_d}, ...\rangle$ in the domain of F: if $w_j \leq_{c_d} w_k$, then $w_j \leq_W w_k$

Nondictatorship (D): F has no dictator.

Nondictatorship states that the global ordering $\leq_W$ would not follow the ordering $\leq_{c_d}$ of a single-center c_d. That is particularly useful for the Social Choice Theory counterpart for evident reasons, but some authors defend the computational applications of having a dictator [105]. In our framework, having a dictator would mean that a given center has a privileged epistemic position. For instance.

To state the following condition, we will need the notion of *restriction*: for any given relation R, and any set S, let $R|S$ be the *restriction of R to S*, that is, the part of R concerning just the elements of S.

Independence of Irrelevant Alternatives (IIA): For all alternatives w_j and w_k in W, and all profiles $\langle \leq_i \rangle$ and $\langle \leq_i^* \rangle$ in the domain of F, if $\langle \leq_i \rangle | \{w_j, w_k\} = \langle \leq *_i \rangle | \{w_j, w_k\}$, then $\leq_W |\{w_j, w_k\} = \leq_{W'} |\{w_j, w_k\}$, that is, if the two profiles are the same concerning the given alternatives, then the respective global orderings concerning those alternatives will be the same.

Independence of Irrelevant Alternatives is usually the first condition to be reconsidered when discussing Arrow's Theorem in the Social Choice Theories. Although it seems intuitive, **IIA** does not consider that the orderings of the other alternatives might play a role in the choice set. Given n alternatives $w_1, w_2, ..., w_n$, suppose we have two possible orderings for given profiles $\langle \leq_i \rangle$ and $\langle \leq_i^* \rangle$: $w_2 \leq_i ... w_n \leq_i ... \leq_i w_1$, and $w_2 \leq_i^* w_1 \leq_i^* ... \leq_i^* w_n$. In our setting, it could be a possible condition. The above orderings lead to the global orderings $w_2 \leq_W w_1$ and $w_2 \leq_{W'} w_1$, respectively, when considering only the alternatives w_1 and w_2. However, the first profile prefers w_1 over *all* other alternatives, while the second prefers w_1 only over w_2.

Finally, we can state an Arrow Impossibility Theorem for Centered Worlds:

Theorem 11 ***Arrow's Impossibility Theorem**:*

Suppose that are more than two alternatives. No function F satisfies U, GO, WP, D, and IIA.

Proof: The original proof can be found in [8]: Arrow first shows that there is a decisive voter and then proceeds to show that that voter is a dictator.

On the one hand, Social Choice Theory has shown that when there are only one or two alternatives, the Impossibility Theorem does not apply. Nevertheless, having more than two accessible possible worlds allows us to reason about theories, counterfactuals, and hypotheses, so it seems intuitive that is having more than two possible worlds would be the case most of the time. On the other hand, when there is an infinite number of voters, Fishburn [30] shows that the problem proposed by Arrow's Impossibility Theorem also disappears. The finite number of centers could have practical motivations. Allowing an infinite number of centers would mean that every imaginable center should be considered, including the impossible ones. Kirman and Sonderman [50] shows that if there are infinite voters or centers, we end up with "invisible dictators". They are considering that we will work with a finite number of possible worlds and contexts, Arrow's Impossibility Theorem states that not all conditions can be simultaneously satisfied. In that case, we debate which conditions we could drop.

7.5 Dropping Arrow's Conditions

If not all Arrow's conditions can be satisfied simultaneously, we might have to choose to drop one or more conditions. Indeed, there is a vast debate on Social Choice literature about which one we should drop or the possible ways out of Arrow's Impossibility Theorem. One famous suggestion was given by Sen [98] and considers that we should attach utility to the alternatives. Since that solution would find a probabilistic counterpart and our focus is on a qualitative version of the problem, we will not discuss that solution. Instead, let us consider each condition and possible implications for dropping it.

The first condition we shall consider dropping is **(U)** Universal Domain, stating we should consider every possible profile. Recall that a profile is a list of individual orderings of possible worlds given by the centers. We can consider, for instance, that not all possible profiles would be acceptable. Some profiles would contain impossible preorders, contradictory with previous background knowledge, or logic violators (like impossible worlds [12]. Dropping **(U)** would mean restricting the possible profiles to the rational ones. That would have significant practical applications when constructing a relevant context-insensitive global ordering. Indeed, condition **(U)** states that the centers can rank the worlds; however they " wish " , there are no restrictions on how the center should order the worlds. Constraints on the ordering, if any, would be extra-logical.

(GO) Global Ordering condition tells us that the resulting aggregated ordering would be a total preorder. As mentioned before, without that requirement, we could not guarantee that we would be able to construct a Belief Revision function that respects AGM Postulates. If we do not require the global ordering to be a total preorder, we could still end up with a preorder that is not total. Constructing weaker Belief Revision functions that do not respect all AGM postulates would be possible. A consequence of relaxing this condition and ending up with a preorder has been presented as a possibility to get out of Arrow's Impossibility Theorem. We, then, could end up having oligarchy [111]. For our initial purposes, respecting AGM postulates seems desirable, and dropping **(GO)** might not be a practical solution.

The most intuitive condition seems to be **(WP)** Weak Pareto. **(WP)** provides us an opportunity of defining a global ordering when all the individual orderings rank two given alternatives the same way. Justifying that the global ordering would disagree is counterintuitive if every profile ranks a possible world w_1 as more preferable than a possible world w_2. It is hard to find a practical example on which the global ordering would rationally go against the individual orderings. Perhaps, that would only happen if a stronger universal law or moral constraints would play a part. Suppose, for instance, that all individual orderings rank $w_1 < w_2$ where w_1 represents a world on which bikes can fly and w_2 a world on which bikes cannot fly. It would not be rational for a global ordering to accept individual orderings that go against the law of gravity. However, this seems like an unpractical example. Therefore, if the individual orderings $\leq_{c_i}$ of the centers given distinct profiles disagree, then we need to find another way of determining $\leq_W$. The standard Social Choice Theory argues that there is no unique way of defining such a $\leq_W$, which is a challenge we also face.

Arrow's theorem is thought for a finite number of voters. However, considering only one voter or center, we could have a weaker version of **WP**, called the *Pareto Principle*, which fits the idea of having an accidental dictator. The Pareto Principle is very intuitive since it states that the global ordering will follow the unique individual ordering available:

(P) Pareto Principle: *If there is exactly one center c, then the global ordering should follow the individual ordering of the center c.*

Suppose we give up the Nondictatorship **(D)** condition. Now we must pick up an epistemically privileged center to be a dictator. Which center should have such power? Extra-logical considerations play a part here. One could say, for instance, that c_2 is our epistemically privileged center and, therefore, our dictator. The global ordering, then, would be $w_2 \leq_W w_1$; that is, the agent believes that the eclipse occurring on March 6th is more plausible than the eclipse occurring on March 5th. Moreover, since c_2 is the dictator, we assume the agent i knows she is located on March 6th. As we can see, we need to rely on additional information to construct such a context-insensitive ordering, and there might be several distinct possible ways of choosing a dictator.

Consider another example with three different centers representing the three possible instants Monday, Tuesday and Wednesday, respectively: c_m, c_t, c_m. Then, we assume the following individual orderings:

$$c_m : w_1 \leq_{c_m} w_3 \leq_{c_m} w_2 \leq_{c_m} w_4$$
$$c_t : w_1 \leq_{c_t} w_4 \leq_{c_t} w_2 \leq_{c_t} w_3$$
$$c_w : w_4 \leq_{c_w} w_1 \leq_{c_w} w_3 \leq_{c_w} w_2$$

Although in the context of Social Choice, having a dictator in a democratic system is not a good idea, in our context of centered worlds, it might not be so bad to drop condition **(D)**: that means we could have a dictatorial center. A dictatorial center could represent an *epistemically privileged* center (it could be an individual or instant of time or even a position). What makes a center epistemically privileged may vary: in some situations, for instance, a later time could have an epistemically privileged position of having more information about past events; when talking about beliefs about myself, the center which I am the individual might be the epistemically privileged position; and so on. One would require some extra logical information to determine that epistemically privileged position, as stated above. For instance, we could learn that c_w is our privileged epistemic center and, therefore, our dictator. In that sense, the global ordering $\leq_W$ would follow $\leq_{c_w}$: $w_4 \leq_W w_1 \leq_W w_3 \leq_W w_2$. Interpreting the epistemically privileged position of the center c_w is extra-logical and not particularly the focus of this article. After constructing such global ordering, we can revise using traditional AGM methods. However, it does not necessarily have to be that way.

Instead of having a single center as a dictator, it could be useful to consider some oligarchy where more than one center plays a role as a dictator. Formally, such an oligarchy could be represented as a multi-centered world, an ordered pair consisting of a possible world w, and an n-tuple of centers or individuals. In that sense, it might be interesting to construct a multi-centered world to be a dictator. For instance, when most centers agree on an ordering given some alternatives, that becomes the individual ordering of the multicenter. For precise and straightforward cases that might work; however, that is not possible in some cases, and the centers will not agree like in our example. If that is the case, we cannot construct the global ordering and, therefore, would only be able to revise with centered information (revision on our original $\leq_W$). Moreover, we should consider what that would mean to the resulting global ordering [111].

Another solution to Arrow's Impossibility Theorem is to follow several criticisms presented in Social Choice Theory and drop condition *(IIA)* instead of *(D)*. In that case, if two profiles are the same concerning the given alternatives, then the respective global ordering concerning those alternatives does not have to be the same; that is, intuitively, **(IIA)** says that when the function F is applied to two distinct profiles given a restriction to given alternatives, if the profiles agree on the ordering of the restricted alternatives, then the global ordering must yield the exact preference ordering and any differences between the profiles considering the other alternatives are irrelevant. Criticizing condition **(IIA)** is a primary objection to Arrow's Impossibility Theorem on Social Choice theories. Indeed, the other ordering possibilities might not be relevant to constructing the global ordering we want.

Consider our example again. The individual ordering that constitutes the profile is given according to the individual rankings given by the centers. As *(U)* states, centers could rank alternatives as they " wish". Thus, we could have another individual ordering of possible worlds. To clarify the next point, let us restrict the number of alternatives to two. If we restrict the orderings to the alternatives w_1 and w_2, we would have the following ordering for the first profile and also for the second profile:

$c_m : w_1 \leq_{c_m} w_2$
$c_t : w_1 \leq_{c_t} w_2$
$c_w : w_1 \leq_{c_w} w_2$

Since the ordering of the two alternatives is the same for both profiles, the respective global orderings would be the same concerning only the worlds w_1 and w_2. That would give us a way of relating different orderings of an agent's dispositional beliefs. Suppose more than one relevant profile agrees on ordering two given alternatives that give rise to global orderings. Those global orderings could be interpreted as "more plausible to be believed " by the agent. Other interpretations might also be considered if we consider that we might have privileged epistemic centers and that could be explored in further works.

As we can see, selecting a unique procedure to determine aggregation policies or global ordering that fulfill some Arrow-like conditions like those stated above is challenging. Social Choice theories have presented many voting methods to provide such procedure, including the *Plurality Rule*. In our language, the Plurality Rule can be stated as follows:

(Pl) Plurality Rule: *Each center selects one world to be the most preferred, and the world selected by the most centers is the most preferred on the global ordering.*

The Plurality Rule seems a plausible way of selecting a world given individual orderings of the centers as defined above. Nevertheless, it has an evident limitation: it does not mention the global ordering of the other alternatives on the ranking. We need a ranking to give rise to a Belief Revision function as we saw above so that it could be more interesting for our purposes. Another problem with the Plurality Rule is that it can select a *Condorcet Loser*.

However, other principles could be helpful for an agent doing Belief Revision. The *Principle of Satisfying* [2] in [90], for instance, states that a rational agent should take any alternative (world) that meets some fixed criteria of acceptability even if she knows that are better options. That means, for instance, that she could accept some center as epistemically privileged even if she knows there might be another more epistemically privileged center. The Principle of Satisfying will not give us which fixed criteria of acceptability an agent should adopt but might provide a possible path to construct a global ordering when the individual orderings do not agree.

Intuitively, we also might face the issue that the ordering of possible worlds might be influenced by the agent's ranking of the centers c. Suppose, for instance, an agent has a ranking of centered possible worlds. Given that ordering $\leq_{W_C}$ of centered worlds, every center c_i will determine an individual preorder $\leq_{c_i}$ of possible worlds. Formally:

Definition 7.5.1 $w \leq_{c,\leq_{W_C}} w'$ *if and only if* $\langle c,w\rangle \leq_{W_C} \langle c,w'\rangle$.

In order to not depend on the centers, constructing a context-insensitive ordering might be a good option. If we end up with a total pre-order, we can construct a Belief Revision function following AGM postulates. One might argue, for instance, that we could define such Belief Revision function given the following preorder correspondences: $w \leq_W w^*$ if, and only if, there are c, c' such that $\langle c,w\rangle \leq_{W_C} \langle c',w^*\rangle$ or, perhaps, $w \leq_W w^*$ if, and only if, for all c, c' such that $\langle c,w\rangle \leq_{W_C} \langle c',w^*\rangle$. However, which would be the right way to define? Are those the only possible

[2]Herbert Simon 1959

definitions? Do those definitions give us global orderings that are total preorders? There is no unique way of constructing a context-insensitive ordering $\leq_W$ of possible worlds given an ordering $\leq_{W_C}$ of centered worlds. However, it might be interesting to investigate the possibility. In Social Choice Theory, ranking pairs of voters and alternatives could show, for instance, that the vote of some specialists towards a determined subject might be more preferred than others. So far, that solution should have been investigated in the standard literature.

One criticism of Arrow's Impossibility Theorem is Sen's [98] "enriched information basis ". Sen argues that the profiles alone do not provide enough input to apply a function F as above, mainly because preference orders are purely ordinal. Sen suggests we use utility functions for individuals instead of individual preference orderings and argues that it is possible to construct a global ordering given enough information. For our purposes, it is intuitive to think that we could use probability functions, that is, credences, as the new input. However, a more qualitative approach might be more philosophically interesting. We could investigate which information might enrich our profile in that case.

Given a profile of individual orderings, partial solutions can be investigated since there is no unique method to construct the global ordering of possible worlds. One option is to see if the so-called *Abstract Aggregation*, that is, the aggregation of binary evaluations like those presented by Wilson in [113] and explored by Dokow and Holzman in [26] could provide a possible solution.

We show that for each possible way of assigning a reference to an indexical expression, we can construct a Belief Revision operator based on the individual ordering of possible worlds provided by each possible center. We want to construct a global ordering of possible worlds from that profile that allows us to define a Belief Revision operator for a sublanguage without indexical expressions. Arrow's Impossibility Theorem shows there is no unique way of constructing such a global ordering. Since it is impossible to fulfill all the Arrow conditions above simultaneously, one or more conditions must be dropped. Dropping the condition **(D)** and **(IIA)** might be interesting in our context of Belief Revision, as we saw.

It is possible, then, to construct a non-unique rational context-insensitive preference ordering of possible worlds facing analogous issues to Arrow's Impossibility Theorem; that is, the Arrow's Theorem applies when we try to aggregate preference orderings of distinct profiles of individual-centered orderings and to keep certain desirable conditions we might end up with dictatorial centers or even oligarchies. Dictatorial centers are epistemically privileged ones, and extra-logical criteria are required to determine them. From individual orderings given by the centers, we have possible profiles to which the aggregation rules must be applied. We might rule out profiles portraying impossible lists of orderings, for instance. In that case, constructing a global ordering would be more straightforward. A challenge we might face is investigating the possibility of recovering the individual orderings of possible worlds given by the centers. In our framework, considering the agent i and the solar eclipse example, that might be possible if we know which center is the privileged one. One possible application for recovering individual orderings is investigating how an agent's beliefs about the centers might change their choices.

Arrow's Impossibility Theorem applied to the centered possible worlds frame shows us that it is impossible to aggregate individual-centered orderings when generating a global ordering of ordinary possible worlds. Consider the example below to see the impact of this result on Belief Revision. Consider Jane, the detective, is investigating a murder. She gathered all the evidence needed, and now she has five days to write a final report on who the killer is. Suppose we have four ordinary possible worlds to be considered:

w_A: the individual A is the killer;

w_B: the individual B is the killer;

w_C: the individual C is the killer;

w_D: the individual D is the killer.

Suppose also that there are four possible centers for the detective, representing the days before the deadline on Friday. On the first day, she will write down her initial hierarchy of beliefs about who the killer is. The next day, she will analyze the data again and write down her hierarchy of beliefs about who the killer is the same for the third and fourth days. Finally, on the deadline, she will consider her orderings from previous days to construct a global ordering for the final report. Thus, we will have the following centers:

c_1: first day to write the report, that is, Monday;

c_2: second day to write the report, that is, Tuesday;

c_3: third day to write the report, that is, Wednesday;

c_4: fourth day to write the report, that is, Thursday;

c_5: Deadline, that is, Friday;

Consider $\geq_{c_n}$, for any $n = \langle 1, 2, 3, 4 \rangle$, to be a plausibility ordering on which $w \geq_{c_n} w'$ indicates that the center c_n considers the ordinary possible world w to be more plausible than the world w'. Let us, then, consider the following centered individual orderings:

$\geq_{c_1}$: $w_A \geq_{c_1} w_B \geq_{c_1} w_C \geq_{c_1} w_D$;

$\geq_{c_2}$: $w_A \geq_{c_2} w_B \geq_{c_2} w_C \geq_{c_2} w_D$;

$\geq_{c_3}$: $w_A \geq_{c_3} w_C \geq_{c_3} w_D \geq_{c_3} w_B$;

$\geq_{c_4}$: $w_C \geq_{c_4} w_D \geq_{c_4} w_A \geq_{c_4} w_B$.

Those orderings tell us the following story:

- On Monday, the detective believes that it is more plausible that the killer is individual A; if that is not the case, then the closest plausible killer is individual B, then C, and; finally, the least plausible killer is individual D.
- On Tuesday, she still believes the same as the day before.
- On Wednesday, upon further consideration [3], she considers that individual A is the most plausible killer. However, now she believes that B is the least plausible option.
- On Thursday, she believes that individual C is the killer; if not, then the closest option would be individual D. Notice that now she does not believe the most plausible killer is individual A like the days before.

How should she aggregate those or whatever four orderings she might have come up with on the four days into a global order for the final report on Friday?

In standard AGM Belief Revision, we would answer that the detective should only consider her last ordering, that is, the ordering provided by c_4. However, we are talking about the possibility of sending someone innocent to prison: she wants to be certain about it. Evidence justified her previous ordering of beliefs but resulted in distinct orderings. In order to be fair, the detective decides to consider the last orderings of beliefs when writing her final report. First, let us construct the possible profile considering the individual orderings above:

$\langle \geq_{c_i} \rangle$: $\langle \geq c_1, \geq c_2, \geq c_3, \geq c_4 \rangle$

Here, we will consider a single profile; according to Feldman and Serrano, a Single-Profile is enough to provide the necessary diversity for applying Arrow's theorem [28]. In this profile, $\langle \geq_{c_i} \rangle$ four individual orderings do not agree: thus,

[3] Note that here she could change her mind with autoepistemic tools due introspection, she is not learning anything new.

we cannot apply the *Pareto Rule*. Furthermore, we need to drop one condition to construct a global ordering.

