Aggregating Credences into a Belief

Minkyung Wang

2022 München

Aggregating Credences into a Belief

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Philosophie an der Ludwig-Maximilians-Universität

München

vorgelegt von Minkyung Wang

aus Busan, Südkorea

Referent: Prof. DDr. Hannes Leitgeb

Korreferent: Prof. Dr. Christian List

Tag der mündlichen Prüfung: 06.05.2022

Acknowledgements

This thesis could not have been completed if there had been no assistance and support from many people. Words might not describe how thankful I am, but I want to express my sincere gratitude.

First and foremost, I am very grateful to my supervisor, Hannes Leitgeb, for his excellent guidance and brilliant comments on my work. Not only his support and help but also his compliments and encouragements have made me proceed further. From his invaluable feedback, I have been inspired and got a lot of ideas, and from his philosophical insights, I have learned how to be a good philosopher and how to present my thoughts. He has also taught me what is important in philosophy and how to present my thoughts, which made my writings balanced from too many math-oriented ones. He has been, since I heard his talk in Berlin and decided to move to Munich, and will be the philosopher that influences my work the most and that I want to be like.

I also want to thank Christian List, the second examiner, and my mentor, for giving me many ideas about developing this work further. His insightful and crystal clear works have influenced me a lot, and so I am very pleased to have him as the examiner. I am also indebted to my master's thesis supervisor Stephan Hartman for kindly agreeing to be a second supervisor at the beginning of my Ph.D. studies. I presented parts of this work at the formal epistemology workshop 2022, and I thank all the organizers and participants. I am particularly grateful to Hanti Lin. He was the commentator and made really interesting, productive, and helpful comments about how to situate my research in a broader philosophical context. I also thank Catrin Campbel-Moore for follow-up research questions and discussions. In addition, I am grateful to all the people I met at the MCMP. Especially, I thank Jean Baccelli for weekly discussions during his lectures about social epistemology.

I want to extend my gratitude to those who taught me mathematics at LMU Munich and formal philosophy at HU Berlin before my Ph.D. studies. I want to thank especially Hans-Dieter Donder and Helmut Schwichtenberg. They taught me not only mathematics for many years but also how mathematical languages can be so succinct and beautiful. I am also indebted to Karl-Georg Niebergall, who first taught me various logics and offered a daily reading group even during summer vacation. His pure and honest love for logic motivated me to do formal philosophy. Moreover, I sincerely thank Min Geum, who introduced me to philosophy and recommended studying philosophy in Germany. In particular, I am grateful to him for reminding me, through his activities, that my research should contribute to society's progress.

I also want to thank my dear family —Bokhee Kim, Jee-Seok Wang, Minhye Wang, and Minjeoung Wang — and Chisu's family — Sunyeop Choi, Dongsin Kim, Youngran Kim, and Nawon Park — for their unconditional sustained support. Further, I would like to thank my wonderful friends for their consistent friendship over the years, especially, Jeonghyun Lee, Hongyeol Ahn, Byeonghun Lee, Taeho Noh, Nari Yun and Mijeong Lee. I thank all my family members and friends for cheering me up and helping me forget difficulties.

Last but certainly not least, I dedicate this work to my best friend, closest family, and the most congenial study- and work-mate Chisu Kim. He has been with me on this long journey, from south Korea to Berlin to Munich, studying philosophy and mathematics from scratch and sharing all thrills and adventures. He has been discussing everything not only on this work but also on all I am interested in and giving me insightful and fabulous answers to my relentless questions. It is not an overstatement that this work is some part of the results of my collaboration with him. Actually, he proofread all of this work, and many parts of this work are based on joint work with him. If it weren't for ceaseless exciting conversations with him, I might not have been so excited and pleased throughout my studies. I learn a lot from his passion for learning and his affection for wisdom, enabling me to withstand adversity and make bold challenges.

This work has been supported generously by the German Academic Scholarship Foundation (Studienstiftung des Deutschen Volkes).

Abstract

This thesis proposes a new research topic of how we should aggregate multiple individual credences on logically connected issues into a collective binary belief: heterogeneous belief aggregation. We argue that heterogeneous belief aggregation is worth studying because there are many situations where credences and binary beliefs are more appropriate as inputs and outputs of aggregation procedures, respectively. The main problem is that it is vulnerable to a dilemma like the discursive dilemma or the lottery paradox: issue-wise independent procedures might not ensure deductive closure and consistency. Confronting this situation, we have two main questions: how to formulate and generalize the dilemma, and what kinds of aggregation procedures can avoid the dilemma and obtain rational collective beliefs.

To answer the first question, we employ the axiomatic approach to deal with general aggregation procedures as in judgment aggregation and social choice theory. We investigate which kinds of individual and collective rationality requirements and which properties of aggregation procedures should be imposed on heterogeneous belief aggregation, and which of their combinations are impossible. We mainly assume deductive closure rather than completeness, in contrast with most of the judgment aggregation literature. Moreover, we address impossibility results without anonymity conditions, which cannot be considered in belief binarization. This leads to three kinds of impossibility results, and we also determine the sufficient and necessary agenda condition for each of the results. Furthermore, we analyze similarities and differences between our proofs and other related proofs and conclude that the problem of heterogeneous belief aggregation is not reducible to the other related problems. Moreover, we show that our methods can be applied to other similar impossibilities.

For the second question, we explore specific heterogeneous belief aggregation procedures and their properties. There can be two kinds of heterogeneous belief aggregation procedures: collective belief binarization combined with a probabilistic opinion pooling method, and direct rules.

As for collective belief binarization, belief binarization theories are applicable. To this end, we first analyze the existing threshold-based procedures, especially those that relax the Lockean thesis and preserve rationality. We categorize them as local-threshold rules — where thresholds depend on probability measures — and world-threshold rules — where thresholds are applied not to an issue but to a possible world. Their characteristics are captured by the property of local monotonicity and world monotonicity, respectively. We compare and relate these properties with other existing properties like being stable in the stability theory of belief and with new —

to be introduced — properties. Whether some existing rational procedures, like the camera shutter rule, satisfy these properties is an interesting and philosophically important question. We provide geometrical characterizations of some of the properties to answer this question. Furthermore, we propose that convexity norms should be discussed in the context of belief binarization. We introduce various kinds of convexity norms and examine whether the relevant procedures satisfy them.