Let us start by considering dropping the condition **(U)**. We would have to exclude one of the individual orderings $\geq_{c_n}$. However, there is no reason to do so: the detective wrote down those possible orderings based on the evidence she gathered during the investigation. As discussed before, dropping **(GO)** is not an option: without it, we could not guarantee adequate conditions to apply Belief Revision operators on the global ordering. Moreover, we want **(GO)** to be the ordering of the final report. We cannot drop **(WP)** either: it would be inconsistent for the detective to believe that it is the case that two or more individuals were the killer when evidence suggests only one was responsible. As expected, we turn our attention to the two remaining possibilities:

(1) We drop **(D)**;

(2) We drop **(IIA)**.

On the one hand, if we drop **(D)**, we would have to select a center to behave as a dictator. Thus, we would need a justification to select such a dictator. The global ordering would follow the dictator's ordering. On the other hand, if we drop **(IIA)**, we could consider two or more profiles. That is not possible in this case since we only have a single-profile frame.

Therefore, the detective cannot construct a global ordering that simultaneously fulfills Arrow's conditions. She would be forced to pick a center to be a dictator. In that case, she could pick the last possible center, namely c_4, as her global ordering when writing the final report. By doing so, she will act according to standard AGM Belief Revision (although she could choose any other center to be the dictator as they are equally preferred).

7.6 Revising with Uncertainties

Consider that no agent knows everything there is to know about the actual world. That means an agent has an ordering of beliefs about how the world is or could be. Indeed, since we have several possible centers within the same possible world, not all times, spatial positions, or individuals will be relevant for an agent's revision policies. If we want to decide which w we believe in, we would have to aggregate specific centered information to select our best alternative. There is no unique answer for how we could do that. If Stalnaker [99] is correct, then this approach intends to provide a way of characterizing the proposition just in terms of ordinary possible worlds. It is that way, a matter of what the world looks like. We assume that every uncentered possible world w can be extended to a centered possible world $\langle w, c \rangle$, but not all $\langle w, c \rangle$ can be reduced to w. However, we can consider several centers to create a global ordering of w that we can consider while revising. Thus, we are suggesting a generalized approach that yields no unique solution. Instead, it will consider extralogical information to determine which centers should be considered plausible or relevant while constructing the global ordering.

It is irrelevant the kind of sensitiveness we are considering: about time, spatial position, or individuals. The main challenge is to choose the contextual parameters we need to address adequately. For instance, the point of evaluation is not necessarily the point of utterance. Then, we need to consider that both centers are relevant while constructing the global ordering. Like in Quine [83], we focus on centered content and, once a center is fixed, the problem disappears because we have our *dictatorial center*. However, because it is not always possible to fix a unique center, we need to have a rational ordering of plausible contexts in case we need to select a *new dictator*. The distance between contexts presented by Schulz [96] provides us a chronological way of rationally ordering contexts. With this, we can fix the point of evaluation to be the one that might intervene with selecting the *epistemically privileged center*. Moreover, from K-Contexts we can get to S-Contexts: first, we need to consider plausible centers to select individual orderings and construct an ordering of ordinary possible worlds. However, how we choose will depend on the uncertainty we are considering.

We explored the implications of revising beliefs in the presence of uncertainties, focusing on the role of centered worlds and the challenge of constructing a global ordering. We discussed how agents' beliefs about the world depend on both time, space, and individual centers, and how decisions about which belief to revise are influenced by the contexts in which those beliefs are held. The key takeaway is that revising beliefs with context-sensitive information, especially in the presence of indexicals, creates challenges in creating a unique and rational revision process. We also highlighted the importance of contextualizing beliefs by introducing ideas from Social Choice Theory, particularly Arrow's Impossibility Theorem. Arrow's conditions provide important insights into how we might attempt to aggregate the individual rankings of possible worlds based on different centers of evaluation. The key issue, as discussed, is that no method exists to guarantee a unique global ordering of possible worlds, especially when centered information is introduced. However, Arrow's theorem presents a framework for understanding the limitations and challenges in creating such an ordering and emphasizes the need for rational choice in a context-sensitive setting.

Decision-making becomes a particularly complex process when we introduce uncertainties and multiple contexts into the equation and it is a very interesting approach to be investigated in further works. The challenge is to ensure that decisions are made rationally, even when the information available is partial, uncertain, or context-dependent. Arrow's theorem highlights the inherent difficulties in aggregating preferences and making collective decisions in such contexts. The debate around revising beliefs in the presence of indexical information suggests that rational decision-making must take into account not only the agents' beliefs but also the centers and contexts in which those beliefs are formed. We must decide which centers to prioritize, which contextual information is most relevant, and how to aggregate the individual beliefs into a coherent global ordering that can guide rational revision. Decision making in the presence of uncertainty is inherently complex, as it involves the challenge of making rational choices without complete information. Uncertainty arises when agents lack full knowledge about the current state of the world, the future outcomes of their decisions, or even the context in which they are operating. In these situations, agents must rely on available information, beliefs, and preferences to make the best possible choice. The uncertainty in decision-making processes can take many forms, including probabilistic uncertainty, where outcomes are known but the likelihood of each outcome is uncertain, and epistemic uncertainty, where the very knowledge about the world is incomplete or vague.

The application of Arrow's impossibility theorem to centered worlds introduces a novel framework for understanding the aggregation of preferences and beliefs in context-sensitive settings. Traditional formulations of Arrow's theorem address the difficulties of constructing a fair and rational collective decision-making system when individual preferences must be aggregated into a single social ordering. The impossibility result demonstrates that, under reasonable conditions of rationality, independence, and non-dictatorship, no voting system can perfectly aggregate individual preferences while satisfying all fairness criteria simultaneously. However, standard formulations of Arrow's theorem operate under the assumption that preferences are uncentered—meaning they do not inherently depend on the self-locating perspective of individual agents within a given informational or epistemic context. In contrast, in a centered world framework, preferences and beliefs are not merely abstract rankings but are anchored in agents' self-location within a possible world, influencing their epistemic access to information and their perception of the choice space. This contextual dependence creates new challenges in defining social choice functions that respect both individual self-locating beliefs and collective rationality constraints. The adaptation of Arrow's theorem to centered worlds thus raises key questions: How does the structure of self-locating beliefs influence the aggregation process? Can traditional fairness and consistency criteria be maintained when individual beliefs about their own position in the world play a role in their preferences? And, crucially, does the impossibility result extend to contexts where agents' preferences depend on their shifting epistemic standpoint? To explore these issues, we now state and analyze the conditions required for an *Arrow's Impossibility Theorem for Centered Worlds*, formalizing how preference aggregation operates in scenarios where self-location plays a fundamental role in decision-making.

The specificity of uncertainty is crucial in shaping how agents approach decision-making. Different types of uncertainty may require different decision-making strategies. For instance, in cases of probabilistic uncertainty, Bayesian decision theory provides a useful framework by updating beliefs based on new evidence to form more accurate predictions. However, in situations where the uncertainty is epistemic, more sophisticated strategies are required that account for how an agent's knowledge evolves over time, especially when new, context-sensitive information is introduced. Indexical information such as "today," "here," or "I" further complicates the uncertainty, as the meaning of such expressions depends on the context in which they are used. In such cases, the agent's decision-making process must be sensitive to the center (such as time or location) from which the information is being evaluated. Considering the particularities of uncertainties helps to ensure that the decision-making process is not only rational but also adaptable to the nuances of the situation. For example, when dealing with epistemic uncertainty, where the agent's knowledge of the world is incomplete, the agent may need to rely on context-sensitive Belief Revision to update their beliefs in light of new information. In contrast, when facing probabilistic uncertainty, the decision-maker may rely more heavily on probability distributions to weigh the potential outcomes of each choice. By recognizing the specific nature of the uncertainty involved, the agent can tailor their decision-making process to better accommodate the complexities of the situation. This leads to more informed, contextually appropriate decisions, which are particularly important in dynamic, real-world scenarios where uncertainty is pervasive.

Moreover, the process of aggregating uncertain information, especially when considering multiple centers or perspectives, presents an additional challenge in rational decision-making. In scenarios where the decision depends on conflicting or incomplete information from different sources, such as in the aggregation of centered beliefs, the agent must carefully weigh the relevance and plausibility of each source. This requires a global ordering of beliefs that accounts for the specific contexts in which each piece of information is relevant. As discussed in Arrow's Impossibility Theorem, constructing a rational aggregation of preferences under uncertainty is difficult and may lead to non-uniqueness, further complicating the decision-making process. Understanding these challenges helps to frame the problem of decision-making under uncertainty as one of not just choosing the best alternative but also navigating the complexities of how uncertainty interacts with the decision environment. These issues set the stage for the next chapter, where we explore how context-sensitive uncertainties interact with Belief Revision, particularly within a two-dimensional framework. The need to model belief change in response to shifting contextual features leads us to a structured approach for two-dimensional Belief Revision, incorporating both world-relative uncertainty and agent-relative self-location. By integrating these insights, we aim to bridge the gap between individual belief dynamics and the broader challenges of collective decision-making under uncertainty.

Chapter 8

Revising Context-Sensitive Uncertainties

Human reasoning is inherently full of uncertainties. However, not all types of uncertainty stem from the same sources. Sometimes, we are uncertain because of vagueness, leading to borderline cases. Other times, uncertainty arises from ambiguity, and in some situations, we are simply uncertain due to a lack of sufficient information to support our beliefs. Despite this, the absence of certainty does not seem to prevent us from drawing conclusions and taking action. We often reserve the right to change our minds whenever we encounter new contradictory information.

Uncertainty regarding how the world *actually* is could be alleviated by verifying the world and revising our beliefs accordingly. However, when it comes to context-sensitive claims, we face an additional layer of complexity. In these cases, it is not enough to check the world alone; we also need to consider the changes of context that might occur within the world.

At a fundamental level, we can categorize uncertainty concerning centered propositions into three main types:

- We can be uncertain about the world w,
- We can be uncertain about the context, i.e., the center c,
- We can be uncertain about both the world w and the center c, which pertains to uncertainty about the entire centered proposition $\langle w, c \rangle$.

These types of uncertainty have been illustrated by the Sleeping Beauty Experiment. Recall that the *Centered Possible Worlds* approach extends the traditional possible worlds framework in modal logic. While possible worlds generally represent complete ways the world could be, centered possible worlds introduce the notion of a *center*, focusing on a particular individual, point in time, and spatial position within that world. This is often represented as a triple $\langle t, p, i \rangle$, where t denotes the time, p denotes the spatial position, and i represents the individual. This approach is particularly useful for capturing the subjective perspective of an individual within a possible world, which is essential for representing context-sensitivity or, more specifically, indexicality.

When we are uncertain about the world w, but a certain or fixed center c is given, this situation reduces to being uncertain about an uncentered proposition, and standard AGM Belief Revision is sufficient. The problem arises when we are uncertain about the center c, meaning that the contextual information is imprecise, incomplete, or unknown. In such cases, we face two additional possibilities:

- (1) We can be uncertain about the center c, but certain of the world w,
- (2) We can be uncertain about both the center c and the world w.

For instance, consider the following table with two plausible worlds and two plausible centers:

	c^1	c^2
w^1	$\langle w^1, c^1 \rangle$	$\langle w^1, c^2 \rangle$
w^2	$\langle w^2, c^1 \rangle$	$\langle w^2, c^2 \rangle$

As we can see in the table, when there is a fixed center, there is only one possible ordering of worlds given by that particular center. If that is the case, no further steps are needed to generate a global ordering. However, things become complicated when we do not have a fixed center. If we are certain of the world w^1, we now have two possible centered propositions, $\langle w^1, c^1 \rangle$ and $\langle w^1, c^2 \rangle$, which corresponds to the situation described in case (1). In this case, we must select the best element for our belief set in a non-standard way by creating a global order that considers both centers. This could happen, for example, if both centers agree on the ranking of the worlds. However, if the order given by one center does not agree with w^1 being preferred, we would choose the other center.

If we are still uncertain about both the centers and the worlds (case 2), we will need to generate a global ordering. But, as we have seen in previous chapters, constructing a global ordering that satisfies desirable conditions, such as those in Arrow's Impossibility Theorem, is impossible.

What should we do to rationally revise our beliefs when we are uncertain about the contextual information?

To address this, we will examine several examples presented throughout the work, considering only the *K-contexts* that are plausible for our framework. Recall that a *K-context* is a quadruple $\langle w, t, p, i \rangle$, where w is an ordinary possible world, t

is a time, p is a spatial position, and i is an individual. Uncertainties about t, p, and i will relate to uncertainties about the centered information, that is, uncertainties about c, regardless of the certainty status of w. These are the *context-sensitive uncertainties*. On the other hand, when the agent is uncertain only about w but not about c, we deal with context-insensitive uncertainty.

This distinction is crucial because uncertainties that are not contextual fall into the realm of standard Belief Revision, which requires little to no changes. However, when we face uncertainties about contextual information, we must adopt a new approach to Belief Revision. We know that constructing an ordering of ordinary possible worlds from a list of centered individual orderings is impossible, as shown by our application of Arrow's Impossibility Theorem. Nonetheless, we must explore alternative solutions to deal with this challenge effectively. In our case, we face context-sensitive uncertainties that cannot be dealt with by merely applying traditional Belief Revision methods. The solution requires understanding how contextual information interacts with Belief Revision, considering both the world and the center. Arrow's Impossibility Theorem shows the inherent complexity of defining a unique global ordering from individual-centered orderings. Hence, this chapter has demonstrated that revising beliefs in the presence of context-sensitive uncertainties requires a more nuanced approach that accommodates the intricacies of context and the impossibility of constructing a single, definitive global ordering. Therefore, the next steps involve exploring alternative models and approaches that allow us to rationally revise beliefs while acknowledging the limitations posed by context-sensitive information. Let us keep in mind that decision-making processes must incorporate the recognition of these uncertainties and the non-monotonic nature of context-sensitive Belief Revision. The choice of context, like any decision, is shaped by a dynamic interplay of rational criteria and contextual relevance, which must be addressed in any decision-making process that involves revising beliefs in the face of uncertainty.

8.1 Revising Context-Sensitive Uncertainty

When an agent is uncertain about which possible world she is in, the agent might need access to sufficient elements to determine which world it is. For Stalnaker [102], all cases would be reduced to uncertainty about the world. Moreover, this is not the case: some kinds of uncertainty cannot be concentrated on identifying the actual ordinary possible world.

Consider, for instance, our Sleeping Beauty example. She was certain of her position at the possible world w when she woke up, no matter if the coin landed Tails or Heads, because she did not forget about being in an experiment. Beauty is still determining her position within that possible world, her center. Thus, she is uncertain about c.

Because she is uncertain about c, the Sleeping Beauty case raised much debate. Other examples of uncertainty about c also allowed much discussion in the literature. There are cases in which the agent might be uncertain not only about the ordinary possible world but also about her own location in that world. Thus, the agent is uncertain of $\langle w, c\rangle$. Given C as the set of all centers c, W the set of all ordinary possible worlds w, and any proposition A, B, in a Language, we will start by assuming that any given agent has a hierarchy of plausible beliefs as follows:

- 1. An agent has an ordering of plausible ordinary beliefs on which the most plausible world (that is, the one she is certain of) is the smallest member of the set of centered possible worlds on which makes a proposition A true.
- 2. An agent has an ordering of plausible centers within C.
- 3. There is a finite number of plausible possible worlds and centers.

Those assumptions intuitively say that, given an ordering of centered beliefs, the agent will have a Two-Dimensional hierarchy as follows:

	c^1	c^2	...
w^1	$\langle w^1, c^1\rangle$	$\langle w^1, c^2\rangle$	...
w^2	$\langle w^2, c^1\rangle$	$\langle w^2, c^2\rangle$	...
...	...	...	...

Recall that if an agent is certain about her own center c, she will fix that center as the only plausible one on the ordering of centered possible worlds $\langle w^n, c\rangle$, for $n \geq 1$, and revise over the ordering determined by the ordinary possible worlds. If an agent is certain about the ordinary possible world w, then all possible centers should yield that w is the most preferred ordinary possible world.

Let us think about the possible cases, considering our three kinds of uncertainty. First, if an agent is certain of c, then she is fixing her ordering of beliefs to be the same as the centered individual ordering provided by c. By fixing c, unless you have explicit evidence that c is not the case, we are in the realm of Standard AGM Belief Revision and should consider only the ordering of ordinary possible worlds w. Things get more complicated if an agent is certain of w, but uncertain about c, since she knows how the world is but might need to be more certain about her own position within that world. For instance, this can happen when we lose track of time or position. Recall Alice and Bob's example: Alice is not uncertain about how the world is; she just lost track of the hours. Moreover, it is worth noticing that she also knows the ordering of instants, which makes her correct her beliefs and adapt the indexical expression used to utter them. For those cases, we might resource to testing an ordering of *intended contexts* to check if a given center yields the world w we are certain of. This ordering of intended contexts might consider the chronological ordering of instances, the spatial distance between positions, and/or the limited list of plausible individuals in a context. If that is the case, we have first an ordering of intended or plausible contexts, represented by centers, on which impossible cases are excluded, and some contexts are considered most-preferred over others. Nevertheless, for each of those centers, we also have a centered individual ordering of acceptable/preferred/plausible ordinary possible worlds. That is the ordering on which we can apply Belief Revision operators; thus, our goal is to rationally determine which ordering of ordinary possible worlds we should consider when talking about context-sensitive uncertainties.

The last case is when the agent is clueless about w and about c. Consider, for instance, the Rudolf Lingens example. Even though Lingens has learned about the world, there is something he does not know: that he is Rudolf Lingens. That means that he does not know something about the world w, namely, that an individual Rudolf Lingens is at the Stanford Library. However, also he does not know something about his own location within that world, namely, that he is Rudolf Lingens. To solve a case like that, we might want to assume that a given c is certain. But not all centers will be considered plausible to be assumed. Thus, before we continue, let's define what a plausible center is.

Definition 8.1.1 (Plausible Centers)

Given a center $c = \langle t, p, i \rangle$ on which $t \in \mathcal{T}$ is an instant, $p \in \mathcal{S}$ is a spatial position and $i \in \mathcal{I}$ is an individual, we will say that a plausible center are those c' on which it is possible to measure the distance between successive instants and positions.

This definition intends to exclude some impossible possibilities and constrain our framework to relevant ones. Notice that there is no guarantee about the individuals within the center. If an individual can be placed in $\langle t, p \rangle$, then she is a plausible individual. That is intended to represent the physical possibility of individuals within a center as time passes, for instance. It is part of the dynamics of *contextual changes*. In t_1 the individual i might be in position p_1, but in t_2 an individual i' might be in p_1, for instance. Moreover, no matter what kind of uncertainty one might have, we need to start applying specific strategies.

- i) If doubting $\langle w, c \rangle$, we should check if we are uncertain about w, about c, or about w and c.
- ii) If doubting w, we can fix centers and check if contextual information will not interfere with ordering possible worlds.

- iii) If doubting c, we need to select the plausible centers and think about their *intended plausible ordering*.

This all means that we need an extra step before applying AGM when encountering indexical information. The standard treatment, as we saw, cannot simply be applied to centered worlds. Because the agent can hold different kinds of uncertainty, the strategy we choose will depend on the type of uncertainty the agent holds. The interesting cases, however, are those involving doubting contextual information. Doubting c can often happen when we lose track of contextual changes when we have intended contexts or in conversational situations requiring recentering. We will say that uncertainty about c is *indexical uncertainty*. Because we are already in a context-sensitive framework, we must ensure we are revising over a consistent belief set when conflicting information appears.

8.2 Centers as Conditionals

In the context of Belief Revision, *centers* represent the standpoint or perspective from which an agent evaluates possible worlds. A *center* is a pair consisting of a possible world and an agent's position within that world, typically accounting for indexical information such as "I," "here," and "now." These centers are useful in capturing how beliefs are updated when new, context-sensitive information is received. We can interpret *centers* in a manner analogous to *conditionals*. Specifically, each center can be seen as representing a conditional stance about the world. For example, a center c might encapsulate a conditional belief of the form "If the world were as described by c, then such-and-such would be the case."

Formally, let $c = \langle w, i \rangle$ represent a centered world, where w is a possible world and i is the agent's position within that world. The agent, in adopting center c, evaluates propositions based on the ordering of possible worlds $\geq_c$, which represents how plausible or accessible other worlds are from the perspective of c. Now, we can define the conditional relationship:

$$c \vdash (A > B) \quad \text{iff} \quad \text{“If } A \text{ holds at center } c, \text{ then } B \text{ must also hold.”}$$

The conditional nature of centers becomes more evident when we consider *contextual shifts*. Each center provides a context within which an agent evaluates the truth of propositions, and Belief Revision operates by shifting between different centers depending on new information. Suppose we have a conditional of the form $A > B$, meaning "if A were true, then B would be true.". To evaluate this conditional, we can conceptualize the process as choosing a center c_A where A

holds. The plausibility or truth of B is then evaluated within the context of this center. If B follows from the center c_A, we accept the conditional. The center c_A corresponds to the conditional world where A holds, and we assess whether B follows given that center. If a contradiction arises (i.e., B does not follow), we reject c_A as a plausible center and shift to another center that better fits the new information. Thus, the evaluation of conditionals can be seen as a process of moving between centers, where each center c serves as a conditional context. We accept or reject conditionals based on whether the consequent follows in the new centered world. In formal terms, we represent this process using conditionals and revision operators:

- **Center as Conditional**: A center c can be viewed as representing a conditional belief $A > B$, where A is true in the world associated with c, and B is evaluated in the revised belief set that emerges after accepting A.

- **Belief Revision at Centers**: The Belief Revision process, denoted $\star$, involves updating the belief set according to the hierarchy $\geq_c$ of ordinary possible worlds associated with the center. If A holds at c, we revise our belief set B to B^A, and check whether B follows:

$$B \star A = B_c^A \quad \text{and} \quad B \in B_c^A.$$

- **Revising Centers as Conditionals**: When faced with new information that contradicts the current center, we reject the current center as a plausible conditional and move to another center. This can be seen as revising the conditional $A > B$ by stepping back and considering a different center c', i.e., a different context where A holds.

Using centers as conditionals allows us to model *context-sensitive Belief Revision* more dynamically. Instead of evaluating a conditional strictly in terms of a single world, we evaluate it across *centered worlds*, where each center reflects a different perspective, indexed by the agent's position. For example, if we are considering a conditional like "If I were in Berlin, it would be raining," the center representing "I am in Berlin" would be different from the actual center. Shifting to this new center involves updating beliefs to reflect the conditional scenario, after which the truth of the consequent (whether it is raining) is evaluated. If this revised context leads to a contradiction (e.g., Berlin is known not to have rain today), the center would be rejected, and another plausible center would be sought.

Viewing centers as conditionals allows us to treat each center as representing a conditional context where specific beliefs are evaluated. This perspective aligns with Belief Revision theory by conceptualizing belief change as a process of shifting between context-sensitive conditionals, where new information prompts the selection of different centers. These centers are ordered by plausibility, and the conditional reasoning process iteratively refines the agent's belief state based on their acceptance or rejection of various centers.

8.3 Rational Belief Revision with Centered Worlds

In the centered possible worlds framework, each possible world w is not just a set of propositions, but a pair $\langle c, w \rangle$, where c is the center of the world. This center typically consists of the perspective-sensitive features of the agent, as defined in K-contexts. The agent's belief is evaluated not just with respect to which world is actual, but from the specific perspective they occupy in that world, that is, their center. The problem emerges when the agent has to revise their beliefs in light of new information that directly affects the center (*e.g.*, the agent's belief about where they were, or what day it is). This is distinct from standard Belief Revision because changes in context-sensitive beliefs affect the center itself, not just the content of the world.

To consider how this would play in a Belief Revision process, we will first need to represent an initial state of beliefs. Initially, the agent holds a set of centered possible worlds representing their beliefs. Let's call this set B, where each center $c \in B$ is a centered world compatible with the agent's current beliefs. Then, new information comes in the form of a proposition that might involve indexicals. The next task is to revise the agent's belief set B to accommodate this new piece of information. Unlike classical Belief Revision, where you'd simply intersect the agent's beliefs with the new information, here you need to revise the center itself. The challenge is that indexicals somewhat tie the belief to the perspective of the agent, so revising the belief must also adjust the centered perspective. It is necessary, thus, to expand Standard AGM Belief Revision to centered possible worlds, taking into account both the propositional content and the shifting centers. Moreover, given the context-sensitivity of indexicals, the revision process should respect certain constraints. Considering the standard Belief Revision operators, those constraints can be defined as follows:

Success: The revised belief set B' must include the new information (i.e., all centered worlds in B' are consistent with the new evidence E).

Minimal Change: The revision should make minimal changes to the agent's prior belief set, particularly when it comes to retaining centers that aren't directly contradicted by the new information.

Temporal and Spatial Indexical Consistency: When revising beliefs involving indexicals like “today” or “here”, the process must consider their shifting features, as these terms can change their reference depending on the time or location. For example, if the agent updates their belief about where they were yesterday, the indexical “ yesterday” must correctly update in the revised set of centered worlds and consider the ordering for temporal and spatial information.

A possible way of expanding Stardard AGM Belief Revision to Centered Worlds is to consider the notion of *Distance* from the probabilistic approaches [96]. When changing from one centered world to another, the key challenge is updating the center while preserving as much of the agent’s prior knowledge as possible. Using distance metrics over centered worlds, we will say that the *distance between two centered worlds* is a measure of how much the agent’s perspective differs across worlds. A Belief Revision process can then aim to minimize this distance when selecting the new set of centered worlds B'. For example, if the messy shopper revises their belief from “Someone in the supermarket is making a mess” to “I am in the supermarket making a mess” the minimal revision might involve only changing the individual component of the center while preserving other features of the centered world (e.g., the location and the time). The notion of Distance can be used to address some of the challenges in applying Standard AGM to centered worlds.

As we saw, Standard AGM doesn’t directly offer an account deal with indexicality. In the AGM framework, you would revise your belief state with a proposition like “The coin landed heads” but not with “I am the Messy Shopper” because the latter is a self-locating belief. Moreover, centered possible worlds introduce a dimension of temporal or spatial uncertainty that standard AGM revision does not typically address. Standard AGM models typically assume that Belief Revision deals with non-temporal and non-self-locating uncertainty, which can be revised away by new propositions. In centered worlds, even after revision, self-locating uncertainty might remain. Indeed, in centered possible worlds approach, revision must take place at the level of centers within worlds. This is more complex than AGM’s approach, where the revision only deals with possible worlds, not specific centers within those worlds. Moreover, we proved by Arrow’s Impossibility Theorem for Centered Worlds that we cannot reduce an ordering of centered possible worlds into an ordering of ordinary possible worlds in a simple matter.

To reconcile these issues, we might think of a *Centered AGM-like Framework*, where the agent’s belief state is a set of centered possible worlds, not just possible worlds; the revision process involves updating not just which the ordinary possible worlds, but also which centers within those worlds are most plausible; and the postulates of Minimal Change, Consistency, and Success must be adapted to handle the additional indexical content. We can, now, think about a *Rational Belief Revision in Centered Worlds*.

Recall that a *centered world* is a pair $\langle w, c\rangle$, where:w is a *possible world*; c is the *center*, representing the agent's perspective in w, defined by parameters such as time, location, and individual. Let B be the agent's current belief set, represented by a set of centered worlds compatible with the agent's current knowledge. We begin by proving that **Success** and **Minimal Change** can be applied to centered possible worlds, ensuring that the revised belief set incorporates new information while altering the center minimally.

The **Success** postulate in our new framework must state that, after revising the belief set B by some new evidence E, the resulting belief set B' must include E. Formally:

Success $E \in B'$ *for any revision* $B' = B * E$

In standard Belief Revision, this means that every possible world in the revised belief set is consistent with E. For centered worlds, we extend this postulate as follows:

Centered Success Let B be a set of centered worlds. Revising B by evidence E, the revised set $B' = B * E$ must include only centered worlds $\langle w, c\rangle$ such that $w \models E$ (i.e., E is true in world w).

Justification:

1. Let $B = \{\langle w_1, c_1\rangle, \ldots, \langle w_n, c_n\rangle\}$, the initial set of centered worlds.
2. The revision $B * E$ requires us to remove any worlds w_i where $w_i \not\models E$, i.e., those where E does not hold.
3. After removing these worlds, the remaining belief set B' consists of only those centered worlds $\langle w_i, c_i\rangle$ where $w_i \models E$.
4. Hence, B' satisfies the Success postulate. Thus, $E \in B'$ as required.

The **Minimal Change** postulate ensures that the revision process makes the smallest possible change to the agent's previous belief set. We define a distance measure between centered worlds to quantify the degree of change, denoted as $d(\langle w_1, c_1\rangle, \langle w_2, c_2\rangle)$, where d represents how much the center (agent's perspective) differs across these worlds, like in [96].

Minimal Change in Centered Worlds Let B be a set of centered worlds, and E a new piece of evidence. The revised set $B' = B * E$ must minimize the change in the agent's center, preserving as much of the original center c as possible while incorporating E.

Justification:

1. Suppose we have two belief sets, $B = \{\langle w_1, c_1\rangle, \ldots, \langle w_n, c_n\rangle\}$ and $B' = B * E = \{\langle w'_1, c'_1\rangle, \ldots, \langle w'_m, c'_m\rangle\}$.
2. Define $d(\langle w, c\rangle, \langle w', c'\rangle)$ as the distance between two centered worlds, where d measures the shift in the agent's perspective. Specifically, we consider changes in indexicals like "I", "here", "today", etc.
3. The revision $B * E$ ensures that the agent's center changes minimally, subject to E being true in all w'_i.
4. Hence, the belief set B' is the one that minimizes:

$$\sum_{\langle w_i, c_i\rangle \in B} \min_{\langle w'_j, c'_j\rangle \in B'} d(\langle w_i, c_i\rangle, \langle w'_j, c'_j\rangle)$$

   This ensures that the new centers c'_j preserve as much of the original agent's perspective as possible.

We now formalize the constraint for **Temporal and Spatial Indexical Consistency**, which ensures that revisions involving indexicals like "here", "today", or "I" respect their reference shifts. Let $f(ic)$ be a function that maps centered worlds to their indexical content.

Indexical Consistency Given a set of centered worlds B and a new indexical expression $E_{f(ic)}$, the revised belief set $B' = B * E_{f(ic)}$ must preserve the consistency of indexical terms. Specifically, if $E_{f(ic)}$ refers to "today", then B' must update the center to reflect the correct temporal shift.

Justification:

1. Let $B = \{\langle w_1, c_1\rangle, \ldots, \langle w_n, c_n\rangle\}$, where each center c_i includes parameters for time t_i and location l_i.
2. If the new evidence $E_{f(ic)}$ involves an indexical like "today", then revision must update the time parameter t_i in each center c_i.

3. Let $f(ic)(c_i)$ be the indexical content of c_i. After revising with $E_{f(ic)}$, the belief set must respect the temporal shift, such that: $f(ic)(c_i') = f(ic)(c_i)$ for all indexicals that remain unchanged.

4. The revision adjusts only the indexicals directly referenced by $E_{f(ic)}$. Hence, consistency is maintained across the revised set B'.

Thus, indexical consistency is preserved during Belief Revision.

We can now formalize how **Distance** can guide Belief Revision in centered worlds, particularly when dealing with indexicals. This approach helps us preserve the agent's prior knowledge while minimizing changes to the center and it is consistent with theories such as Grove's Spheres [37].

Distance between Centered Worlds Let $d(\langle w, c\rangle, \langle w', c'\rangle)$ be a distance function that measures the difference between centers c and c'. We define the revision operator $*$ to minimize the total distance when updating beliefs.

Justification:

1. Let $B = \{\langle w_1, c_1\rangle, \ldots, \langle w_n, c_n\rangle\}$ be the initial belief set, and E the new evidence.

2. Define the distance d as the sum of differences in the indexical parameters (time, location, individual).

3. The Belief Revision process $B * E$ selects a revised set B' that minimizes:

$$\sum_{\langle w_i, c_i\rangle \in B} \min_{\langle w'_j, c'_j\rangle \in B'} d(\langle w_i, c_i\rangle, \langle w'_j, c'_j\rangle)$$

4. This ensures minimal deviation from the agent's previous centered worlds, satisfying the principle of Minimal Change.

Thus, Belief Revision minimizes the shift in the agent's perspective while incorporating the new evidence.

Extending Standard AGM Belief Revision for Centered Information might look straightforward for the cases on which we consider the whole centered proposition as uncertain. However, as we discussed so far, to capture the distinctions between other kinds of uncertainty that might arise from indexical belief change. Thus, we might turn our attention towards a Two Dimensional approach to AGM Belief Revision.

8.4 Two-Dimensional Belief Revision

Belief Revision is a fundamental topic in epistemology, logic, and artificial intelligence, as it provides a formal framework for modeling how rational agents update their beliefs in light of new information. Traditional models of Belief Revision, such as the AGM (Alchourrón, Gärdenfors, and Makinson) framework [2], assume a static perspective where belief updates occur within a fixed world structure. These classical models primarily focus on non-indexical beliefs—propositional content that remains the same regardless of the agent's perspective. However, they do not fully capture the complexities involved when dealing with self-locating beliefs or context-sensitive propositions. In such cases, a more nuanced approach is required, one that accounts for both the *world-relative* and *center-relative* dimensions of belief.

Two-Dimensional Semantics, originally developed by Kaplan [46] to analyze the behavior of indexicals and demonstratives, was later expanded by Chalmers [21, 23] and Jackson [42] to distinguish between epistemic and metaphysical necessity. The core idea behind 2D semantics is that the meaning of a sentence is determined not only by its reference in a given possible world (the secondary intension) but also by its epistemic characterization across counterfactual scenarios (the primary intension). This distinction is particularly useful for understanding self-locating beliefs , where belief content is not merely about the world but also about the agent's perspectival center —i.e., their location, identity, and temporal standpoint. The key insight of 2D semantics is that the meaning of an expression or proposition is not solely determined by the possible world in which it is evaluated, but also by the specific context from which it is uttered or believed. This distinction is crucial when dealing with self-locating beliefs, where the content of a belief is not merely about the state of the world but also about the agent's *epistemic center*—i.e., their location, identity, and time. Incorporating Two-Dimensional approach to Belief Revision allows us to distinguish between two types of belief change:

1. **World-sensitive Revision**: This type of revision involves updates based on new information about the actual world. In traditional Belief Revision, this corresponds to situations where an agent revises their beliefs to accommodate factual information that modifies their understanding of the external environment. The revision operates over the standard domain of propositional beliefs, where contradictions are resolved by removing conflicting statements to maintain consistency.

2. **Center-sensitive Revision**: Unlike world-sensitive revision, center-sensitive revision occurs when new information affects the agent's understanding of

their *epistemic center*. This includes changes in self-locating beliefs, such as learning about one's location in space and time or updating beliefs about one's identity. Unlike standard updates that alter what the agent believes about the world, center-sensitive revisions modify the agent's perspective *within* the world. This is particularly relevant in cases involving indexicals such as "now," "here," and "I," where a shift in context necessitates an update to beliefs without necessarily changing the underlying facts about the world.

8.4.1 The Role of Centered Beliefs in Two-Dimensional Revision

The central challenge in Two-Dimensional Belief Revision is accommodating changes to both the agent's beliefs about the world and their beliefs about their self-location within it. Traditional Belief Revision approaches treat belief states as sets of propositions that can be revised by the introduction of new information. However, when beliefs are *centered*, they depend not only on the world but also on the agent's epistemic perspective. This creates unique challenges for revision, particularly in cases where information updates are *context-sensitive*.

Consider the *Sleeping Beauty problem* [27], where an agent undergoes memory erasure and must update their beliefs based on newly acquired contextual information. Beauty's credence in a given proposition must account for both the information she receives about the world and the shifting context in which she finds herself. If she awakens on Monday and later finds out that it is Tuesday, the update does not merely involve learning a new fact about the world—it involves *re-centering* her beliefs to accommodate the change in her epistemic standpoint.

Similarly, in the *Alone in the Wilderness* scenario, an agent who initially lacks knowledge of their precise location must revise their beliefs upon discovering new information. The revision does not involve a change in the external world but rather a modification of the agent's understanding of their own place within it. In such cases, a purely world-sensitive model of Belief Revision is insufficient; what is needed is a Two-Dimensional framework that accounts for shifts in both world-relative and center-relative knowledge.

8.4.2 Formalizing Two-Dimensional Belief Revision

A formalized Two-Dimensional Belief Revision framework must explicitly distinguish between *world-sensitive* and *center-sensitive* updates, ensuring that belief changes respect both external factual modifications and shifts in self-locating knowledge. One possible approach is to extend AGM-style postulates by introducing two revision operators: one for updating beliefs about the world while keeping the agent's self-location fixed, and another for updating self-locating information while preserving consistency with previously accepted world-relative beliefs. Another direction involves integrating *probabilistic models* that account for belief uncertainty at both levels, allowing for weighted belief adjustments based on new information. Modal logic extensions, particularly those that build on *Two-Dimensional Modal Semantics*, could provide a structured way to model how belief states evolve along both axes, ensuring that revision processes reflect both objective and subjective dimensions of information change.