What is more, we propose two novel kinds of belief binarization methods that preserve rationality but are not based on thresholds: *distance-based binarization* and *epistemic-utility-based binarization*. The first is a holistic one minimizing the distance from a given probability measure to the resulting binary belief. The second one is based on an accuracy norm minimizing expected inaccuracy. We devise novel ways to measure the required distances and inaccuracies. Moreover, we study distance minimization with *Bregman divergence*, utility maximization with *strict proper scores*, and their relationship.

Direct heterogeneous belief aggregation rules will also be proposed and studied regarding threshold, distance and epistemic-utility. We provide a new classification and characterization of them. Furthermore, we investigate some norms that are especially relevant in social contexts, such as various unanimity norms and convexity norms interpreted in social contexts, and commutativity norms, which govern the relationship between direct rules and combinations of probabilistic opinion pooling and collective belief binarization.

Putting all this together, we conclude that heterogeneous belief aggregation is an philosophically fruitful topic that deserves attention. Heterogeneous belief aggregation can be seen as a general framework, where not only heterogeneous belief aggregation but also probabilistic opinion pooling, judgment aggregation, and belief binarization are studied in connection to each other. First, studying heterogeneous belief aggregation is by itself interesting and cannot be reduced to other research fields: we can deal with different rationality norms in social contexts and address properties characteristic for heterogeneous belief aggregation. Moreover, it is not only the direct rules but also the different possible combinations of methods from different research areas that makes this whole endeavor to be more sum of its parts. Indeed, second, this framework bridges independently developed research areas: first, we can apply well-developed formal theories in formal epistemology like belief binarization theories and epistemic decision theories to the belief aggregation problem. Second, this framework enables us to add social contexts to belief binarization problems and epistemic decision theories, which can be extended to cover also social beliefs. Our theory of heterogeneous belief aggregation can be applied to the (collective) belief binarization problem and epistemic (collective) decision theory. In this way, the thesis fills, or at least narrows, the gap between individual epistemology and collective epistemology.

Contents

Chapter 1

Introduction

It is a fundamental problem in society and politics how a group of agents makes decisions rationally and fairly. Simple aggregation methods, such as majority voting, are used everywhere to form various kinds of collective attitudes. For example, an expert panel of scientists aggregates beliefs about what the real world is like, and political voting is usually concerned with aggregating voter preferences about what the society should be like. When aggregating beliefs, we prioritize epistemic norms such as the veritistic norm and consistency norm, while we usually emphasize fairness norms in the case of aggregating preferences or moral judgments. This thesis focuses on how a group should aggregate their beliefs.[1][2]

In philosophy, belief aggregation has been studied in social epistemology (Goldman (1999), Dietrich (2021)).[3] In particular, aggregating individual binary beliefs into a

[1]Belief aggregation is a function of individuals' beliefs. To form a collective belief, it is not necessary for the input to be individuals' beliefs. If we use belief elicitation mechanisms such as the *prediction market* (Wolfers & Zitzewitz (2004)), the input can be individuals' actions, which are supposed to reveal individuals' beliefs.

[2]Belief is a thinner concept than knowledge in that it neither needs to be true nor requires justification. Therefore, our thesis will not deal with collective justified beliefs or collective knowledge.

[3]Outside of social epistemology, belief aggregation has been studied in political philosophy too. According to *epistemic democracy* (Goodin & Spiekermann (2018)), democratic procedures can

collective binary belief and aggregating individual credences into a collective credence have been the main topics. This thesis presents a new belief aggregation problem: *aggregating credences into a binary belief.* This problem has not been discussed in social epistemology, although it has long been pointed out that the two types of beliefs are distinct and governed by different norms and their bridge principles deserve attention. An in-depth discussion of this issue can be found in formal epistemology, which has mainly dealt with individual epistemology. Unlike traditional epistemology, formal epistemology has emphasized credences and investigated rational relations between binary beliefs and credences. This thesis will study the new problem in social epistemology using existing and new results, to be developed in this thesis, in formal epistemology. To elaborate on our research project, we first summarize how each of the two kinds of beliefs has been studied in formal epistemology.[4]

Binary belief — which is also called all-or-nothing-belief/categorical belief/outright belief/qualitative belief/full belief — allows only two options regarding a proposition, say A: she believes that A or she does not believe that A. If we consider two propositions A and not-A[5] together, three doxastic attitudes are permissible: (1) belief — she believes that A and she does not believe that not-A; (2) disbelief — she does not believe that A and she believes that not-A; (3) suspension — she believes neither that A nor that not-A. Different kinds of models and norms governing binary beliefs have been suggested: standard propositional logic can be used to represent a belief set that collects all believed propositions; *epistemic logic* internalizes the belief modal into the logical language, and therefore can represent higher-order beliefs (Hintikka (1962)); *AGM belief revision theory* focuses on revision methods of a belief set, taking inconsistency management into account, in light of a newly believed proposition (Gärdenfors (1988), Hansson (1999)).

In contrast, credence — also called degree of belief/quantitative belief/graded belief/partial belief — usually allows infinitely many options: she believes with a degree of certainty x that A, where x is a numerical value. Numerous ways of modelling

be vindicated using, e.g., *Jury theorems*, which state that it is more likely that democratically aggregated beliefs from a larger group, under certain conditions, outperform the ones from smaller groups in terms of tracking the truth (Dietrich & Spiekermann (2021)).