At first glance, Rott‘s [91] Bounded Revision proposal aims to offer a constraint-based approach that balances conservative and moderate belief updates, making it a potential candidate for a Two-Dimensional Belief Revision framework. Its emphasis on structured, non-arbitrary modifications aligns well with the idea that self-locating beliefs should not be revised in the same manner as world-relative beliefs. However, bounded revision as originally formulated does not inherently distinguish between world-sensitive and center-sensitive belief changes; it primarily governs the degree of change allowed rather than categorizing belief updates based on their contextual nature. Therefore, while bounded revision provides a useful constraint mechanism, it requires additional formalization to explicitly handle the Two-Dimensional structure necessary for context-sensitive Belief Revision.

Bounded Revision: A Two-Dimensional Model of Belief Change

Hans Rott's framework of *bounded revision* [91] offers a refined model of belief change that introduces an additional parameter to traditional Belief Revision methods. Unlike classical AGM-style revision, which assumes a single-dimensional modification of the belief state, bounded revision incorporates a second constraint that governs the extent to which new information is integrated. This model is particularly useful in addressing the dynamics of belief change when beliefs interact across different epistemic commitments. The approach extends beyond traditional methods by introducing a reference point that restricts the revision process, preventing radical shifts in the belief system while ensuring epistemic coherence.

Rott emphasizes that the standard AGM paradigm, though widely accepted, fails to account for varying degrees of entrenchment among beliefs. In real-world reasoning, not all beliefs are revised with equal flexibility; some are more resilient due to their foundational role in the agent's epistemic system. Bounded revision acknowledges this distinction by incorporating a *reference sentence* that determines the level of epistemic commitment associated with a belief update. As Rott states, "Belief Revision should be constrained in such a way that the new information is accepted no more strongly than the reference belief itself" [91, p. 12]. This constraint ensures that new beliefs do not override established ones indiscriminately but rather integrate within the existing belief framework in a controlled manner.

In standard Belief Revision, an agent revises their belief set $\mathcal{B}$ upon receiving new information φ by incorporating φ in a way that ensures consistency and minimal change. The AGM model [2] assumes that Belief Revision occurs under uniform principles, treating all beliefs as equally susceptible to change. However, this assumption does not reflect the nuanced nature of epistemic commitments, particularly when belief updates involve self-locating information or indexical expressions.

Bounded revision introduces a Two-Dimensional modification process, where belief updates are assessed relative to a reference sentence ψ. The function that governs bounded revision thus operates as follows:

- φ: the *input sentence*, representing the new belief to be incorporated;
- ψ: the *reference sentence*, which dictates the degree to which φ is integrated into the belief system.

This reference-based structure prevents revisions from occurring in a manner that disrupts core epistemic commitments. The degree of entrenchment associated with ψ determines how significantly φ modifies the belief state. When ψ is strongly entrenched, φ is integrated more conservatively, whereas when ψ is weakly entrenched, φ may be more readily adopted.

Given a belief set $\mathcal{B}$, the bounded revision of $\mathcal{B}$ by φ relative to ψ is denoted as $\mathcal{B} *_{\psi} \varphi$. This revision function is formally defined as follows:

Definition 8.4.1 ([91]) *Let $\mathcal{B}$ be a belief set, and let φ and ψ be sentences. The bounded revision $\mathcal{B} *_{\psi} \varphi$ satisfies:*

1. *If $\mathcal{B} \vdash \psi$, then $\mathcal{B} *_{\psi} \varphi = \mathcal{B} * \varphi$ (classical AGM revision applies).*
2. *If $\mathcal{B} \nvdash \psi$, then $\mathcal{B} *_{\psi} \varphi$ ensures that φ is no more entrenched than ψ.*
3. *If $\mathcal{B} \vdash \neg\varphi$, then ψ imposes constraints on how $\mathcal{B}$ can accept φ while preserving consistency.*

This approach ensures that when a belief system fully accepts ψ, then the revision behaves as a standard AGM update. However, when ψ is uncertain, φ is incorporated in a constrained manner, maintaining epistemic stability. Rott's model of *bounded revision* [91] provides an intermediate strategy between two well-established Belief Revision paradigms: *conservative revision* and *moderate revision*. Conservative revision, as discussed by Boutilier [15, p. 10], prioritizes the preservation of the original belief state, integrating new information only when it does not disrupt entrenched commitments. Moderate revision, on the other hand, adopts a lexicographic ordering approach, where new beliefs can displace prior ones even at the cost of altering deeply held commitments [71, p. 149]. Bounded revision mediates between these two extremes by ensuring that belief integration is always relative to a reference belief rather than applying unrestricted updates. As Rott [91, p. 24] emphasizes, “An appropriate balance between conservative and moderate belief change allows for the integration of new information without destabilizing an agent’s epistemic framework.” This constraint-based mechanism prevents radical shifts in belief structures, maintaining coherence while allowing for updates based on epistemic priorities.

One of the critical challenges in Belief Revision theory is accounting for *iterated revision*, where an agent must update beliefs multiple times while preserving logical consistency. The Darwiche-Pearl postulates [25, p. 6] formalize constraints on successive belief updates to ensure rational coherence. Bounded revision aligns with these principles by preventing abrupt shifts in epistemic commitment, thus preserving stability across iterative updates. However, a more pressing issue arises when belief changes involve *context-sensitive updates*. Traditional Belief Revision frameworks largely assume that belief states exist independently of the agent’s position within a world. Yet, as we have seen in cases involving indexicals and self-locating beliefs, the context of an agent significantly influences how beliefs should be revised. The Sleeping Beauty problem provides a clear example: when Beauty receives new world-relative information (such as the result of a coin flip), she engages in *world-sensitive revision*. However, when she updates her belief about what day it is, she undergoes *center-sensitive revision*, which directly modifies her self-locating knowledge. Standard Belief Revision models fail to distinguish these two dimensions, requiring a more refined approach.

Bounded revision offers a partial solution to this challenge by introducing a constraint-based mechanism for belief updates, but it remains insufficient in handling Two-Dimensional belief changes. While it accounts for the preservation of entrenched beliefs, it does not explicitly formalize the distinction between world-sensitive and center-sensitive updates. A comprehensive model must include additional structural principles that ensure self-locating beliefs are revised appropriately without collapsing into standard uncentered updates. Thus, we propose extending bounded revision into a formalized *Two-Dimensional Belief Revision* framework that accounts for both dimensions of belief change.

Beyond individual belief updates, bounded revision has significant implications for collective decision-making in uncertain contexts. In social choice theory, aggregating individual beliefs and preferences poses a fundamental challenge, particularly when contextual factors influence preference formation. Arrow’s impossibility theorem [7, p. 33] highlights the inherent difficulties in constructing a rational aggregation mechanism that maintains consistency while respecting individual differences in preference structures. When dealing with context-sensitive beliefs, the issue becomes even more complex. If different agents within a decision-making body hold partially overlapping but distinct beliefs, bounded revision offers a constraint-based mechanism for integrating these perspectives. However, without a Two-Dimensional framework, there is no clear method for distinguishing how belief aggregation should account for the agent's self-location in a given decision scenario. Consider a voting system where preferences depend not only on objective world conditions but also on the agent's position in a particular social context. Standard models of social choice struggle with incorporating self-locating perspectives, often treating them as anomalies rather than essential components of rational decision-making. Thus, a Two-Dimensional Belief Revision model is necessary to refine bounded revision and extend it to collective decision-making. This would allow us to formalize how different agents revise beliefs under constraints imposed by both external world factors and their individual perspectives. In the next chapter, we introduce a framework for *Two-Dimensional Belief Revision*, which extends bounded revision to handle both world-sensitive and center-sensitive belief changes. This model ensures that belief updates are not only logically coherent but also contextually appropriate, providing a more refined toolset for reasoning under uncertainty in both individual and collective epistemic contexts.

8.4.3 A Two-Dimensional Framework for Belief Revision

Traditional approaches to Belief Revision, such as AGM theory [2], focus on updating an agent's beliefs about the world while assuming a fixed epistemic perspective. However, as we have seen throughout this thesis, self-locating beliefs introduce a crucial additional dimension: the agent's position within the world, which includes their temporal, spatial, and epistemic standpoint. This necessitates a *Two-Dimensional Belief Revision* framework, one that allows for the simultaneous revision of world-relative information and self-locating (center-relative) information.

An evident motivation for this approach can be found in classical examples of self-locating uncertainty. Consider the *Sleeping Beauty Problem*: upon waking, Beauty must revise her beliefs about both (i) what the actual world is (has the coin landed heads or tails?) and (ii) where she is within that world (is it Monday or Tuesday?). A traditional Belief Revision operator would allow Beauty to update her beliefs about the coin flip but would not account for her need to revise her self-locating information. Similarly, in the *Alone in the Wilderness* scenario, an agent waking up in an unfamiliar location must revise both their factual understanding of their environment and their self-locating beliefs concerning their own position in that world. These examples illustrate why a one-dimensional approach to Belief Revision is insufficient: belief updates must not only accommodate changes in external reality but also shifts in the agent's epistemic center. To formally capture this distinction, we introduce a Two-Dimensional Belief Revision function, which explicitly differentiates between *world-sensitive* and *center-sensitive* revisions. A Two-Dimensional Revision Function $*$ can thus be expressed as follows:

Definition 8.4.2 *Let $\mathcal{B}$ be a belief state represented by a pair $\langle W, C \rangle$, where W is a set of possible worlds and C is a set of centers, each representing a particular self-location within a world. Given new information ϕ, the Two-Dimensional Belief Revision function $*$ is defined as:*

$$\mathcal{B} * \phi = \langle W', C' \rangle$$

where:

- *W' is the revised set of worlds obtained by standard AGM Belief Revision on W, reflecting changes in the agent's factual knowledge;*
- *C' is the updated set of centers, modified to reflect shifts in the agent's self-location.*

This definition ensures that Belief Revision accommodates both external and internal updates. When an agent receives new information, they must not only determine whether it alters their understanding of the actual world but also whether it affects their position within that world. In standard cases of Belief Revision, such as acquiring new empirical knowledge (e.g., "It is raining"), only W is updated. However, in cases where the agent learns self-locating information (e.g., "I am in Berlin" or "It is Monday"), C must also be revised accordingly.

To illustrate this framework, consider a modified version of the *Lingens Example*. Suppose Lingens wakes up in a library, unaware of his identity or location. Over time, he gathers new information: first, he learns that he is in a library (W is revised), and later, he discovers that he is Lingens himself (C is revised). These updates occur in distinct dimensions: the first is a world-sensitive revision, adjusting his knowledge about external reality, while the second is a center-sensitive revision, shifting his self-locating beliefs.

Our proposal extends traditional Belief Revision theories by providing a structured method for handling context-sensitive epistemic updates. By incorporating both world-sensitive and center-sensitive revision, this Two-Dimensional Belief Revision function offers a more comprehensive account of belief change in contexts where self-location matters. In the following sections, we will explore additional postulates governing the interaction between W and C, and we will examine how our model can be extended to iterated revision and probabilistic settings. Two-Dimensional Belief Revision provides a robust framework for addressing the limitations of traditional Belief Revision approaches, particularly when dealing with context-sensitive information. By distinguishing between world-sensitive and center-sensitive revisions, this approach allows for a more comprehensive understanding of how beliefs change in response to new information. The ability to track epistemic centers is essential for understanding cases involving indexicals, self-locating beliefs, and dynamic belief updates.

Consider a scenario in which an agent is uncertain about the location of a solar eclipse, and the agent's belief is tied to their understanding of "today," which depends on their current context. In traditional Belief Revision, this situation would involve revising beliefs about the actual world based on new information about the eclipse. In a Two-Dimensional model, however, we must account for both the content of the belief (the actual state of the world) and the context (the agent's current time, location, and perspective). Let us formalize those necessary steps using Two-Dimensional Belief Revision as follows:

Step 1: Contextualize the Belief The agent holds a belief about the world, but this belief is tied to a specific context, denoted as a centered world $\langle w, c \rangle$, where w is the world and c is the center. For example, an agent may believe that the eclipse happens on March 5th, but this belief is contextualized by their understanding of which day it is (i.e., their center).

Step 2: Update the Belief When new information is introduced, it can affect either the world w or the center c. A revision in the world (e.g., the eclipse occurring on a different date) will lead to a change in belief about the world itself. However, a revision in the center (e.g., realizing that the agent was mistaken about the date) will lead to a shift in the agent's perspective, requiring an update in their belief about the relevant center.

Step 3: Adjust the Belief System In Two-Dimensional Belief Revision, the agent must adjust both the content and the context of their belief. This process involves selecting which possible worlds are relevant under the new context, and how the revised belief fits into the global ordering of worlds and centers.

This framework enriches the traditional AGM Belief Revision model by integrating context-sensitive beliefs with world-sensitive ones. In AGM Belief Revision, the focus is primarily on updating a belief set in the face of new, non-contradictory information. Two-Dimensional revision, however, necessitates a dual update: both the content of the belief (the world) and the agent's position (the center) must be updated. This leads to a more nuanced form of Belief Revision, where the agent must consider not only what is true in the world, but also how the world relates to their specific perspective within it. The Two-Dimensional approach allows for more flexible and accurate modeling of Belief Revision when dealing with context-sensitive claims. By recognizing the role of the center in shaping beliefs, we can better understand how agents update their beliefs about the world, especially when their position or perspective shifts. The Two-Dimensional framework for Belief Revision provides a more comprehensive model for understanding how agents revise their beliefs in the face of uncertainties, particularly when dealing with context-sensitive information. By incorporating both world-sensitive and center-sensitive revisions, we can more accurately capture the complexities of self-locating beliefs and contextual changes. This approach builds on traditional AGM revision, extending it to accommodate the nuances of indexicality and the agent's shifting perspective within the world. As we continue to explore these foundations, we can refine how Belief Revision operates in more complex, context-sensitive scenarios.

Moreover, in the centered possible worlds approach, an agent has beliefs not just about which possible world is actual but also about their place (center) in the actual world. Two-Dimensional Semantics provides a natural framework for modeling this. The *primary intension* represents how the agent's beliefs vary across different centers within a given world, while the *secondary intension* represents how their beliefs vary across possible worlds. This allows us to systematically integrate the revision of both types of beliefs: beliefs about the world and beliefs about the self's location in the world. This can be used to develop a *Two-Dimensional Belief Revision*. In standard AGM, revision $B * \alpha$ modifies the belief state B' minimally to include the proposition α. In our Two-Dimensional setting, however, α could express a statement about the agent's location within the world (primary intension) or a statement about the world (secondary intension). Those distinctions are based on the definitions for *K-Contexts* and *S-Contexts*. In our framework we will say that *K-contexts* (Kaplanean contexts) inspire the notion of *K-intensions* (primary intensions) as *S-contexts* (Stalnakerian contexts) inspire the notion of *S-intensions* (secondary intensions) as follows:

K-Intension In K-contexts, the center is fixed by the speaker's position, including indexicals like "I". This reflects an agent's beliefs about their own specific location in the actual world. The *K-intension* (or Primary Intension), in that sense, is the function that captures how these context-sensitive beliefs vary across different centers in the same possible world. It represents how the agent's beliefs would change depending on different values of these indexicals (e.g., different times or places within the same world). K-contexts, thus, give us the framework for modeling Belief Revision in terms of primary intensions, where the focus is on updating or revising beliefs about the *centers* within a single world.

S-Intension In S-contexts, the emphasis is on the set of possible worlds that constitute the common ground or conversational background. Here, Belief Revision involves how the agent's beliefs about the *world* itself change in light of new information. The *S-intension* (or Secondary Intension) represents how an agent's beliefs vary across different possible worlds. It deals with more general, non-indexical propositions that hold regardless of the agent's specific location or perspective. In S-contexts, Belief Revision is concerned with updating these non-indexical beliefs across different possible worlds, in line with the dynamics of discourse and common ground.

By linking K-contexts to K-intensions, we capture Belief Revision that focuses on the agent's indexical, self-locating beliefs. Meanwhile, connecting S-contexts to S-intensions allows us to model Belief Revision concerning the broader set of possible worlds and general propositions. This Two-Dimensional distinction gives us the tools to handle both kinds of belief updates systematically. We can now define AGM-style postulates in a Two-Dimensional framework as follows:

S-Success (for Secondary Intension):

$$\langle w, c\rangle \in B * \alpha \Longrightarrow w \models \alpha_s.$$

This ensures that the revised belief state only contains worlds where the proposition α is true in the S-Intension sense.

K-Success (for Primary Intension):

$$\langle w, c\rangle \in B * \alpha \Longrightarrow c \models \alpha_k.$$

This ensures that the revised belief state only contains centers where the proposition α is true in the K-Intension sense.

Consistency: If $B \cup \alpha_s$ is consistent, then $B * \alpha$ is consistent.

S-Minimal Change: If possible, retain as many worlds w from the original belief state as are consistent with α. That is, revise the worlds w in B' minimally.

K-Minimal Change: If possible, retain as many centers c from the original belief state as are consistent with α.

Let us apply this to some of our examples. Let us start with the Messy Shopper Example. Recall that in Perry's original Messy Shopper Example [73], an agent (the shopper) has different beliefs about where they are shopping (the context) and the available products in the store. Their beliefs may change based on new information about what they see in front of them (the specific products) or their position in the store (self-locating beliefs).

Let B represent the original belief state. Let α represent a proposition about the shopper's location or the products. Let w be a possible world. Let c be a center (the agent's specific location in the actual world). Formally:

Belief State: $B = \{(w_1, c_1), (w_2, c_2), \ldots, (w_n, c_n)\}$

We will assume that the gained information is a proposition α such that α_k is a centered proposition about the shopper's location (K-intension); and α_s is a ordinary or centered proposition about the store itself (S-intension). In this sense, the K-contexts is given by the center and reflects the shopper's current position in the store; while the S-contexts is the set of possible worlds reflecting different shopping scenarios in the store. We can define the Belief Revision in the context of the Messy Shopper Example as follows:

K-Success: $\langle w, c\rangle \in B * \alpha \Longrightarrow c \models \alpha_k$

S-Success: $\langle w, c\rangle \in B * \alpha \Longrightarrow w \models \alpha_s$

Consistency: If $B \cup \alpha_s$ is consistent, then $B * \alpha$ is consistent.

S-Minimal Change: Retain as many worlds w from B as possible, consistent with α.

K-Minimal Change: Retain as many centers c from B as possible, consistent with α.

Consider Perry's Messy Shopper example [74]. Now suppose the shopper initially believes they see the trail of sugar in Aisle 7 (S-intension) but is unsure if they are in the Aisle 7 (K-intension). Thus, we have:

Initial Belief State: $B = \{\langle w_{\text{sugar}}, c_{\text{aisle 7}}\rangle, \langle w_{\text{sugar}}, c_{\text{aisle 8}}\rangle\}$

New Information: The shopper learns that the trail of sugar is coming from their cart. Thus, α: "I am making a mess."

The next step says we need to analyze the revising process according to our Two-Dimensional approach: for **S-intension**, the shopper's belief about the world remains intact as they still see the trail of sugar; for **K-intension**, the shopper's belief about their specific location is updated; they recognize their role in making the mess, indicating their proximity to the cart. The resulting Belief is: $B * \alpha = \{(w_{\text{sugar}}, c_{\text{cart}})\}$

Two-Dimensional Belief Revision framework intends to offer several key advantages over standard AGM Belief Revision. To start, separating two kinds of intensions allows us to better understand that Kaplan [46] and Stalnaker's [102] accounts for *Contexts* can coexist in the same framework. This separation enables agents to manage beliefs about their own position (*self-locating beliefs*) and beliefs about the external world independently, addressing context-sensitive expressions more effectively. As we saw, not all centered uncertainties are the same. When incorporating both K-contexts and S-contexts to the same framework, we allow agents to navigate and update their beliefs in a dynamic and context-sensitive manner. This capability reflects real-world situations more accurately, where an agent's understanding may shift based on changes in their perspective or information not only when Revising, but also when Updating their beliefs. Because of its context-sensitivity feature, the 2D-framework for Belief Revision is particularly well-suited for dealing with indexicals, which standard AGM approaches struggle with. By accommodating how beliefs change depending on different contexts and centers, this framework provides fine-grained modeling of scenarios where the meaning of propositions is context-dependent. Moreover, by introducing K-minimal and S-minimal change principles, the framework emphasizes retaining as much relevant information as possible during Belief Revision. This flexibility can lead to less radical changes in belief states, maintaining continuity and coherence in the agent's belief system. We do not need *necessarily* to give up all the information within a centered proposition: sometimes it is just a matter of updating the contextual information, such as Alice in our *Annoying Bob Example*. Overall, the Two-Dimensional Belief Revision framework enhances the standard AGM approach by providing a more comprehensive and flexible means of modeling belief updates, particularly in complex, context-sensitive situations.

8.5 When Do We Need a Two-Dimensional Approach?

Belief Revision is not always a straightforward process of updating one's knowledge about the external world; it often involves an agent reassessing their own epistemic position within that world. Consider once again the *Alone in the Wilderness* scenario: an individual lost in the wilderness faces two types of uncertainty. On the one hand, they are uncertain about external facts—what is the best path forward, how far they are from civilization, or whether there is a water source nearby. On the other hand, they are also uncertain about their own position relative to those facts—how long they have been lost, whether they have traveled in circles, and whether it is morning or afternoon. These two forms of uncertainty parallel the distinction between traditional (one-dimensional) and Two-Dimensional Belief Revision. The former deals with knowledge of the world, while the latter requires the agent to revise their own self-locating beliefs.

Much like an explorer navigating a landscape with incomplete information, an agent revising their beliefs must determine whether the uncertainty they face is purely factual or also self-locational. In many cases, a standard, one-dimensional approach to Belief Revision is sufficient. When an agent acquires new, non-indexical information—such as "It will rain tomorrow"—traditional AGM-style revision suffices to integrate this knowledge into the belief set. However, when the new information is indexical, such as "I am in Berlin" or "It is Tuesday," a single-dimensional update is inadequate. These beliefs are inherently tied to the agent's perspective, requiring an additional dimension to account for their shifting epistemic position. Two-Dimensional Belief Revision thus provides a structured framework for handling belief changes that involve both world-relative and self-locating updates.

The need for a 2D framework, however, does not arise in every case. Introducing a more complex formalism comes with computational and conceptual costs. If the context of an utterance is stable or if indexical expressions do not play a significant role in the belief change, a simpler one-dimensional model remains more efficient. The challenge, then, is determining when the complexity of Two-Dimensional Belief Revision is warranted and when a one-dimensional approach is sufficient. In this chapter, we explore this fundamental question: under what conditions does an agent need to adopt a Two-Dimensional framework for Belief Revision, and when can they rely on a simpler model? Much like a lost traveler choosing the right tools to navigate their journey, a well-chosen framework ensures that Belief Revision remains both precise and manageable. We will examine concrete examples where self-locating uncertainty necessitates a Two-Dimensional approach and contrast these with cases where a one-dimensional revision strategy is more appropriate. In doing so, we aim to establish clear criteria for when the richer structure of Two-Dimensional Belief Revision is indispensable.

Let us once more think about the Alone in the Wilderness Example. Imagine yourself lost in the wilderness, uncertain of your position in the world and uncertain of your place in time. You know where you are relative to the landscape around you, but not in terms of precise coordinates. Is it morning or afternoon? How far have you traveled? These are the kinds of questions we often face in life questions not just about the world, but also about our position within it. As we navigate through our beliefs, sometimes we face a similar uncertainty. Our beliefs aren't always about the objective facts of the world alone they're often about our relationship to those facts. This chapter will explore a similar dilemma in Belief Revision: when should we need a Two-Dimensional approach to handle the complexities of context-sensitive beliefs, and when can we rely on a simpler, one-dimensional model? When you're lost in the wilderness, the way you perceive the landscape can change depending on where you stand. Similarly, Belief Revision isn't always a simple matter of adjusting beliefs in response to new information about the world. Sometimes, it's also about adjusting our beliefs based on our position within the world our location, our time, and our perspective. This is where the concept of Two-Dimensional (2D) semantics comes in. By adopting a 2D approach, we can distinguish not only between what's true in the actual world and what's true in possible worlds, but also between what's true from the perspective of the agent (their center) and what's true in the world itself. But no always we need to make use on a very powerful tool. In many cases, simpler treatments are more efficient. In this chapter, we will address the central issue: When should we adopt a 2D framework for Belief Revision, and when is a simpler 1D model sufficient? Much like a hiker choosing which path to take in the wilderness, we must decide which model will allow us to better navigate the revision of our beliefs will it be a Two-Dimensional model that accounts for context-sensitive information, or can a simpler model suffice when the context is less relevant?

Two-Dimensional Semantics provides a framework that accounts for the way context influences the meaning of propositions. Chalmers [24] introduced this framework by distinguishing between the actual world dimension, which represents the truth of propositions in the actual world, and the possible worlds dimension, which reflects the truth in various possible worlds, given the agent's context and perspective. This distinction is especially useful when dealing with self-locating beliefs, such as “I am here" or “It is now," where both the world and the agent's position within the world must be taken into account. When we revise beliefs using a 2D approach, we must account for both the content of the world and the context of the agent. For instance, in Belief Revision with self-locating beliefs, like the wilderness example where our belief about time or space is tied to where and when we are, we need a mechanism that updates both the content (the world) and the context (the agent's position). This approach is more complicated than a one-dimensional Belief Revision because it considers not just what is true in the world but also what the agent believes about their own position.

However, not all Belief Revision scenarios require such complexity. As we move forward, we'll explore scenarios in which the world can be revised without considering the agent's position in it situations where a one-dimensional model is sufficient. For instance, when revising beliefs about objective facts, like the weather or the state of a distant event, the agent's position is irrelevant. In these cases, a simpler, one-dimensional revision is both sufficient and more efficient. As we've discussed in this thesis, Belief Revision with 2D semantics becomes essential when the agent's belief depends on both the world and their own position within it. For example, in self-locating beliefs, where the proposition is inherently tied to the agent's perspective (e.g., "I am standing in front of a tree"), the Belief Revision process requires not only updating the content of the world (e.g., the tree) but also updating the agent's context (e.g., where the agent is standing). In these cases, the 2D model allows for a more nuanced and accurate Belief Revision process. In contrast, when the belief is context-insensitive such as a belief about objective facts where the agent's perspective doesn't affect the truth value the one-dimensional model works just fine. The Belief Revision simply involves updating the belief set without considering the agent's context. In the following sections, we will explore these two approaches in more detail, comparing the contexts where a Two-Dimensional model is necessary and when a one-dimensional approach will suffice. As we review different Belief Revision scenarios, we will consider how 2D semantics helps us handle context-sensitive beliefs, while also examining situations where the additional complexity of 2D semantics is unnecessary. One of the key challenges in applying a 2D approach to Belief Revision is determining when the additional complexity is truly necessary. The issue becomes especially clear when we try to aggregate beliefs that are sensitive to context, such as self-locating beliefs. If the agent's beliefs are dependent on both the content of the world and their context (location, time, etc.), the revision process becomes more complicated. We must consider how to aggregate multiple individual-centered beliefs into a global ordering of possible worlds. This is similar to navigating a complex terrain in the wilderness, where every step you take alters your perspective. Just as your position influences how you perceive the landscape, an agent's position in time and space influences their beliefs. The challenge is determining when this added complexity is required for Belief Revision and when simpler models will suffice.

By understanding the principles of Two-Dimensional Semantics and applying them to Belief Revision, we can better understand when context-sensitive revisions are necessary. However, as we will see, there are cases where simpler, one-dimensional Belief Revision models remain adequate. Avoinding uncesesary and overcomplications is always interesting. Through the lens of self-locating beliefs and context-sensitive information, we explore the distinction between one-dimensional and Two-Dimensional models and highlight the circumstances under which each is appropriate. In the next sections, we will delve deeper into these models, examining specific scenarios that require a 2D approach and others that are adequately addressed by a 1D model. The wilderness example will continue to guide our exploration, helping us understand how context shapes our beliefs and how Belief Revision must adapt accordingly.

In some cases, the use of Two-Dimensional revision is indispensable. For example, when the belief being revised depends on the agent's self-locating information, such as beliefs about their current position in time, space, or their personal identity, a Two-Dimensional model is required. This is because the agent's beliefs are inherently tied to their own perspective, which must be incorporated into the revision process. As we will explore in this chapter, the ability to distinguish between the actual world dimension and the possible worlds dimension provides the necessary tools for revising beliefs about the agent's position and perspective. However, in other cases, a simpler one-dimensional model may suffice. When the belief being revised does not depend on the agent's context, and when the information being revised is context-insensitive, the one-dimensional approach provides a more straightforward and efficient method of Belief Revision. This is particularly true in situations where the belief concerns objective facts about the world, rather than subjective or self-locating beliefs. In the next sections, we will explore these distinctions further by considering examples where Two-Dimensional Semantics is essential for Belief Revision, and where a simpler approach would be sufficient. Through these examples, we will develop a clearer understanding of when to apply Two-Dimensional frameworks in Belief Revision and when to rely on simpler one-dimensional models.

8.5.1 Revisiting Two-Dimensional Semantics

Two-Dimensional (2D) semantics provide a framework for understanding the meaning of propositions, particularly in the context of indexical expressions and self-locating beliefs. The core idea of 2D semantics is that the meaning of a proposition can be understood along two dimensions: the primary dimension of possible worlds (or "what is actual") and a secondary dimension that reflects the speaker's or agent's perspective or context (i.e., "what could have been the case, given what I know"). This dual approach allows us to model how the content of a proposition might change depending on the context in which it is evaluated, such as the agent's point of view, location, and time. David Chalmers [24], one of the key figures in the development of Two-Dimensional Semantics, provides a clear distinction between two types of meaning. The first is the "a priori" meaning, which captures the truth of a proposition based on its intrinsic structure and logical necessity, independent of any specific context. The second is the "a posteriori" meaning, which captures the truth of a proposition in relation to the context of an agent's knowledge or belief. For Chalmers [22], Two-Dimensional Semantics is a framework for understanding propositions that distinguishes between what is true in the actual world and what is true in other possible worlds, taking into account the agent's context and perspective. In this framework, the meaning of a sentence is given by a pair of dimensions: the actual world dimension, which represents what is true in the actual world, and the possible worlds dimension, which represents what is true in various possible worlds given the agent's context and knowledge. He goes further and elaborates on the significance of the Two-Dimensional approach for understanding epistemic and modal contexts. He suggests that many philosophical puzzles, especially those involving knowledge, belief, and modality, can be addressed by distinguishing between the actual and possible worlds dimensions. This allows us to capture the way in which agents reason about possibilities and counterfactuals, while accounting for their limited perspective on the actual world. Moreover, Two-Dimensional Semantics allows us to understand how meaning varies depending on the speaker's or agent's context. The key insight is that meaning is not fixed but depends on both the actual world and the possible worlds the agent considers, which allows us to account for how agents interpret propositions based on their own perspective and knowledge. This insight into context-sensitive meaning has been pivotal in analyzing self-locating beliefs, which are beliefs about one's own position in time and space. For example, when an agent says, "I am in Berlin," the meaning of this sentence depends not only on the actual state of the world but also on the agent's knowledge and perspective at the time of utterance. In this case, the sentence is context-dependent, and its truth value is determined by both the world and the agent's position within that world. Chalmers' [22] approach has been influential, especially in the development of Two-Dimensional frameworks for Belief Revision. By adopting this perspective, we can better understand how agents update their beliefs based not only on new information but also on the shifting contexts in which that information is interpreted.

Another important contribution to the development of Two-Dimensional Semantics comes from philosophers like Kaplan [46] as we are already familiar with and Rabern [86], that builds on Chalmers' work, emphasizing how the 2D framework can be used to understand the role of self-locating beliefs in modal logic and how it addresses problems in epistemology. According to him, Two-Dimensional Semantics provides the tools necessary to bridge the gap between how we think about the actual world and how we reason about other possible worlds in terms of belief and knowledge. By incorporating both the agent's perspective and the possible worlds in which the agent is situated, 2D semantics provides a nuanced approach to meaning and Belief Revision. This framework is crucial for addressing the challenges of context-sensitive reasoning, especially when beliefs and knowledge are not fixed but change depending on the perspective and knowledge of the agent involved.

Two-Dimensional Semantics offers a robust tool for understanding context-sensitive beliefs and Belief Revision. By distinguishing between the actual and possible worlds dimensions, we can model how agents reason about possibilities and counterfactuals, incorporating both the agent's perspective and their knowledge of the world. This framework is especially valuable in epistemology and modal logic, where self-locating beliefs and context-sensitive reasoning are central concerns. Moreover, in the previous chapter, we proposed distinctions between different types of intensions specifically K-intensions and S-intensions as a central element in our framework for Belief Revision. These distinctions are not merely theoretical but provide essential tools for understanding how agents revise beliefs, especially when context-sensitive expressions like indexicals and self-locating beliefs are involved.

The introduction of these different intensions addresses a crucial gap in existing Belief Revision models, which often oversimplify the dynamics of belief change. By distinguishing between K-intensions and S-intensions, we can offer a more nuanced approach that is sensitive to both the agent's knowledge of the world and their position within that world. This distinction, we argue, is vital for ensuring that Belief Revision processes accurately capture the complexity of real-world reasoning, particularly when dealing with context-sensitive or self-locating beliefs. The introduction of these different intensions addresses a crucial gap in existing Belief Revision models, which often oversimplify the dynamics of belief change. Traditional models of Belief Revision, such as the AGM framework, focus primarily on revising beliefs about the world (K-intensions), typically in response to new information that directly impacts the content of those beliefs. However, these models tend to overlook the role of context in shaping beliefs, which is especially relevant when dealing with indexicals, self-locating expressions, or context-sensitive propositions. By distinguishing between K-intensions and S-intensions, we can offer a more nuanced approach that is sensitive to both the agent's knowledge of the world and their position within that world. K-intensions focus on the agent's understanding of facts and propositions, while S-intensions address how these facts are perceived in relation to the agent's own situation time, place, and perspective. This distinction allows for a more flexible and accurate representation of Belief Revision, particularly in scenarios where context influences what is known or believed.

Furthermore, this distinction is vital for ensuring that Belief Revision processes accurately capture the complexity of real-world reasoning, particularly when dealing with context-sensitive or self-locating beliefs. Real-life reasoning often requires the ability to revise beliefs based on not just the facts of the world but also the agent's subjective experience and situational context. For instance, consider the statement "I am here now." The truth of this statement is not solely determined by the world itself but also by the speaker's current position in space and time. When agents revise such beliefs, they need to account for changes in their own context, which influences how they interpret the facts of the world. Without distinguishing between K-intensions and S-intensions, Belief Revision models would fail to account for this dynamic aspect of reasoning. By introducing both K-intensions and S-intensions, we can create Belief Revision models that better mirror the complexities of real-world decision-making. These models not only address the agent's knowledge of facts but also incorporate the agent's changing perspectives, providing a more comprehensive framework for understanding Belief Revision in contexts where knowledge and perspective are intertwined. This distinction enables a more accurate and sophisticated approach to belief change, offering a richer understanding of how agents update their beliefs in the face of both new information and shifting contexts.

8.5.2 More on K-Intensions and S-Intensions

K-intensions are associated with the epistemic dimension of Belief Revision. They capture beliefs that an agent holds about the world based on the knowledge they have at a given time. In the context of Belief Revision, K-intensions represent the agent's subjective understanding of the world, constrained by their available information. This is important because an agent's belief may change when new information becomes available, thus modifying their understanding of the world. For example, consider a scenario in which an agent, Alice, believes that a certain event will occur based on her current knowledge. Alice's belief is represented by a K-intension, reflecting her epistemic state in the actual world. As new evidence is introduced, Alice's belief undergoes a revision. This process is essential for Belief Revision, as it allows agents to update their beliefs based on new, reliable information about the world. K-intensions are critical when dealing with uncertainty about the world because they focus on the agent's knowledge and available evidence. These distinctions allow us to formalize Belief Revision in a way that respects the limitations of what the agent knows at any given time. Without this separation, we risk oversimplifying the revision process by conflating the agent's knowledge of the world with the broader set of possibilities the world could hold. In contrast, S-intensions refer to the contextual dimension of Belief Revision. These are beliefs about the world that depend on the agent's specific context, including their position in time, space, and identity. S-intensions are essential for understanding self-locating beliefs, where an agent's perspective affects their interpretation of a proposition. These beliefs are sensitive to changes in the agent's context, which means that revising them requires taking into account not just the world, but also the specific context in which the belief is held.

Consider the example of indexicals like “I am here now." The truth of this statement depends on the agent’s position in space and time, and is thus context-sensitive. When an agent revises their belief about such a statement, the revision is not purely about the world; it involves adjusting their understanding based on the contextual information they possess at the time. Therefore, in Belief Revision, the distinction between K-intensions and S-intensions is essential for handling cases where beliefs depend not only on the world but also on the agent’s situatedness within it. S-intensions are particularly useful when dealing with indexicals, self-locating beliefs, and other context-sensitive expressions. By distinguishing S-intensions from K-intensions, we can formalize the Belief Revision process in a way that acknowledges the fluidity of context and the agent’s evolving perspective. This distinction enables a more nuanced and accurate model of Belief Revision, where changes in the agent’s context are properly accounted for.