[4]We can distinguish metaphysical questions about beliefs from normative ones. Typical metaphysical questions are as follows: do binary belief and credence refer to the same doxastic state or distinct doxastic states? How are they related? Can a binary belief be reduced to a credence? Although these questions are crucial, especially in cognitive science, we will not deal with them. For a survey of metaphysical issues of beliefs, see Jackson (2020) and Chapter 1 in Leitgeb (2017a). The other questions about beliefs are normative ones: how should we represent and organize our beliefs rationally, and update our beliefs in light of new evidence? Which norms should we impose on distinct doxastic states? What are rational bridge principles between distinct doxastic states or between different representations of a doxastic state? These questions are independent of metaphysical ones since normative models can be criteria when evaluating existing ones, even if they cannot be realized. This thesis focuses on normative questions about (aggregating) beliefs, and we do not assume any metaphysical position. We also hope that our results are compatible with any existing positions. We also remark that the level of rationality depends on background assumptions — e.g., we may provide different epistemic norms depending on whether the agents are supposed to store a finite or infinite number of propositions.

[5]not-A means that it is not the case that A.

credences have been suggested: the mainstream approach employs a probability function, which is mathematically well-developed[6]; some approaches such as the theories of imprecise probabilities and the *Dempster-Shafer theory* provide generalizations of probability functions[7]; some approaches are radically different from probability functions, such as Spohn's ranking functions (Spohn (2012)). In our thesis, we accept *probabilism*, and presuppose that credences should be modelled as probability functions or numerical functions extendable to probability functions.

Binary beliefs and credences have their own merits and disadvantages. On the one hand, credences are informative and sophisticated, while binary beliefs are computationally efficient and thus human-friendly. On the other hand, credences are usually computationally demanding even for ideal reasoners, while binary beliefs are uninformative in some complex contexts, e.g., decision contexts in uncertain environments. Because one does not dominate the other, we have a good reason to embrace them in belief aggregation contexts. Considering two types of belief, we have the following diagram, where all possible belief aggregation problems are depicted.[8]

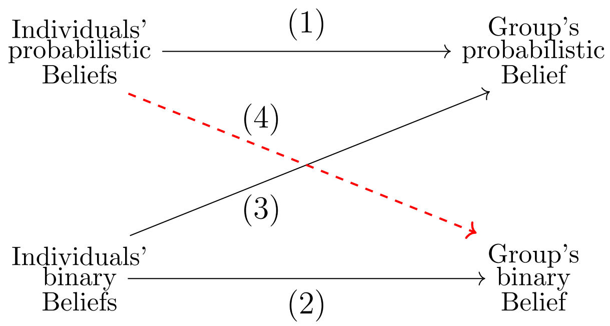


Figure 1.1

In the cases of (1) and (2), the input and output data types are the same, while they are different in the cases of (3) and (4). The existing research about belief aggregation has focused on the cases of (1) and (2). In case (1), both the input and output are credences, specifically probability functions, and this problem has been extensively studied in *probabilistic opinion pooling*. In case (2), binary beliefs are the input and output, and this problem has been investigated in the *judgment aggregation* and *belief merging* literature. In contrast, the cases (3) and (4) have rarely been studied. We claim that there are some situations where the case (4) is needed. To this end, we will deal with the input and output data types separately.

[6]Different kinds of arguments have vindicated this approach: some invoke practical rationality, such as *the Dutch Book argument* (Pettigrew (2020b)); epistemic rationality such as *epistemic decision theory* (Joyce (1998), Pettigrew (2016)); partition invariance (Leitgeb (2021)).

[7]For an overview of various models of credences, see Huber & Schmidt-Petri (2009) and Halpern (2017)

[8]More complex belief aggregation methods might include different dynamic processes; sequential evidence learning (Blackwell and Dubins (1962)); deliberation processes (the *consensus formation model* (DeGroot (1974), Lehrer and Wagner (1981)); higher-order evidence learning (*peer disagreement* literature, *supra Bayesianism* (Morris (1974)).

For the input data type, we claim that credences are generally better than binary beliefs. We can treat the input data as evidence for the resulting collective belief, and we expect that sophisticated and informational input data are more likely to track the truth. This view supports that, if possible, probabilistic beliefs should be provided as the input data type. In this sense, the input data in cases (2) and (3) already include information loss.

For the output data type, we claim that there are some cases where binary beliefs are appropriate. Although the following arguments supported by examples are not rigorous and include exceptions, this will be enough to support our claim. First, the output data type depends on the extent that the group can be seen as a genuine group agent. Individual beliefs can be gathered without supposing a group agent to which the aggregate beliefs belong. For example, consider a voting result of South Korean women in their 30s in the last presidential election. Although some meaningful patterns and information can be extracted from the statistical data, it is difficult to say that the set of all South Korean women in their 30s is a genuine group agent.[9] For mere summaries of individual beliefs, credences would be a more appropriate output data type rather than binary beliefs. A more complex, structured, and organized collection of individuals can be regarded as a group agent. Indeed, some groups like political parties and companies seem to have collective beliefs as well as consistent plans and goals; they respect some epistemic and practical rationality norms much like an individual agent.[10] In many cases of belief aggregation, the resulting collective beliefs are shared with other members of society, and evaluated by the society's dominant rationality norms. For example, political parties describe society's problems through their program, and if the program makes contradictory claims, the party will be criticized. And in all of these processes, we see collective binary beliefs, and some rationality norms on them, such as consistency and closure under conjunction. This is not surprising since binary beliefs are human-friendly and appropriate for conveying information efficiently and for communicating with other individuals.

Second, the output data type also pertains to the purpose of belief aggregation. Indeed, group beliefs should take different types depending on whether the collective beliefs are provided as a basis for some decision making or released to the public through announcements. For example, consider the case where a working group of various scientists in IPCC[11] publishes an assessment report for policymakers. Since this report is used as the background information for policymaking, probabilistic beliefs may be appropriate. In contrast, binary collective beliefs would be appropriate when a government announces a final position after a discussion based on the report and communicates it to the public. Otherwise, the government's public announcement would not be effectively passed on to the public.

Third, the types of belief in the final judgment are pre-determined in many insti-

[9] List (2014) distinguished three kinds of collective beliefs: *aggregate beliefs*, which are mere summaries of individuals' beliefs; *common beliefs*, which each entertains with common awareness; *corporate beliefs*, which can be thought of as a genuine group's belief.