The distinction between K-intensions and S-intensions becomes especially relevant when an agent faces uncertainty about both the world and their own position within that world. As we have discussed in previous chapters, Belief Revision is often required when an agent is confronted with new evidence that forces them to reassess their beliefs. However, when the evidence is context-sensitive, simply revising based on the world alone is not sufficient. We must account for how the agent’s perspective influences their interpretation of the world and how it affects the revision process. For example, in the Sleeping Beauty thought experiment, the agent’s Belief Revision depends not only on the world (whether the coin flip is heads or tails) but also on the agent’s position in time (whether it is the first or second time they wake up). In this case, both K-intensions (beliefs about the world) and S-intensions (beliefs about the context) must be considered in the revision process. Without the distinction between K-intensions and S-intensions, we would risk conflating the agent’s knowledge of the world with their understanding of their position within it. This could lead to imprecise or inadequate revisions of beliefs, especially in cases involving context-sensitive expressions like indexicals. By maintaining these distinctions, we ensure that Belief Revision remains both flexible and accurate, capable of handling the complexities of real-world reasoning. Thus, the distinction between K-intensions and S-intensions is indispensable for the study of Belief Revision, especially when dealing with context-sensitive beliefs. K-intensions allow us to model an agent’s Belief Revision based on their knowledge of the world, while S-intensions account for the agent’s position in the world and the role of context in shaping their beliefs. By carefully maintaining these distinctions, we can build more sophisticated models of Belief Revision that accurately reflect how agents revise their beliefs in the face of both new information and shifting contexts. This approach allows us to address a broader range of scenarios, from self-locating beliefs to the challenges posed by indexicals and other context-sensitive expressions, ultimately offering a more comprehensive framework for Belief Revision in a dynamic world.

As discussed before, the distinction between K-intensions and S-intensions is deeply connected to the issue of context-sensitivity in Belief Revision. Context-sensitive expressions, such as indexicals and self-locating beliefs, are inherently tied to the agent's perspective in time and space. As discussed earlier, a Belief Revision process that incorporates these expressions must account for not only the content of the proposition (the world) but also the context in which it is held (the agent's position). This is where the concept of Two-Dimensional Semantics (2D semantics) becomes essential. The agent's Belief Revision is not merely a matter of updating the truth values of propositions in the world; it involves a reevaluation of the agent's context and their shifting perspective. The dual-layered structure of Two-Dimensional Semantics provides a formal mechanism for dealing with both the content and context of beliefs, ensuring that Belief Revisions respect both the agent's knowledge and their position within the world. For example, in the context of the Sleeping Beauty problem, the agent's belief about whether Beauty is awake or sleeping cannot be revised solely by updating the content of the belief (whether or not the coin is heads or tails); we must also take into account the agent's perspective (whether it is the first or second time Beauty wakes up). This requires a revision framework that handles both the world and the context, which is precisely what the distinction between K-intensions and S-intensions allows us to do.

Intensions play a critical role in understanding Belief Revision, particularly when dealing with context-sensitive beliefs and indexicals. Intensions can be seen as the modes of presentation of propositions, capturing how a proposition is thought about or understood by an agent. They provide a framework for understanding the mental representations and cognitive significance that an agent attaches to a particular proposition. In this sense, intensions determine not just the content of a belief but also the way in which that content is considered and processed within an agent's belief system. For example, the belief "I am here now" carries a different cognitive load compared to a non-indexical proposition like "The sky is blue," as the former is deeply tied to the agent's personal context time, location, and self whereas the latter can be understood independent of such factors. Moreover, distinguishing between different types of intensions is essential for properly updating beliefs when new information is introduced. This distinction helps to clarify the role of context in the revision process and ensures that belief changes are accurate and consistent with the agent's subjective position in the world. When an agent revises a belief based on new information, it is not merely the content of that belief (the proposition) that matters, but also how that content is understood in relation to the agent's context. Therefore, without distinguishing between the modes of presentation of propositions, a Belief Revision model risks oversimplifying the complexity of the revision process, especially when dealing with context-sensitive or self-locating beliefs.

K-intensions, often thought of as the epistemic content of a belief, represent the way a proposition is understood or believed to be true within the context of the agent's knowledge of the world. They are concerned with the factual content of beliefs and how those facts are represented in the agent's cognitive model of the world. In contrast, S-intensions refer to the way the proposition is understood in relation to the agent's position within the world, encompassing the context-specific factors such as time, place, and individual perspective. S-intensions allow for a more nuanced understanding of Belief Revision, as they account for how changes in the agent's context influence the revision of beliefs. When new information is introduced, it often impacts not just the factual content of a belief but also the agent's perspective on that belief. In such cases, the revision process must account for both the content (K-intension) and the contextual parameters (S-intension) in which the belief is situated. Thus, recognizing and distinguishing between K-intensions and S-intensions is crucial for creating a Belief Revision framework that accurately reflects how agents revise their beliefs in the face of new, context-sensitive information. This nuanced approach ensures that Belief Revision is not only about updating facts but also about understanding how those facts are perceived and understood within a given context.

8.5.3 2D Belief Revision Framework and Applications

We have introduced the Two-Dimensional Belief Revision (2D Belief Revision) framework, which combines the Two-Dimensional Semantics with Belief Revision theory to provide a more nuanced approach to handling context-sensitive beliefs. The main goal of the 2D Belief Revision framework is to address the challenges posed by beliefs that depend not only on the state of the world (the content) but also on the agent's position in time, space, and their own perspective (the context). This framework extends traditional Belief Revision by incorporating a dual-layered model that distinguishes between world-based beliefs and context-sensitive beliefs.

The 2D Belief Revision framework builds upon the Two-Dimensional Semantics introduced by Kaplan [46], Jackson [42], and Chalmers [24], which separates the meaning of a proposition into two distinct dimensions: the actual world and the context of the speaker. The central idea is that beliefs are not just about the world but are also shaped by the agent's perspective within that world. This distinction between the content (world) and context (center) is crucial for understanding how beliefs evolve and change when the agent's perspective shifts or when new information is provided that alters their understanding of both the world and their place in it. The 2D Belief Revision framework consists of two main components:

- **Content-Based Revision:** This refers to the traditional process of Belief Revision based on changes in the world. When an agent encounters new evidence, the content of their belief is revised to reflect this new information. This part of the revision process follows standard AGM (Alchourrón, Gärdenfors, and Makinson) postulates, where beliefs are updated in a consistent manner based on the new evidence.

- **Context-Based Revision:** In addition to content-based revision, 2D Belief Revision takes into account changes in the agent's context. This could involve changes in the agent's position in time (e.g., waking up in a new day), space (e.g., moving to a different location), or perspective (e.g., gaining new knowledge or information). Context-based revision allows us to account for beliefs that are sensitive to the agent's subjective position in the world, such as self-locating beliefs or indexical expressions.

A central aspect of the 2D Belief Revision framework is the distinction between K-intensions and S-intensions, which we introduced earlier in the thesis. K-intensions refer to beliefs about the world based on the agent's knowledge and available evidence, while S-intensions refer to beliefs that are sensitive to the agent's context, including their time, space, and individual perspective. In the context of Belief Revision, K-intensions are useful when an agent revises their beliefs based on new information about the world. For example, when an agent learns that a proposition P is true in a possible world, their K-intension about P is revised to reflect that new knowledge. On the other hand, S-intensions are relevant when an agent's belief is influenced by their position in the world, such as when they update their beliefs based on the realization that they are in a different context (e.g., a different time or location). By maintaining both K-intensions and S-intensions, the 2D Belief Revision framework can model Belief Revision that accounts for both the content of the belief and the agent's perspective.

The 2D Belief Revision framework is particularly effective when dealing with context-sensitive beliefs, such as those involving indexicals and self-locating expressions. For example, consider the statement "I am here now," where the truth of the statement depends on the speaker's identity, location, and time. The 2D Belief Revision framework allows us to model this kind of Belief Revision by distinguishing between the agent's knowledge about the world (K-intension) and their knowledge about their own position in the world (S-intension). When an agent updates their belief about an indexical expression like "I am here now," the revision process involves both content-based and context-based revision. First, the agent may revise their belief about where they are (content-based revision), and then they must update their belief about the context in which the statement was made (context-based revision). The 2D Belief Revision framework allows for this dual revision process, ensuring that both the world and the agent's position are taken into account.

One of the main challenges of the 2D Belief Revision framework is the complexity of managing two dimensions of Belief Revision. While the standard AGM framework is relatively straightforward, the inclusion of context-based revision introduces additional layers of complexity. In particular, when an agent's belief is influenced by multiple possible contexts, it can be difficult to determine which context should be prioritized in the revision process.However, the benefits of the 2D Belief Revision framework outweigh these challenges. By incorporating both world-based and context-based revision, the framework provides a more accurate and nuanced model of Belief Revision that captures the complexities of context-sensitive reasoning. This is especially important when dealing with self-locating beliefs, indexicals, and other expressions that depend on the agent's perspective.

The 2D Belief Revision framework offers a powerful tool for understanding how agents revise their beliefs in the face of context-sensitive information. By distinguishing between K-intensions and S-intensions, the framework captures both the content of beliefs and the agent's position within the world. This dual-layered approach allows us to model Belief Revision in a more flexible and accurate way, addressing the challenges posed by context-sensitive expressions and self-locating beliefs. While the framework introduces added complexity, its ability to handle nuanced Belief Revision makes it an essential tool for understanding the dynamics of belief change in real-world reasoning. Moreover, it aims to provide a robust framework for updating beliefs in the presence of context-sensitive information such as in those examples discussed in this thesis. It combines elements from traditional Belief Revision with the flexibility of Two-Dimensional Modal Logic (2DML), allowing agents to revise their beliefs while considering both the content of the world and the context from which the belief is evaluated. This approach is particularly useful for handling beliefs about self-locating propositions, indexicals, and scenarios where the agent's perspective and the evaluation context play a significant role in determining the truth of a proposition. In this section, we explore how 2D Belief Revision operates, using examples to illustrate its application in real-world scenarios. By addressing cases where the revision process involves both the world and context dimensions, we can better understand the circumstances under which a 2D approach is necessary and how it enhances the process of rational belief update.

One of the most common applications of 2D Belief Revision arises in the context of indexicals and self-locating propositions. Indexicals, such as "I," "here," and "now," are expressions whose meanings depend on the context of utterance. For instance, the sentence "I am here now" depends on both the speaker's identity and their location in time and space. In traditional Belief Revision, such propositions are treated as context-independent, but this is inadequate for handling self-locating beliefs where context plays a crucial role. To illustrate how 2D Belief Revision addresses this issue, let us consider the following example:

> *Suppose John, who is lost in the wilderness, makes the statement "I am here now." At some point, he realizes she is at a fork in the woods, but she is uncertain about which path to take next.*

In this case, John's belief, "I am here now," is context-sensitive because it refers to her specific location in both time and space. When John revises her belief with new information such as learning the specific time or gaining knowledge of his exact location the context changes, which impacts the truth value of the proposition. Using 2D Belief Revision, we represent John's belief as a pair of dimensions: the actual world (the specific time and place where she makes the statement) and the possible worlds (alternative scenarios she could consider, such as paths she could take in the woods). The revision process accounts for the changes in context by updating the relevant centered proposition, which includes both the world and the context of John's belief.

$$\Box_P(\Diamond_S(I\ am\ here\ now)),$$

indicates that John's belief holds in all possible contexts (P) and is evaluated across all possible worlds (S).

Another important application of 2D Belief Revision involves cases where the agent is uncertain not only about the world but also about the context from which they are evaluating the proposition. Consider the Sleeping Beauty problem, a famous thought experiment involving self-locating beliefs. In this scenario, the agent (Beauty) is placed in a situation where she is uncertain both about the day of the week and whether or not the coin toss was heads or tails.

> *Beauty wakes up in a room every day, but she does not know whether it is the first day or a subsequent day after the coin toss. She must revise her beliefs about the day and the coin toss each time she wakes up.*

In this case, Beauty's beliefs are influenced by two types of uncertainty: 1. She is uncertain about the world (i.e., whether it is Heads or Tails). 2. She is uncertain about the context (i.e., whether today is Monday or Tuesday).

Applying 2D Belief Revision, we represent the belief as a Two-Dimensional framework where the primary context (C) represents the different possibilities for which day it could be (Monday or Tuesday), and the secondary world (W) represents the two possible outcomes of the coin toss (Heads or Tails). Each time Beauty wakes up, she must update her belief set according to both dimensions:

$$\Diamond_P(\Box_S(\textit{It is Monday or Tuesday})),$$

indicating that Beauty believes the proposition "Today is Monday or Tuesday" is true in some accessible primary contexts (P), and is necessary in all secondary worlds (S).

In this example, the revision process takes into account both the agent's uncertainty about the actual world and the contextual uncertainty, leading to a more nuanced Belief Revision process. Moreover, 2D Belief Revision also helps in scenarios where an agent is reasoning about epistemic (knowledge-based) and metaphysical (possibility-based) modalities. For example, consider the case of a scientist revising their belief about the existence of a specific particle.

> *The scientist initially believes the particle exists in all possible worlds, but after conducting an experiment, they revise their belief to reflect new evidence, concluding that the particle might only exist in some worlds.*

Here, the scientist's Belief Revision depends on both epistemic modality (the belief about the existence of the particle in the actual world) and metaphysical modality (the belief about the possible worlds in which the particle might exist). Using 2D Belief Revision, we can formalize this as:

$$\Box_P(\Diamond_S \textit{Particle exists}),$$

indicating that in all primary contexts, it is possible to find a secondary world where the particle exists. This allows the scientist to rationally revise their belief, incorporating new knowledge (epistemic modality) and adjusting their understanding of what is possible (metaphysical modality).

In the previous examples, we have seen that 2D Belief Revision allows agents to revise beliefs in the presence of context-sensitive information. Whether dealing with indexicals, self-locating beliefs, or epistemic and metaphysical modalities, 2D Belief Revision provides a formal system that accommodates the nuances of context. This dual-layered approach ensures that the agent's beliefs are updated consistently, accounting for both the actual world and the context in which the beliefs are held. The core advantage of 2D Belief Revision is that it allows agents to refine their beliefs by considering both the content of the world and their own perspective within that world. This provides a richer framework for dealing with uncertainty and complexity in Belief Revision, offering more flexibility than traditional one-dimensional models. Two-Dimensional Belief Revision offers a sophisticated way of handling context-sensitive beliefs, providing a formal structure for updating beliefs that depend on both the agent's position in the world and the worlds under consideration. Through the use of examples such as Sleeping Beauty and the scientist's particle belief, we have seen how 2D Belief Revision can be applied to various scenarios, addressing both epistemic and metaphysical uncertainty. By integrating context-sensitive reasoning into the Belief Revision process, this approach enhances our understanding of how agents revise their beliefs in complex, uncertain environments.

8.6 2D or not 2D?

One of the fundamental challenges in Belief Revision is determining when a Two-Dimensional (2D) approach is necessary and when a simpler, one-dimensional model suffices. In this framework, belief revision operates within two distinct but interrelated categories: K-intensions, which are context-dependent in the sense of Kaplan, and S-intensions, which are world-dependent in the sense of Stalnaker.

A K-intension captures how an expression's meaning depends on the agent's knowledge and the particular contextual setting in which a belief is formed. It reflects how an agent interprets a given proposition based on their epistemic state at a given time and place. An S-intension, on the other hand, concerns how a proposition is evaluated across different possible worlds. This distinction is crucial because it determines whether belief revision should be treated as a conventional, world-sensitive process (operating within a stable set of S-intensions) or whether it requires an additional dimension to account for the agent's self-location (shifting across different K-intensions).

For example, consider the Alone in the Wilderness scenario. Suppose Alice wakes up in a dense forest with no recollection of how she got there. She knows she is in a forest (a belief based on her S-intension), but she does not know where in the world the forest is or how long she has been there (uncertainties tied to her K-intension). In this situation, revising her beliefs requires distinguishing between updates that modify her understanding of the external world (e.g., discovering she is in Canada) and those that refine her self-locating perspective (e.g., realizing it is late afternoon). This example illustrates how K-intensions introduce an additional layer of complexity, necessitating a 2D approach to belief revision.

8.6.1 Step 1: Identifying the Context of the Belief

The first step in determining the appropriate belief revision process is assessing whether the belief is context-sensitive. A belief is context-sensitive if its truth value or relevance depends on external situational factors such as time, place, or the identity of the agent. This includes statements involving indexicals such as:

- *"I am here now"*—depends on the speaker's location and time of utterance.
- *"It is Monday"*—relevant only within a particular temporal frame.
- *"This is true"*—requires reference to a specific proposition in discourse.

Once a belief is identified as context-sensitive, the next step is categorizing it based on K-intensions and S-intensions:

- K-intensions (context-dependent, Kaplan-style): When a belief concerns knowledge about the world but needs to be considered under the context of the agent's epistemic state. For example, "I was born in Berlin" reflects the knowledge-based truth conditions of a belief.
- S-intensions (world-dependent, Stalnaker-style): When a belief is independent of any particular agent's epistemic state and instead concerns general truths across possible worlds. For instance, "Berlin is the capital of Germany" holds in all accessible worlds where the statement is true, regardless of who evaluates it.

Recognizing whether a belief is world-sensitive (S-intension) or self-locating (K-intension) is crucial for determining whether a 2D belief revision approach is necessary.

8.6.2 Step 2: Identifying the Type of Uncertainty

Once the context of the belief has been established, the next step is classifying the type of uncertainty involved in the revision process. Uncertainty can manifest in the following ways:

- Uncertainty about the world (S-intension sensitive): The agent is unsure about external facts but remains certain about their position within the world. For instance, "It will rain tomorrow" expresses uncertainty about a future event but does not require updating the agent's self-location.
- Uncertainty about the context (K-intension sensitive): The agent is uncertain about their temporal or spatial self-location. For example, in the Sleeping Beauty experiment, the belief "Today is Monday" requires a shift in K-intensions rather than merely updating propositional knowledge.

- Uncertainty about both the world and the context: The agent is uncertain about both external facts and their self-location. A prime example is the Alone in the Wilderness scenario, where Alice does not know her geographical location or the time of day.

If the uncertainty is purely world-related, a standard one-dimensional belief revision model is sufficient. However, if self-location is involved, a two-dimensional belief revision framework must be applied to accommodate the shifting K-intension.

8.6.3 Step 3: Applying Two-Dimensional Modal Logic to Update the Belief

Once the type of uncertainty is identified, the belief revision process operates across two dimensions:

- Primary Context (C): The agent's current epistemic state, determining how the belief holds relative to their knowledge.
- Secondary World (W): The possible world under consideration, capturing how the belief remains valid across alternative scenarios.

To formalize this, we introduce 2D modal operators:

- $\Box_P$ (necessary across primary contexts): The proposition holds in all accessible epistemic states.
- $\Diamond_P$ (possible in at least one primary context): The proposition is true in some accessible epistemic states.
- $\Box_S$ (necessary in all secondary worlds): The proposition holds in all accessible external environments.
- $\Diamond_S$ (possible in some secondary worlds): The proposition holds in at least one alternative scenario.

This modal framework ensures that beliefs are revised consistently across both knowledge-based and situation-based dimensions.

8.6.4 Step 4: Revising the Belief Set

After applying the necessary modal logic, the agent revises their belief set:

$$K \circ \mu$$

where:

- K represents the current set of beliefs.
- μ is the new proposition being incorporated.

Consider the Sleeping Beauty experiment again: If Beauty revises her belief based on new information that it is Monday, this is a S-intension update requiring a standard revision process. However, if Beauty remains uncertain whether it is Monday or Tuesday, a K-intension shift is needed, requiring a two-dimensional approach. In this case, both her knowledge of the world (the day of the week) and her position within the experimental setting must be considered.

The structured approach to identifying K-intensions and S-intensions ensures that belief revision is conducted in a way that accurately reflects both the agent's knowledge of the world and their self-locating perspective. By distinguishing between world-sensitive and context-sensitive uncertainties, we provide a formal foundation for Two-Dimensional Belief Revision, ensuring that belief updates remain epistemically coherent. This framework has broad applications, from formal epistemology to AI-driven decision-making, and highlights the necessity of refining standard Belief Revision models to accommodate context-dependent reasoning. Future work should focus on refining the modal operators for K-intension shifts and further integrating probabilistic reasoning into 2D Belief Revision.

Chapter 9

Conclusion

This thesis has explored the complex dynamics of Belief Revision, specifically in the context of context-sensitive beliefs, such as those expressed by indexicals and self-locating propositions. Traditional Belief Revision frameworks, such as the AGM model, have been extended throughout this work to account for the specific challenges that arise when beliefs are context-dependent. The central innovation of this thesis lies in the introduction of a Two-Dimensional approach to Belief Revision, which incorporates *K-contexts* (knowledge contexts) and *S-contexts* (self-locating contexts). By distinguishing between these two types of context, we have provided a more nuanced understanding of Belief Revision processes, particularly in situations where the agent's position within the world whether it be their identity, time, or spatial location plays a crucial role. The framework presented extends beyond classical Belief Revision by integrating defeasible reasoning, which allows for conclusions to be retracted or revised in light of new contradictory information. This is especially important in scenarios where beliefs are not static and where new contextual information can radically alter our understanding of the world. Using the wilderness example as a guiding model, this thesis demonstrated how beliefs can shift depending on changes in context and the knowledge available to the agent. The proposal also incorporated probabilistic reasoning and invoked Arrow's Impossibility Theorem to further refine the Belief Revision process, offering a more comprehensive view of how to handle uncertainty in Belief Revision models.

Moreover, the present work aimed to explore the complexities of rational Belief Revision when faced with context-sensitive information, particularly involving indexicals such as "I", "today", and "here". The objective was to identify how agents revise their beliefs when confronted with context-sensitive expressions, which classical systems fail to address adequately. By reviewing the literature on Indexicals, Belief Revision, and Self-locating Credences, we have developed a novel framework that integrates contextual reasoning into Belief Revision theory, addressing the gaps in traditional approaches.

At the outset, we introduced the standard logical nomenclature and relevant discussions surrounding possible worlds and belief systems. It quickly became apparent that existing frameworks, while robust in many respects, lacked the expressive power to deal with the complexity of indexical information. Notably, AGM Postulates for Belief Revision offered a powerful structure for handling changes in belief sets, particularly in the case of expansion. However, they fall short when faced with the challenges posed by the shifting nature of indexicals, which require an approach capable of capturing both *contraction* and *revision* in a context-sensitive manner. Standard Belief Revision approaches treat propositions as sets of ordinary possible worlds, which is insufficient when dealing with propositions containing indexicals. Indexical expressions, by their very nature, have contents that shift from context to context, and some occurrences of these expressions are, as Perry argued, essential and irreplaceable with non-indexical coreferential terms. Kaplan's [46] Logic of Demonstratives, while a foundational system for understanding indexicals, does not focus on the mechanisms through which agents revise their beliefs upon encountering conflicting indexical information. More recent systems, such as the *Indexical Hybrid Tense Logic* [43] and the *Logic of Indexicals* [87], although useful in some aspects, also fail to address the dynamics of rational belief change when contextual features are involved.

To address these shortcomings, we introduced two kinds of contexts that play a crucial role in Belief Revision. The first, *S-Contexts*, based on Stalnaker's [102] idea of common ground and shared beliefs, captures the social and communicative aspects of context. The second, *K-Contexts*, builds on Lewis [56] and Kaplan [46], treating contextual features as parameters of centered possible worlds. By incorporating these two kinds of contexts into our formal framework, we can better account for the shifts in belief that occur when indexical information is encountered. One key insight is that communication itself is a rich source of opportunities for Belief Revision. As agents interact, they are constantly adjusting their beliefs based on centered propositions they previously held as certain. This process of revision is defeasible, meaning it is open to reevaluation as new information becomes available. In such scenarios, agents not only revise their beliefs about how the world plausibly is but also revise their understanding of the plausible intended contexts of interpretation, as suggested by Predelli [79]. Lewis's [58] qualitative representation of orderings through *counterfactual spheres* provides a useful framework for understanding how Belief Revision operators work in these contexts.

We extended this ordering system to include not just ordinary propositions but also possible contexts. By adopting a probabilistic approach to weighted contexts, we can account for the uncertainty inherent in context-sensitive propositions. In doing so, we showed that it is possible to construct a plausible ordering of propositions and contexts, which facilitates rational Belief Revision. This approach, inspired by the concept of *social choice theory*, allowed us to borrow from Arrow's [8] conditions and the impossibility theorem to construct an ordering of possible worlds that respects the rationality criteria outlined by the AGM Postulates.

Centered assertions and self-locating beliefs present a unique challenge, particularly in terms of communication and Belief Revision. To better represent these dynamic, context-dependent scenarios, we proposed a multicentering approach, which could provide a way to model conversations and other interactions that shift within specific contexts. Though this suggestion remains underdeveloped, it has the potential to help us determine which centers of context are plausible and which are not. While several probabilistic approaches to self-locating uncertainties have been proposed, they tend to focus primarily on temporal context sensitivity, often neglecting the challenges posed by demonstratives like "that', which require extralogical information.

Titelbaum's [106] work on self-locating credences and Perry's [73] insights on the irreplaceability of indexicals contributed to our understanding of why some indexical expressions cannot be substituted with non-indexical equivalents. This thesis expands on these insights by developing a formal framework that allows for rational Belief Revision in cases where indexicality is essential. The implications of this work extend beyond formal epistemology and Belief Revision theory. Our proposed framework has significant applications in Artificial Intelligence (AI) and Decision-Making Theories. In AI, systems tasked with modeling agent-based reasoning in dynamic, context-sensitive environments stand to benefit from our approach. Autonomous agents or personal assistant AI systems, which must constantly adjust their "beliefs" or knowledge bases in response to new contextual information, could employ this framework to handle indexical shifts effectively. By doing so, AI systems can achieve greater adaptivity and context awareness, which is especially relevant in fields such as robotics, natural language processing, and real-time decision-making. Moreover, decision-making theories particularly those involving temporal shifts or shifting preferences can be strengthened by incorporating indexical reasoning. The ability to model context-dependent choices is crucial in economics, behavioral science, and public policy, where agents must make decisions based on incomplete or evolving contextual information. The framework introduced here could serve as the foundation for more robust decision models that account for dynamic contexts, offering greater precision in both individual and collective decision-making.

Further research may involve extending this framework to handle more complex forms of indexicality, such as de se attitudes or nested indexicals. Such extensions would enhance the expressivity of the system, making it applicable to an even wider array of practical and philosophical problems. In conclusion, this thesis has introduced a new formal framework for handling rational Belief Revision in the context of indexical expressions. By integrating context-sensitive reasoning into Belief Revision theory, we have developed a system that accounts for the dynamic nature of real-world reasoning. The implications of this work reach beyond the theoretical, with potential applications in AI, decision theory, and other fields where Belief Revision and context sensitivity play a crucial role. The contributions of this thesis represent a significant step toward understanding how rational agents navigate the complexities of indexicality, and the future directions of this research open exciting possibilities for modeling the interaction between belief and context in both human and artificial systems.

9.1 Further Considerations

The use of the Two-Dimensional (2D) Belief Revision framework has proven especially useful in addressing complex uncertainties, such as those faced by individuals navigating real-world situations involving context-sensitive information. For instance, consider the case of John, the lone ranger in a wilderness setting, used as a central example throughout this thesis. John's beliefs about his location, the time of day, and his surroundings are subject to constant revision, as new contextual information becomes available. Initially, John may hold the belief that he is on a well-known path, but his location could be uncertain due to the lack of precise landmarks. In such a scenario, his beliefs are influenced not only by the information he has about the world but also by his position within that world the time, place, and his perspective. The introduction of K-contexts and S-contexts helps model this uncertainty by distinguishing between what John knows about the world (the world's state, w) and where he stands in relation to that world (his position, c).

By distinguishing between K-intensions and S-intensions, we offer a clearer view of John's Belief Revision process. In the beginning, John's Belief Revision would be based on K-intensions his knowledge about the world. However, once he becomes aware of the uncertainty surrounding his location and the context shifts (perhaps due to nightfall or disorientation), his Belief Revision will rely on S-intensions, taking into account both the time and place of his current situation. This shift reflects the importance of context in updating beliefs, highlighting the need for a Two-Dimensional framework when dealing with beliefs that are sensitive to the agent's position within the world. In a more structured environment, such as the investigative scenario involving Jane the detective, the need for 2D Belief Revision is also evident. Jane's beliefs about the events she is investigating are influenced by both her knowledge of the world (the set of facts, or K-contexts) and the specifics of her investigation, which are constantly evolving (S-contexts). As she uncovers new evidence, Jane updates her beliefs not just about what has happened, but about the relevance of that evidence in the context of the broader investigation. The use of a 2D framework helps illustrate how her beliefs about the world change as her understanding of the context her position in relation to the case evolves.

Similarly, the example of Lingens, who finds herself lost and disoriented, mirrors these dynamics. Lingens' uncertainty is rooted not only in the vagueness of the surrounding environment but also in her lack of a fixed reference point within it. In this case, the application of 2D Belief Revision allows us to track how her beliefs evolve as she shifts between different contextual perspectives. Initially, Lingens may revise her beliefs based on K-intensions (what she knows about the world), but as she navigates the unfamiliar environment, she must also account for changes in S-intentions, such as her position within that environment and the new information that arises as she gains or loses access to landmarks or navigational clues. This distinction between K-intensions and S-intensions also directly connects with the uncertainties that surround John, the lone ranger in the wilderness. His uncertainty about both his surroundings and his position within those surroundings is a prime example of the need for Two-Dimensional Belief Revision. John's situation exemplifies how the 2D framework can accommodate belief updates that consider both what is known about the world (the state of the world) and the agent's position within that world (the context or center of perception). As John revises his beliefs, he must adjust not only to the information he learns about the world but also to shifts in his perspective and understanding of where he stands in relation to that world.

In contrast, there are situations where a simpler, one-dimensional Belief Revision model might suffice. For example, in less context-sensitive scenarios, where an agent's belief is based solely on the state of the world and does not depend on the agent's position or perspective, the simpler one-dimensional approach might be both adequate and efficient. However, as illustrated through the wilderness example and the other cases considered, the complexity introduced by context-sensitive beliefs necessitates the use of a Two-Dimensional framework in certain circumstances. The practical applications of this framework in real-world reasoning scenarios such as navigation, investigation, and decision-making underscore the value of distinguishing between different kinds of intensions. By employing K-intensions and S-intensions, we not only provide a more accurate model of Belief Revision but also offer a more flexible approach that can accommodate the fluidity of real-world belief changes.

As we've explored throughout this thesis, the challenge of Belief Revision in the context of uncertainty, especially when dealing with context-sensitive or self-locating beliefs, necessitates a more refined approach than traditional one-dimensional models. The Two-Dimensional (2D) framework provides a structured method for accounting for both the agent's knowledge of the world and the agent's position within that world. This distinction between K-intensions (knowledge about the world) and S-intensions (knowledge about the agent's position) enriches our understanding of Belief Revision, particularly in scenarios where the context of belief such as time, space, and perspective plays a crucial role in shaping the belief update process.

The examples from earlier in the thesis, such as John, the lone ranger in the wilderness, Jane the detective, and Lingens lost in the forest, all illustrate the practical utility of this Two-Dimensional framework. In each case, the revision of beliefs involves not only the acquisition of new information about the world but also a shift in the context or perspective from which that information is understood. For John, revising his beliefs about his location in the wilderness requires not just the knowledge of where he is (K-intension) but also how he perceives his position in relation to the surrounding environment (S-intension). Similarly, Jane's investigation requires constant shifts in both what she knows about the case (K-intensions) and how that knowledge is framed within the evolving context of her investigation (S-intensions). In both cases, the Two-Dimensional framework allows for a more accurate and nuanced representation of belief change.

Importantly, the use of K-intensions and S-intensions directly addresses a gap in traditional Belief Revision models, which often fail to account for the complexity of context-sensitive beliefs. By introducing the distinction between these types of intensions, we provide a formal mechanism for revising beliefs that is sensitive to both the content of the world and the context of the agent's perspective. This is crucial when dealing with indexicals or self-locating expressions, where the meaning of a proposition depends on both the world and the agent's position within that world. The Two-Dimensional model offers a way to track these shifts, ensuring that Belief Revision can adapt as new context-dependent information becomes available. The 2D Belief Revision framework has important implications for real-world reasoning. As we have seen through the examples, Belief Revision is not just about updating the content of an agent's beliefs but also about adjusting those beliefs based on shifting contexts. In scenarios such as navigation (John's experience in the wilderness), investigation (Jane's detective work), or even interpersonal communication, the context in which information is acquired can dramatically affect how that information is interpreted and used. For instance, an agent may initially hold the belief that a certain event is true, but when faced with new information that alters the context (e.g., a change in time, location, or the agent's perspective), the agent may revise their belief accordingly. In these situations, a Two-Dimensional model is essential to accommodate both the knowledge about the world and the evolving context.

Moreover, the ability to apply this model to decision-making processes in complex, uncertain environments can improve the accuracy and adaptability of agents' beliefs. Whether in dynamic environments like the wilderness, as shown in the case of John, or in the shifting landscapes of detective investigations or personal relationships, agents are constantly faced with the need to update their beliefs based on both the state of the world and the context in which they find themselves. The Two-Dimensional approach provides a formal tool to handle such updates in a rational and structured manner.

Despite the strengths of the Two-Dimensional framework, it is not without its challenges. As with any formal system, managing the dual dimensions of Belief Revision can introduce complexity, especially in large-scale applications where the contexts themselves may be fluid or poorly defined. Furthermore, while this framework offers a robust method for updating beliefs, it may still need refinement when dealing with highly dynamic or multi-agent systems where multiple centers of perception must be reconciled. Future research could explore ways to simplify the application of the 2D framework in practical scenarios, such as by integrating hybrid logics or introducing new semantics for modeling complex belief interactions in multi-agent environments. Additionally, there may be room to refine the relationship between K-intensions and S-intensions, potentially developing more nuanced models for handling the interplay between knowledge and context in Belief Revision. At one glance, the Two-Dimensional Belief Revision framework introduced in this thesis provides a powerful tool for addressing the complexities of belief change in the presence of context-sensitive information. By distinguishing between K-intensions and S-intensions, we are able to capture the full spectrum of Belief Revision processes, from the agent's knowledge of the world to their position within that world. This distinction is crucial in ensuring that Belief Revision models accurately reflect real-world reasoning, especially when dealing with indexicals, self-locating beliefs, and other context-sensitive phenomena. As we have seen through various examples, such as the lone ranger, the detective, and the lost individual, the 2D framework allows for a more nuanced and accurate representation of how beliefs evolve over time and across contexts. By building on these ideas, this thesis has contributed to the development of a more comprehensive model of Belief Revision that can handle the complexities of real-world reasoning.

The Two-Dimensional framework for Belief Revision not only enhances our understanding of how beliefs evolve in the face of new context-sensitive information, but it also provides a deeper insight into rational decision-making in uncertain environments. In real-world scenarios, decisions are often made under conditions of uncertainty, where the agent's beliefs are shaped by both the content of the world and the specific context in which they find themselves. The framework we have developed in this thesis allows us to model this dual influence, providing a tool for decision-making that accounts for both the agent's knowledge and their perspective. The challenge of rational decision-making in complex environments whether in navigation, investigation, or negotiation becomes significantly more tractable when we acknowledge the importance of context. For example, in the wilderness, John's decisions about how to navigate and where to seek shelter depend not only on his knowledge of the world (e.g., whether he is near a water source) but also on his position within the world (e.g., his sense of direction, the time of day). By adopting a 2D approach, we are able to capture these dynamic shifts in belief and update them accordingly as new context-sensitive information is introduced.

Similarly, in detective work, Jane's ability to revise her beliefs about the case depends on both the information she gathers and her understanding of how that information fits into the broader context of her investigation. By treating beliefs as being both world-dependent and context-dependent, the 2D framework ensures that revisions are made in a way that reflects both the factual content of the case and Jane's position in relation to the evidence. The same applies to Lingens, who, upon being lost in the forest, must adjust her beliefs not only based on what she knows about the world but also considering the shifts in her own perception and knowledge as she experiences different environments. In each of these cases, adopting the 2D Belief Revision framework provides a way of making decisions that takes into account not only the facts but also the context in which those facts are interpreted. This dual evaluation allows for more informed and rational decisions that are responsive to changes in both the world and the agent's perspective, making the model especially useful for handling complex decision-making processes in uncertain environments.

While the Two-Dimensional framework for Belief Revision provides a robust theoretical model, its application in practice, especially in large-scale or multi-agent systems, presents challenges. One of the primary difficulties lies in managing the dual dimensions of context and content across complex scenarios. For instance, in multi-agent environments, where different agents may hold different contextual perspectives, reconciling these perspectives to achieve a global ordering of beliefs becomes a non-trivial task. The interplay between the various centers of perception (K-intensions and S-intensions) in these multi-agent systems requires further investigation to develop more refined models that can account for the shifting perspectives of multiple agents. Moreover, while we have demonstrated the utility of the 2D model in scenarios involving indexicals and self-locating beliefs, there remains a need for further work in simplifying its application. In particular, integrating the 2D framework with decision-theoretic models, such as expected utility theory or game theory, could offer new ways to formalize rational decision-making in uncertain and dynamic environments. This could allow for the development of hybrid models that combine the strengths of both logical reasoning and probabilistic decision-making.

Another avenue for future work involves exploring the potential for integrating the 2D Belief Revision framework with other areas of research, such as cognitive science, psychology, and artificial intelligence. Understanding how humans revise their beliefs in real-world scenarios, and how artificial agents can be made to reason in a similar manner, would provide valuable insights into the practical application of Belief Revision theories. The development of computational models that implement 2D Belief Revision could also open up new possibilities for designing intelligent systems capable of reasoning in dynamic, context-sensitive environments. Looking ahead, the Two-Dimensional framework for Belief Revision offers promising directions for future research, particularly in multi-agent systems, decision theory, and the integration of cognitive and computational models. The challenges of managing dual dimensions of context and content are non-trivial, but the rewards of developing a more refined and flexible model for Belief Revision are significant. As we continue to refine and apply these ideas, we can expect to see more sophisticated and rational approaches to decision-making in complex, uncertain environments.