[10] In this thesis, we will not discuss further whether/when a group agent exists. For a more discussion about the *group agency* problem, see List & Pettit (2011).

[11] Intergovernmental Panel on Climate Change

tutions, whether written or customary. For example, the jury's verdict in a criminal case is supposed to be expressed as a binary belief: "guilty" or "not guilty".

For this reason, it is worth studying case (4), and we call it *heterogeneous belief aggregation*, which is the topic of this thesis. In Figure 1.2, heterogeneous belief aggregation is illustrated in relation to other research fields. In the figure, binarization indicates rational bridge principles between credences and binary beliefs, which can be used as a function or a correspondence taking credence and giving binary beliefs. It is noticeable that individual belief binarization and collective belief binarization can be distinguished[12], which can be combined with judgment aggregation and opinion pooling methods respectively to devise a heterogeneous belief aggregation method. This thesis focuses on collective belief binarization rather than individual belief binarization since the heterogeneous belief aggregation methods employing the latter — judgment aggregation after individual belief binarization — would result in information loss relatively prematurely.

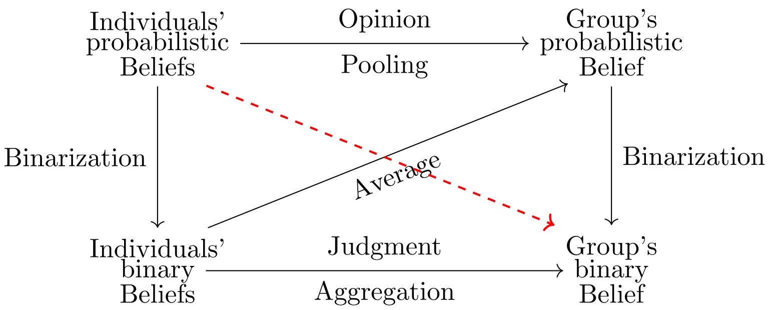


Figure 1.2: Heterogeneous Belief Aggregation is depicted by the red dotted line.

In the remainder of this chapter, we will illustrate heterogeneous belief aggregation in connection with three related research fields: judgment aggregation, probabilistic opinion pooling, and belief binarization. The first two are similar to heterogeneous belief aggregation because they are about aggregating beliefs. The latter addresses two heterogeneous types of belief, which can be used for collective beliefs as well, and thus combined with an opinion pooling procedure for heterogeneous belief aggregation. Our theory of heterogeneous belief aggregation will be developed based on many achievements and obstacles in these research areas.

[12] One may wonder if the relationship between group credences and group binary beliefs is fraught with metaphysical problems. In our opinion, collective belief binarization is rather relatively free from the metaphysical aspect. In general, an artifact does not need to presuppose human psychology. For example, it is not a metaphysically serious problem whether a knowledge base stored in computer memory is written in the symbolic or probabilistic language. However, it may be necessary for social and political institutions to consider and use human psychology since group agents, albeit an artifact, have social relationships with other individuals. Nevertheless, this is not a metaphysical burden; it serves only a practical purpose. For example, internal reports, meetings minutes, and press conferences constitute the corpus of the group's beliefs, and the type of belief depends on the purpose for which those beliefs were formed and the rationality norms that those beliefs are supposed to respect.

We begin by introducing a typical example of belief aggregation or belief binarization problems:

> *A political party wants to establish its position on the basic income policy. To this end, the party asks some experts for opinions on how artificial intelligence will affect the labor market in the future and collects the results in order to use as an argument for the party's position. The experts present their beliefs on the following logically interconnected issues.*

A : *"Artificial intelligence will outperform humans in all areas by 2050."*

B : *"Artificial intelligence will replace humans in the labor market."*

$A \to B$: *"If artificial intelligence outperforms humans in all areas by 2050, then it will replace humans in the labor market."*

Judgment Aggregation Let us assume that the following profile of three experts' beliefs is given. Here, every belief is complete (every agent holds each issue to be true or false, and suspending is not allowed). The collective belief is formed by majority voting for each issue.

Issues	A	$A \to B$	B
Agent 1	T	T	T
Agent 2	T	F	F
Agent 3	F	T	F
Collective Belief	T	T	F

Table 1.1: The Discursive Dilemma

It shows that the issue-wise majority rule on logically interconnected issues might generate an inconsistent collective belief, even though every individual has a consistent belief. This is called the *discursive dilemma*. Motivated by this, the judgment aggregation theory was founded and has been extensively studied.[13] The problem can be generalized beyond majority voting, as shown in List & Pettit (2002) — there is no anonymous, neutral, and independent procedure if the issues have some minimal logical interconnections. Various kinds of impossibility theorems, on the one hand, generalize this kind of dilemma, saying that seemingly reasonable properties and conditions cannot be jointly satisfied, like Arrow's impossibility theorem in social choice theory. On the other hand, there have been different suggestions to avoid impossibility theorems. To formulate impossibilities and find escape routes, we need to understand why this dilemma arises. It is a common view that the main source of the dilemma is the conflict between the issue-wise independence norm (a collective belief should be formed issue by issue), the logical interconnections between the issues, and rationality requirements such as consistency (a binary belief should not

[13]For a survey, see List & Puppe (2009), Mongin (2012), and for a textbook, see Grossi & Pigozzi (2014).

entail a contradiction), deductive closure (a binary belief should contain its logical consequences) or completeness. Let us mention each in turn:

One of the main research directions in judgment aggregation theory is to find the exact agenda conditions under which different (im)possibility results arise. First, Nehring & Puppe (2010) proved that *path-connectedness*[14] is the exact agenda condition under which any independent judgment aggregator satisfying completeness and monotonicity leads to a dictatorship. Next, Dokow & Holzman (2010a) proved that *path-connectedness* and *even-negatability*[15] is the agenda condition under which any independent judgment aggregator satisfying completeness leads to a dictatorship.

One seemingly natural answer to the Discursive Dilemma is to weaken rationality requirements, in particular the completeness condition.[16] This route can be well supported because *epistemic logic* and *AGM belief revision theory* — the most successful models of qualitative beliefs — do not demand the completeness condition. However, Gärdenfors (2006) showed that relaxing completeness also results in a degenerate aggregation rule, namely *oligarchies*.[17] In contrast to the completeness condition, Briggs et al. (2014) proved that the discursive dilemma could be avoided by weakening the consistency condition to probabilistic coherence.

The most often mentioned way to avoid the discursive dilemma is to relax the independence norm[18]. New kinds of *holistic judgment aggregation methods* have been suggested, e.g., premise-based procedures (Pettit (2001), Dietrich & Monin (2010)); distance-based rules (Miller & Osherson (2008)); sequential rules (List (2004)).

Probabilistic Opinion Pooling Probabilistic opinion pooling is a research field devoted to the question of how to aggregate the individual probabilities into a collective probability.[19] Let us illustrate the opinion pooling problem with an example: This time, three experts present their credences on the same issues as before. The collective belief is formed by issue-wise averaging of the individual credences. It is noticeable that although the issues do not constitute an algebra, each agent's numerical beliefs can be extended to a (finitely-additive) probability measure on the algebra generated by the issues, and the issue-wise averaging procedure guarantees that the group's numerical belief also can be extended to a probability measure.

The main finding in opinion pooling is the characterization theorem of *linear*

[14]Path-connected is also called *total blockedness*.

[15]Even-negatability is also called *non-affineness* .

[16]This route is important to this thesis because we do not require the completeness condition except for some parts in Chapter 2.

[17]Gärdenfors (2006) assumed a demanding condition that the agenda set should be an atomless Boolean algebra. Dietrich & List (2008) and Dokow & Holzman (2010a) proved that given an (in)finite agenda set — Dietrich & List (2008) deals with infinite and finite agenda sets at the same time whereas Dokow & Holzman (2010a) assumes only the finite case —, relaxing the completeness condition on the (input and) output yields oligarchies, which is, strictly speaking, a stronger notion than those in Gärdenfors (2006). Furthermore, they found the exact agenda condition — path-connectedness and even-negatability — under which the oligarchy result arises.

[18]For relaxing the requirement of *universal domain* see List (2003).

[19]For a survey see Genest & Zidek (1986), Dietrich & List (2016)

Issues	A	$A \to B$	B
Agent 1	0.9	0.7	0.6
Agent 2	0.8	0.4	0.2
Agent 3	0.4	0.7	0.1
Collective Belief	0.7	0.6	0.3

Table 1.2: Linear Pooling

pooling.[20] McConway (1981) proved that given a σ-algebra, the only opinion pooling function that is independent and preserves zero probabilities is linear pooling (weighted averaging by issues). Dietrich & List (2017a, 2017b) generalized this result to general agendas that need not be a σ-algebra. However, linear pooling does not satisfy all the requirements of probabilism or Bayesianism.[21] It is well-known fact that the linear pooling procedure creates a tension with *Bayesian conditionalization* because Bayesian conditionalization after linear pooling may not coincide with linear pooling after Bayesian conditionalization in individual beliefs, which is said to violate the *external Bayesianity.*[22] *Geometric pooling* has an advantage in this respect. Genest (1984) and Genest et al. (1986) proved that geometric pooling is the only opinion pooling function satisfying the external Bayesianity under some minor conditions.

Belief Binarization Belief binarization is a study concerning rational bridge principles between credences and binary beliefs. In most of the literature, belief is associated with high credence, and thus principles based on some kinds of thresholds have been considered natural. One of them is the well-known *Lockean Thesis*, which states that an agent should believe an issue/a proposition iff its probability exceeds a threshold. However, the *lottery paradox* shows that the Lockean thesis might result in a contradictory belief:

> *Consider a fair 1,000-ticket lottery that has only one winning ticket. An agent believes that one ticket will win. She also believes that each ticket*

[20]It does not mean that probabilistic opinion pooling has only positive results. For example, Lehrer & Wagner (1983) showed that the only probabilistic opinion pooling functions satisfying independence, zero probability preservation, and preservation of probabilistic independence are dictatorships.

[21]Probabilism usually indicates that a rational agent ought to use probability functions and Bayesian conditionalization. In contrast, Bayesian rationality is used with various meanings. For example, Bayesian rationality may mean that rational agents ought be an expected utility maximizer who has preferences and beliefs simultaneously. The *Bayesian social aggregation* views a rational group agent as an expected utility maximizer. However, in this research field, collective beliefs are studied only concerning aggregation of the preferences of the individuals. Although this problem has been extensively studied and is essential in social ethics, we will study the belief aggregation problem independently of the preference aggregation problem. For a survey about Bayesian social aggregation, see Mongin & Pivato (2016).

[22]This does not mean that linear pooling does not fit well with every update method. For example, according to Leitgeb (2017a), combining linear pooling with *probabilistic imaging* can solve the commutativity problem.

will lose since her probabilistic belief is a uniform distribution over the tickets, and her threshold is 0.99. Then it is deduced that no ticket will win, which leads to a contradictory belief.

Note that unless the threshold is 1, one can find examples that shows that the Lockean thesis does not guarantee rational — here, consistent and deductively closed — binary beliefs. There have been many suggestions to resolve this paradox. Some argued for relaxing closure under conjunction (Kyburg (1961), Leitgeb (2021)) or weakening consistency to probabilistic coherence (Easwaran & Fitelson (2015)). Some suggested giving up probabilism and using alternative models of credences such as ranking functions (Spohn (2009)). Some proposed new bridge norms relaxing the Lockean thesis. The *stability theory of belief* (Leitgeb (2017a)) provides a rational way to relate credences and binary beliefs: they should satisfy the *Humean thesis*, which says that an issue should be believed iff its probability remains above a given threshold under conditionalization on any issue whose negation is not believed. The theory proves that this amounts to choosing a threshold for a given probability measure so as to preserve rationality. The *camera shutter rule* (Lin & Kelly (2012)) recommends a different way to preserve rationality and probabilism. According to the rule, a belief core whose supersets are believed should consist of the worlds whose probability ratio to the maximal probability is above a given threshold.