9.2 Closing Remarks

The development of a two-dimensional framework for belief revision marks a significant contribution to the field of formal epistemology, particularly in how we handle context-sensitive and self-locating beliefs. By distinguishing between K-intensions and S-intensions, we have provided a more nuanced way of understanding how beliefs evolve when agents must consider both the content of the world and their position within that world. This dual-layered approach allows us to model belief revision more accurately, ensuring that the revision process is sensitive to both the world and the agent's context. In practical terms, this framework is applicable in a wide range of scenarios, including decision-making in uncertain environments, such as navigation, investigation, and social interaction. By capturing the dynamics between content and context, this framework offers a way to model and reason about belief changes that occur when agents update their beliefs based on both new information and shifts in their perspective. The examples of John in the wilderness, Jane the detective, and Lingens lost in the forest illustrated how the two-dimensional approach can be used to model belief revision in these situations. In each case, context-sensitive reasoning plays a crucial role in the process of rational belief update. Moreover, the application of two-dimensional belief revision can extend to multi-agent systems, where different agents may hold different contextual perspectives. This opens up exciting possibilities for artificial intelligence, where systems must reason about and adapt to the perspectives of multiple agents interacting with each other. As decision-making becomes increasingly complex, especially in dynamic and multi-agent environments, the need for more sophisticated models of belief revision becomes clear. By incorporating the two-dimensional framework, we provide a robust tool for making rational decisions in uncertain and shifting contexts.

As we conclude this work, it is important to reflect on the broader implications of the two-dimensional belief revision framework and its potential applications in future research. The challenges associated with managing the dual dimensions of context and content are non-trivial, but the benefits of a refined model for belief revision far outweigh these challenges. Through this framework, we have been able to account for both the agent's knowledge about the world and their perspective within that world, offering a more comprehensive and realistic approach to belief revision. In future work, one promising direction is the integration of the two-dimensional belief revision framework with decision-theoretic models, such as expected utility theory and game theory. This could provide a more formalized approach to decision-making in uncertain environments, allowing for the development of systems that can reason about belief revision and make decisions based on both the content and context of their beliefs. Additionally, applying the framework to multi-agent systems could yield valuable insights into how agents with different perspectives can rationally update their beliefs in a shared environment. Another interesting possibility for future research is the application of this framework to cognitive science and psychology. Understanding how humans revise their beliefs in real-world scenarios and how artificial agents can be made to reason in a similar manner could provide a deeper understanding of the processes of belief formation and revision. This could inform the development of intelligent systems capable of reasoning in dynamic, context-sensitive environments.

In sum, the two-dimensional belief revision framework presented in this thesis offers a robust and flexible model for handling belief updates in the presence of context-sensitive information. By distinguishing between K-intensions and S-intensions, we have provided a clearer understanding of how beliefs evolve in uncertain environments. The framework proposed here is just one step toward understanding how agents adapt their beliefs in response to shifting contexts and has the potential to continue revising our beliefs.

Bibliography

[1] ÅGOTNES, T., VAN DER HOEK, W., AND WOOLDRIDGE, M. On the Logic of Preference and Judgment Aggregation. *Autonomous Agents and Multi-Agent Systems 22*, 1 (2011), 4–30.

[2] ALCHOURRÓN, C. E., GÄRDENFORS, P., AND MAKINSON, D. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic 50*, 2 (1985), 510–530.

[3] ALCHOURRÓN, C. E., GÄRDENFORS, P., AND MAKINSON, D. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic 50*, 2 (1985), 510–530.

[4] ANTONELLI, G. A. Non-Monotonic Logic. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., summer ed. 2010.

[5] ANTONIOU, G. *Nonmonotonic Reasoning*. MIT Press, 1997.

[6] ARLÓ-COSTA, H., AND PEDERSEN, A. P. Social Norms, Rational Choice and Belief Change. In *Belief Revision Meets Philosophy of Science*, E. J. Olsson and S. Enqvist, Eds. Springer, 2010, pp. 163–212.

[7] ARROW, K. J. *Social Choice and Individual Values*. Wiley, 1951.

[8] ARROW, K. J. *Social Choice and Individual Values*. Yale UP, 2012.

[9] BAGARIA, J. Set Theory. In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Spring 2023 ed. Metaphysics Research Lab, Stanford University, 2023.

[10] BATENS, D. Adaptive Logics and Dynamic Proofs: Mastering the Dynamics of Reasoning, with Special Attention to Handling Inconsistency, 2010. Manuscript available at http://logica.ugent.be/adlog/toc.pdf.

[11] BENTHEM, J. V. *Modal Logic for Open Minds*. CSLI, 2010.

[12] BERTO, F., AND JAGO, M. Impossible Worlds. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., fall ed. 2018.

[13] BIANCHI, C. Context of Utterance and Intended Context. In *Modeling and Using Context: Third International and Interdisciplinary Conference CONTEXT*. Springer, 2001, pp. 73–86.

[14] BLACKBURN, P. Arthur Prior and Hybrid Logic. *Synthese 150*, 3 (2006), 329–372.

[15] BOUTILIER, C. Revision sequences and nested conditionals. *Journal of Philosophical Logic 22* (1993), 1–27.

[16] BRAÜNER, T. Hybrid Logic. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., winter ed. 2014.

[17] BRAÜNER, T. Hybrid Logic. In *Handbook of Philosophical Logic*, D. M. Gabbay and F. Guenthner, Eds., vol. 17. Springer, 2014, pp. 1–77.

[18] BREWKA, G., DIX, J., AND KONOLIGE, K. *Nonmonotonic Reasoning: An Overview*. CSLI Publications, 1997.

[19] CARNAP, R. *Logical Foundations of Probability*. University of Chicago, 1962.

[20] CARNIELLI, W., AND PIZZI, C. *Modalities and Multimodalities*. Springer, 2008.

[21] CHALMERS, D. J. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, New York, 1996.

[22] CHALMERS, D. J. On sense and intension. In *Meaning and Modality*, J. Hawthorne and D. Manley, Eds. Oxford University Press, 2002, pp. 56–98.

[23] CHALMERS, D. J. Epistemic two-dimensional semantics. *Philosophical Studies 118* (2004), 153–226.

[24] CHALMERS, D. J. The foundations of two-dimensional semantics. In *Two-Dimensional Semantics: Foundations and Applications*, M. García-Carpintero and J. Macià, Eds. Oxford University Press, 2006, pp. 55–140.

[25] DARWICHE, A., AND PEARL, J. On the logic of iterated belief revision. *Artificial Intelligence 89* (1997), 1–29.

[26] Dokow, E., and Holzman, R. Aggregation of Binary Evaluations. Tech. rep., Citeseer, 2005.

[27] Elga, A. Self-Locating Belief and the Sleeping Beauty Problem. *Analysis 60*, 266 (2000), 143–147.

[28] Feldman, A. M., and Serrano, R. Arrow's impossibility theorem: Two simple single-profile versions. *Brown University Department of Economics Working Paper*, 2006-11 (2008).

[29] Fermé, E., and Hansson, S. O. Belief revision. *Stanford Encyclopedia of Philosophy* (2021). URL: `https://plato.stanford.edu/entries/belief-revision/`.

[30] Fishburn, P. C. Arrow's Impossibility Theorem: Concise Proof and Infinite Voters. *Journal of Economic Theory 2*, 1 (1970), 103–106.

[31] Galton, A. Temporal Logic. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., fall ed. 2008.

[32] Gärdenfors, P. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT press, 1988.

[33] Gauker, C. Contexts in Formal Semantics. *Philosophy Compass 5*, 7 (2010), 568–578.

[34] Genin, K., and Huber, F. Formal Representations of Belief. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., spring ed. 2021.

[35] Gettier, E. L. Is justified true belief knowledge? *Analysis 23*, 6 (1963), 121–123.

[36] Grice, H. P. Logic and Conversation. In *Speech Acts*, P. Cole and J. L. Morgan, Eds. Academic, 1975, pp. 41–58. Volume 3 of Syntax and Semantics.

[37] Grove, A. Two Modellings for Theory Change. *Journal of Philosophical Logic 17*, 2 (1988), 157–170.

[38] Hájek, A. Interpretations of Probability. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., fall ed. 2019.

[39] Hansson, S. O. A textbook of belief dynamics: Theory change and database updating. *Kluwer Academic Publishers* (1999).

[40] Hitchcock, C. Beauty and the Bets. *Synthese 139*, 3 (2004), 405–420.

[41] Hughes, G. E., and Cresswell, M. J. A new introduction to modal logic. *Studia Logica 62*, 3 (1996), 439–441.

[42] Jackson, F. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press, Oxford, 1998.

[43] Jørgensen, K. F., and Blackburn, P. R. Indexical Hybrid Tense Logic. In *Advances in Modal Logic 9: Papers from the Ninth Conference on "Advances in Modal Logic"*, T. Bolander, T. Braüner, S. Ghilardi, and L. S. Moss, Eds. College Publications, 2012, pp. 144–160.

[44] Kamp, H. Formal Properties of 'Now'. *Theoria 37*, 3 (1971), 227–273.

[45] Kaplan, D. An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals. In *Themes from Kaplan*, J. Almog, J. Perry, and H. Wettstein, Eds. Oxford UP, 1977, pp. 481–563.

[46] Kaplan, D. On the Logic of Demonstratives. *Journal of Philosophical Logic 8*, 1 (1979), 81–98.

[47] Katsuno, H., and Mendelzon, A. O. On the Difference Between Updating a Knowledge Database and Revising It. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference (KR91)*, J. Allen, R. Fikes, and E. Sandewall, Eds. Morgan Kaufmann, 1991.

[48] Kern-Isberner, G. *Conditionals in Nonmonotonic Reasoning and Belief Revision: Considering Conditionals as Agents*, vol. 2087. Springer, 2001.

[49] Kindermann, D. Varieties of Centering and de se Communication. In *About Oneself: De Se Thought and Communication*, M. García-Carpinterom and S. Torre, Eds. Oxford UP, 2016, pp. 307–339.

[50] Kirman, A. P., and Sondermann, D. Arrow's Theorem, Many Agents, and Invisible Dictators. *Journal of Economic Theory 5*, 2 (1972), 267–277.

[51] Klawonn, F., Gebhardt, J., and Kruse, R. The Context Model from the Viewpoint of Logic. In *DGOR/ÖGOR: Papers of the 21th Annual Meeting of DGOR in Cooperation with ÖGOR*, K.-W. Hansmann, A. Sachem, M. Jarke, W. E. Katzenberger, and A. Marusev, Eds. Springer, 1993, pp. 288–295.

[52] Kripke, S. A. *Naming and Necessity*. Harvard University Press, Cambridge, MA, 1980. Originally published as lectures in 1970.

[53] Kyburg, H. E., and Teng, C. M. *Uncertain Inference*. Cambridge UP, 2001.

[54] Leitgeb, H. Reducing Belief Simpliciter to Degrees of Belief. *Annals of Pure and Applied Logic 164*, 12 (2013), 1338–1389.

[55] Leitgeb, H. *The Stability of Belief: How Rational Belief Coheres with Probability*. Oxford UP, 2017.

[56] Lewis, D. Attitudes de dicto and de se. *Philosophical Review 88*, 4 (1979), 513–543.

[57] Lewis, D. Scorekeeping in a Language Game. In *Semantics From Different Points of View*, R. Bäuerle, U. Egli, and A. von Stechow, Eds. Springer, 1979, pp. 172–187.

[58] Lewis, D. *Counterfactuals*. Wiley-Blackwell, 2001.

[59] Lewis, D. K. Attitudes de dicto and de se. *Philosophical Review 88*, 4 (1979), 513–543.

[60] Lewis, D. K. *On the Plurality of Worlds*. Blackwell, Oxford, 1986.

[61] Liao, S. What are Centered Worlds? *Philosophical Quarterly 62*, 247 (2012), 294–316.

[62] Lindström, S., and Rabinowicz, W. On Probabilistic Representation of Non-probabilistic Belief Revision. *Journal of Philosophical Logic 18*, 1 (1989), 69–101.

[63] Ludlow, P. Logical Form and the Hidden-Indexical Theory: A Reply to Schiffer. *Journal of Philosophy 92*, 2 (1995), 102–107.

[64] Makinson, D. Bridges Between Classical and Nonmonotonic Logic. *Logic Journal of IGPL 11*, 1 (2003), 69–96.

[65] McDermott, D. Nonmonotonic Logic II: Nonmonotonic Modal Theories. *Journal of the ACM 29*, 1 (1982), 33–57.

[66] Mellor, H. D. I and Now. *Proceedings of the Aristotelian Society 89* (1988), 79–94.

[67] Meyer, U. Time and Modality. In *The Oxford Handbook of Philosophy of Time*, C. Callender, Ed. Oxford UP, 2011.

[68] Milano, S. *De se Beliefs and Centred Uncertainty*. PhD thesis, The London School of Economics and Political Science, London, England, Janary 2018.

[69] Moore, R. C. Semantical considerations on nonmonotonic logic. *Artificial Intelligence 25*, 1 (1985), 75–94.

[70] Morreau, M. Arrow's Theorem. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., winter ed. 2019.

[71] Nayak, A. Lexicographic revision, conditional logic, and nonmonotonic reasoning. *Artificial Intelligence 141* (2003), 139–163.

[72] Park, I. Simultaneous Belief Updates via Successive Jeffrey Conditionalization. *Synthese 190*, 16 (2013), 3511–3533.

[73] Perry, J. The Problem of the Essential Indexical. *Noûs 13*, 1 (1979), 3–21.

[74] Perry, J. Frege on Demonstratives. In *Demonstratives*, P. Yourgrau, Ed. Oxford UP, 1990, pp. 50–70. Oxford Readings in Philosophy Series.

[75] Perry, J. Indexicals and Demonstratives. In *Companion to the Philosophy of Language*, R. Hale and C. Wright, Eds. Blackwells, 1997, pp. 586–612.

[76] Perry, J. Reflexivity, Indexicality and Names. In *Direct Reference, Indexicality and Proposition Attitudes*, W. Künne, A. Newen, and M. Anduschus, Eds. 1997.

[77] Perry, J. *Temporal Indexicals*. Wiley, 2013.

[78] Perry, J. *Revisiting the Essential Indexical*. CSLI, 2020.

[79] Predelli, S. 'I Am Not Here Now'. *Analysis 58*, 2 (1998), 107–115.

[80] Predelli, S. I Am Still Not Here Now. *Erkenntnis 74*, 3 (2011), 289–303.

[81] Prior, A. *Papers on Time and Tense*. Oxford UP, 2003.

[82] Quine, W. V. O. *Word and Object*. MIT Press, Cambridge, MA, 1960.

[83] Quine, W. V. O. Propositional Objects. *Crítica: Revista Hispanoamericana de Filosofía 2*, 5 (1968), 3–29.

[84] Quine, W. V. O. *The Roots of Reference*. Open Court, La Salle, IL, 1974.

[85] Quine, W. V. O. *Quiddities: An Intermittently Philosophical Dictionary*. Harvard University Press, Cambridge, MA, 1987.

[86] Rabern, B. Against the identification of propositions with sets of worlds. *Synthese 179*, 3 (2011), 321–338.

[87] RADULESCU, A. The Logic of Indexicals. *Synthese 192*, 6 (2015), 1839–1860.

[88] REITER, R. A Logic for Default Reasoning. *Artificial intelligence 13*, 1-2 (1980), 81–132.

[89] ROTT, H. Two Dogmas of Belief Revision. *Journal of Philosophy 97*, 9 (2000), 503–522.

[90] ROTT, H. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*. Oxford UP, 2001.

[91] ROTT, H. *Bounded Revision: Two-Dimensional Belief Change Between Conservative and Moderate Revision*. Springer, 2012.

[92] RUFFINO, M. Deferred Utterances and Proper Contexts. *Logic, Norms and Ontology: Recent Essays in Luso-Brazilian Analytic Philosophy 4*, 34 (2012), 807–822.

[93] RUSSELL, G. Lessons from the Logic of Demonstratives: What Indexicality Teaches Us about Logic, and Vice Versa. In *New Waves in Philosophical Logic*, G. Restall and G. Russell, Eds. Springer, 2012, pp. 190–216.

[94] SCHIFFER, S. The Hidden-Indexical Theory's Logical-Form Problem: A Rejoinder. *Analysis 56*, 2 (1996), 92–97.

[95] SCHROETER, L. Two-Dimensional Semantics. In *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., spring 2021 ed. 2021.

[96] SCHULZ, M. The Dynamics of Indexical Belief. *Erkenntnis 72*, 3 (2010), 337–351.

[97] SCHWARZ, W. Changing Minds in a Changing World. *Philosophical Studies 159*, 2 (2012), 219–239.

[98] SEN, A. *Collective Choice and Social Welfare*. Harvard UP, 2018.

[99] STALNAKER, R. Indexical Belief. *Synthese 49*, 1 (1981), 129–151.

[100] STALNAKER, R. A Note on Non-Monotonic Modal Logic. *Artificial Intelligence 64*, 2 (1993), 183–196.

[101] STALNAKER, R. A note on non-monotonic modal logic. *Artificial Intelligence 64*, 2 (1993), 183–196.

[102] STALNAKER, R. *Context*. Oxford UP, 2014.

[103] Stojanovic, I. The Problem of de se Assertion. *Erkenntnis 76*, 1 (2012), 49–58.

[104] Stojanovic, I. Speaking About Oneself. In *About Oneself: De Se Thought and Communication*, M. García-Carpinterom and S. Torre, Eds. Oxford UP, 2016, pp. 200–219.

[105] Tangian, A. Computational Application of the Mathematical Theory of Democracy to Arrow's Impossibility Theorem (How Dictatorial are Arrow's Dictators?). *Social Choice and Welfare 35*, 1 (2010), 129–161.

[106] Titelbaum, M. G. The Relevance of Self-locating Beliefs. *Philosophical Review 117*, 4 (2008), 555–606.

[107] Titelbaum, M. G. Self-locating Credences. In *The Oxford Handbook of Probability and Philosophy*, A. Hájek and C. Hitchcock, Eds. Oxford UP, 2016.

[108] Torre, S. Centered Assertion. *Philosophical Studies 150*, 1 (2010), 97–114.

[109] Weber, C. Centered Communication. *Philosophical Studies 166*, 1 (2013), 205–223.

[110] Weber, C. Being at the Centre: Self-location in Thought and Language. In *About Oneself: De Se Thought and Communication*, M. García-Carpinterom and S. Torre, Eds. Oxford UP, 2016, pp. 246–271.

[111] Weymark, J. A. Arrow's Theorem with Social Quasi-Orderings. *Public Choice 42*, 3 (1984), 235–246.

[112] Williams, D. C. The Myth of Passage. *Journal of Philosophy 48*, 15 (1951), 457–472.

[113] Wilson, R. On the Theory of Aggregation. *Journal of Economic Theory 10*, 1 (1975), 89–99.