The conflict between rationality and the Lockean thesis in the lottery paradox is closely related to the one between rationality and majority voting in the discursive dilemma. The reason is that the threshold rule of the Lockean thesis can be viewed as a quota rule that generalizes majority voting. Let us return to the experts' beliefs in Table 1.1 to see this. This time, the party forms a credence based on the data and then reduces the credence to a binary belief by applying a threshold rule with threshold 0.6 — an issue is believed iff its probability is above 0.6 — as follows.

Issues	A	$A \to B$	B
Credence	0.66	0.66	0.33
Binary belief	Belief	Belief	Disbelief

Table 1.3: The Lottery Paradox

We see that the resulting belief satisfies neither deductive closure nor consistency. Indeed, any anonymous independent judgment aggregation can be seen as a belief binarization problem. The similarity and relation between the judgment aggregation problems and belief binarization problems have firstly been observed by Douven & Romeijn (2007). Dietrich & List (2018, 2021) showed how to relate the impossibility theorems of judgment aggregation and the lottery paradox and found the agenda conditions under which the impossibility results arise in the binarization problems.

Heterogeneous Belief Aggregation We propose a new research topic: *heterogeneous belief aggregation*. It is an aggregation of credences into a binary belief. On the one hand, this topic differs from judgment aggregation and opinion pooling in that the input and output data types are heterogeneous, just as in belief binarization. On the other hand, the topic is distinguished from belief binarization in that multiple credences are taken as input.

Let us illustrate the heterogeneous belief aggregation problem with an example. Three experts' opinions are given in the form of credences as in Table 1.2. Suppose that the party's belief is generated by the following method: an issue is collectively believed iff its average credence is above 0.6.

Issues	A	$A \to B$	B
Agent 1	0.9	0.7	0.6
Agent 2	0.8	0.4	0.2
Agent 3	0.4	0.7	0.1
Collective Belief	Belief	Belief	Disbelief

Table 1.4: A Heterogeneous Belief Aggregation Problem

This new dilemma shows that this aggregation procedure does not ensure deductive closure and consistency, just as in the discursive dilemma and the lottery paradox. Confronting this situation, we have two main questions: how to formulate and generalize the dilemma, and what kinds of aggregation procedures can avoid the dilemma and obtain rational collective beliefs.

To answer the first question, we will employ the axiomatic approach to deal with general aggregation procedures as in judgment aggregation and social choice theory. We will investigate which kinds of individual and collective rationality requirements and which properties of aggregation procedures should be imposed on heterogeneous belief aggregation, and which of their combinations are impossible. We will mainly assume deductive closure rather than completeness, in contrast with most of the judgment aggregation literature. Moreover, we will address impossibility results without anonymity conditions, which cannot be considered in belief binarization. This will lead to three kinds of impossibility results, and we will also determine the sufficient and necessary agenda condition for each of the results.

For the second question, we will explore specific heterogeneous belief aggregation procedures and their properties. There can be two kinds of heterogeneous belief aggregation procedures: (1) collective belief binarization combined with an opinion pooling method and (2) direct rules, as illustrated in Figure 1.3.

As for collective belief binarization, theories of belief binarization are applicable. To this end, we will first analyze the existing threshold-based procedures, especially those that relax the Lockean thesis and preserve rationality. We will categorize them as local-threshold rules — where thresholds depend on probability measures — and world-threshold rules — where thresholds are applied not to an issue but to a possible world. Their characteristics will be captured by the property of local monotonicity and world monotonicity, respectively. We will compare and relate these properties with other existing properties like being stable in the stability theory of belief and with new — to be introduced — properties. Whether some existing rational procedures, like the camera shutter rule, satisfy these properties will be an interesting

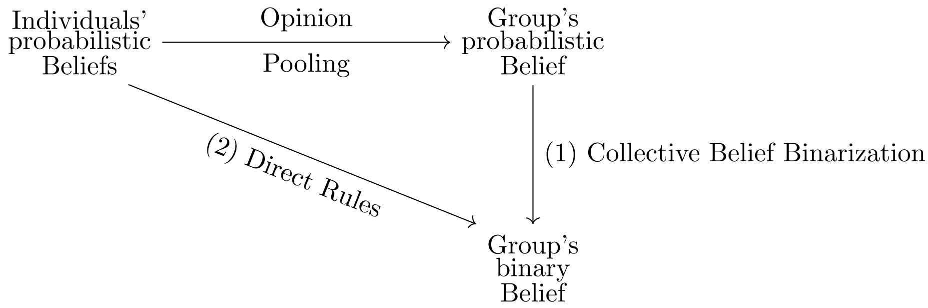


Figure 1.3: Collective Belief Binarization and Direct Rules

and philosophically important question. We will provide geometrical characterizations of some of the properties to answer this question. Furthermore, we will propose that convexity norms should be discussed in the context of belief binarization. We will introduce various kinds of convexity norms and examine whether the relevant procedures satisfy them.

What is more, we will propose two novel kinds of belief binarization methods that preserve rationality but are not based on thresholds: *distance-based binarization* and *epistemic-utility-based binarization*. The first is a holistic one minimizing the distance from a given probability measure to the resulting binary belief. The second one is based on an accuracy norm minimizing expected inaccuracy. We will devise novel ways to measure the required distances and inaccuracies. Moreover, we will study distance minimization with *Bregman divergence*, utility maximization with *strict proper scores*, and their relation.

Direct heterogeneous belief aggregation rules will also be proposed and studied regarding threshold, distance and epistemic-utility. We will provide a new classification and characterization of them. Furthermore, we investigate some norms that are especially relevant in social contexts, such as various unanimity norms, convexity norms interpreted in social contexts, and commutativity norms, which govern the relationship between direct rules and combinations of opinion pooling and collective belief binarization.

These novel problems we are proposing here have not been discussed in social epistemology and formal epistemology thus far. Compared to judgment aggregation and opinion pooling, it has been rarely the case that aggregation problems dealing with heterogeneous belief models have been discussed. As far as we know, the only literature to study the aggregation of credences into a qualitative belief is Ivanovska & Slavkovik (2019). However, they mainly focused on procedures where individual credences are first transformed into qualitative beliefs, to which judgment aggregation methods are then applied. Thorn (2018) also dealt with both probabilities and binary beliefs but investigated the joint aggregation of individual belief states, each of which consists of a quantitative and a qualitative belief, into a collective belief state. Our problem is different because we do not deal with individual belief binarization due to the information loss problem. Instead, we focus on procedures where firstly, the group's credence is formed via an opinion pooling method, after which we apply threshold- or distance-based belief binarization methods to the group's belief state, which include methods from Leitgeb (2014a), Cantwell & Rott (2019) and Lin & Kelly (2012b). We can find literature where belief binarization methods are applied to aggregation problems, e.g., Chandler (2013) and Cariani (2016). However, they used the methods for judgment aggregation. We will apply various belief binarization methods to aggregate credences to a belief.

Last but not least, let us return to the example where the party writes the program including the basic income policy based on some experts' opinions about the effect of artificial intelligence on the labor market. This example satisfies every reason to use heterogeneous belief aggregation procedures. When the experts' probabilistic opinions are given, the program will be better supported, and when binary beliefs are expressed in the program, the party can more effectively convey their beliefs, which will guide the party's and the supporters' reasonable action. Besides this typical example, there can also be broader applications of heterogeneous belief aggregation theory. On the one hand, the theory can be applied to belief binarization of an imprecise probability represented by a finite set of probabilistic beliefs. This shows that heterogeneous belief aggregation can be viewed as a generalization of belief binarization. On the other hand, the theory can also be used for a judgment aggregation problem given subgroups' credences calculated anonymously and independently from the binary beliefs of the members of the subgroups.[23] This indicates that heterogeneous belief aggregation can be interpreted as a generalization of judgment aggregation.

Putting all this together, we conclude that heterogeneous belief aggregation is a philosophically fruitful topic that deserves attention. Heterogeneous belief aggregation can be seen as a general framework, where not only heterogeneous belief aggregation but also opinion pooling, judgment aggregation, and belief binarization are studied in connection to each other. First, studying heterogeneous belief aggregation is by itself interesting and cannot be reduced to other research fields: we can deal with different rationality norms in social contexts and address properties characteristic for heterogeneous belief aggregation. Moreover, it is not only the direct rules but also the different possible combinations of methods from different research areas that makes this whole endeavor to be more than sum of its parts. Indeed, second, this framework bridges independently developed research areas: first, we can apply well-developed formal theories in formal epistemology like belief binarization theories and epistemic decision theories to the belief aggregation problem. Second, this framework enables us to add social contexts to belief binarization problems and epistemic decision theories, which can be extended to cover also social beliefs. Our theory of heterogeneous belief aggregation can be applied to the (collective) belief binarization problem and epistemic (collective) decision theory. In this way, the thesis fills, or at least narrows, the gap between individual epistemology and collective epistemology.

The thesis is organized as follows: In Chapter 2, we define basic properties of heterogeneous belief aggregation and agenda conditions, and prove three impossibility

[23]For example, consider a situation where the political party requested separate opinions of two subgroups, e.g., machine learning programmers and labor economists, on the issues of the previous examples, and the opinions of each subgroup were collected through an anonymous and independent judgment aggregation. This method is appropriate when we want to collect the opinions of the subgroups rather than the opinions of each individual.

results. In Chapter 3, threshold rules are classified and characterized. We also analyze some properties and rules of belief binarization. Convexity norms are addressed as well. In Chapter 4, we propose new binarization procedures minimizing distance and maximizing expected utility, and study their relationships. Moreover, we suggest some properties characteristic for heterogeneous belief aggregation, and investigate the relationships between direct rules and the combination of opinion pooling and belief binarization. Lastly, Chapter 5 concludes the thesis and suggests further research. All of the chapters have one theme and are intimately connected to one another, but each chapter is mostly self-contained and thus can be read independently.

Chapter 2

Triviality Results about Heterogeneous Belief Aggregation

In this chapter, following the old research tradition in social choice theory, we will suggest some rational requirements on heterogeneous belief aggregation and prove several impossibility results. Furthermore, following the relatively recent research tradition in judgment aggregation theory, we will present characterization theorems by which we can determine the exact agenda conditions under which the impossibility results arise.

2.1 Introduction

This chapter aims to formally define heterogeneous belief aggregation and its properties, and investigate which combinations of properties can be satisfied or not. As illustrated in the discursive dilemma and lottery paradox introduced in the last chapter, seemingly desirable or natural properties might not be be satisfied simultaneously. In heterogeneous belief aggregation, we might run into similar difficulties. Thus, to begin our study, we formalize and prove exactly when impossibilities arise, which will provide the exact boundary between possible and impossible heterogeneous belief aggregation.

Heterogeneous belief aggregation is a new research area we are proposing in this study, and, of course, no impossibilities have been established thus far. However, as indicated in the last chapter, heterogeneous belief aggregation is similar to judgment aggregation and opinion pooling since they deal with aggregating individual beliefs into a collective belief. It differs from them in that the input and output types are different, i.e., heterogeneous, which is a commonality with belief binarization. Therefore, reviewing impossibilities in judgment aggregation, opinion pooling, and belief binarization will give some hints for developing the theory of heterogeneous belief aggregation.

In the center of many impossibility theorems in the three research areas lies the norm of independence, which says that procedures to obtain the resulting belief on an issue in the agenda should depend only on the inputs on the issue, not on other issues. This means that the output should be determined issue-wise. However, when issues in the agenda are logically interconnected, and the resulting belief is required to satisfy some rationality norms like deductive closure, consistency, or completeness, issue-wise procedures might violate the rationality norms to respect logical interconnections between issues.

Starting by generalizing the problem of issue-wise majority voting in the discursive dilemma, the axiomatic method in social choice theory, like the method employed to obtain Arrow's impossibility theorem, has been applied to any agenda of which issues are logically represented and interconnected (List & Pettit (2002)). Utilizing this method, much research in judgment aggregation has formalized and generalized the tension mentioned above. For example, it was shown that independent aggregation to generate complete and consistent collective judgments on a given agenda is, under certain minimal conditions, forced to be a dictatorship — the collective judgment is always the same as a fixed individual's judgment — if and only if the issues in the agenda have certain logical interconnection — which is called path-connectedness and even-negatability (Dokow & Holzman (2010a)). If axiomatic requirements on the aggregation are strengthened, we obtain impossibilities more easily; hence the agenda condition can be weakened. For example, if we add anonymity, then the agenda condition for the impossibility can be weakened to what is called blocked (Nehring & Puppe (2010)). While completeness and consistency are usually assumed in judgment aggregation, a few studies have weakened them to deductive closure, where dictatorships are replaced with oligarchies — an issue is collectively accepted if and only if it is unanimously accepted by a fixed subgroup of individuals (oligarchs) (Gärdenfors (2006), Dietrich & List (2008)).

In opinion pooling and belief binarization, the agenda has usually been assumed to have the structure of a (non-trivial) algebra, because they deal with probability measures. Under this assumption, McConway (1981) showed that independent opinion pooling with a certain minor condition (certainty preservation) is restricted to linear pooling. Recently, Dietrich & List (2017a, 2017b) relaxed the assumption of an algebra and investigated general agendas to characterize linear pooling.

Belief binarization is not a problem of aggregating beliefs, but this can be thought of as anonymous issue-wise judgment aggregation (Dietrich & List (2018)). Therefore, there are also impossibilities demonstrating the tension between independence and rationality norms, such as deductive closure, when the agenda has the structure of an algebra, which is common in the belief binarization literature, but can also be relaxed (Dietrich & List (2018, 2021)). It is worth pointing out that deductive closure is the typically required rationality norm in belief binarization as well as in doxastic logic and AGM belief revision theory.

The aforementioned impossibilities in the three research areas shed some light on developing the theory about impossibilities on heterogeneous belief aggregation. Accordingly, one may conjecture that if we combine the strong properties — completeness of collective beliefs and anonymity — with the strong agenda condition of being an algebra, then we can easily obtain an impossibility of independent heterogeneous belief aggregation. The reason is that stronger conditions are more likely to be incomparable with independence; thus we can obtain impossibility results more easily. Our research will make stronger claims. Hence, we investigate not only the impossibility with the above combination but also other impossibilities in the following way:

(i) we relax completeness that is usually required in the theory of judgment aggregation, and assume deductive closure as in belief binarization,

(ii) we drop anonymity in anonymous issue-wise judgment aggregation, which corresponds to the belief binarization problem, and attach social contexts — by taking multiple individuals' beliefs as input — and epistemic contexts — by respecting degrees of expertise — as in judgment aggregation or opinion pooling, and

(iii) we weaken the structure of an algebra commonly required in opinion pooling and belief binarization, and characterize impossibility agendas as in judgment aggregation.

Using this approach, we can contribute to sparking interest in heterogeneous belief aggregation. On the one hand, it has its own interest in the following sense. First, we can deal with different properties: individual rationality should be about probabilistic beliefs, unlike in judgment aggregation, and collective rationality should be about binary beliefs, unlike in opinion pooling. Also, unlike in belief binarization, we should consider rationality norms in a social context. Second, direct procedures might exist that cannot be reduced to a combination of opinion pooling and belief binarization or a combination of belief binarization and judgment aggregation. Through our approach, we can suggest new rationality norms and find impossibilities characteristic for heterogeneous belief aggregation. On the other hand, our approach can provide a general framework to bring together and compare different research areas dealing with binary or probabilistic beliefs, or the aggregation of beliefs. This general framework will enable one to apply our theory of heterogeneous belief aggregation to other research areas and shed new light on them.

Considering the above, our main questions in this chapter are the following:

(Q1) what kinds of impossibility results can we formulate and prove? First, we will determine exact combinations of properties of heterogeneous belief aggregation to yield impossibilities, assuming enough logical interconnection between issues like a (non-trivial) algebra. Our results will involve the cases with deductive closure or completeness, and with or without anonymity;

(Q2) how rich should the logical interconnections be to obtain or avoid each impossibility result? Put differently, what is the necessary and sufficient agenda condition for each impossibility result established above?;

(Q3) how are our theorems and proofs compared with impossibility results in judgment aggregation, opinion pooling and belief binarization? We will relate our results to similar ones in other research areas and analyze similarities and differences between our proofs and other ones.

The remainder of this chapter is organized to answer these questions as follows. First, in Section 2.2, we illustrate our setting and formally define heterogeneous belief aggregation. Next, in Section 2.3, we introduce and formulate some (possible) axiomatic requirements on heterogeneous belief aggregation. In Section 2.4, we present and prove our first main results under the assumption of the agenda being a non-trivial algebra, according to which there is no heterogeneous belief aggregation rule satisfying some presumably rational requirements except for degenerate ones. Then, in Section 2.5, we compare our main results with results in other research areas. Lastly, in Section 2.6, which is based on joint work with Chisu Kim, we present our second main results: three characterization theorems of impossibility agendas.

2.2 Heterogeneous Belief Aggregation

We begin by introducing some notation and terminology that will be maintained throughout this chapter, as well as the formal definition of heterogeneous belief aggregation.

Let W denote a non-empty set that represents a set of worlds. An *agenda* $\mathcal{A}$ on W is a non-empty set of some subsets of W that is closed under complement, that is, for all $A \in \mathcal{A}$, it holds that if $A \in \mathcal{A}$, then $\overline{A} \in \mathcal{A}$ where $\overline{A}$ is the complement of A. We call an element A of $\mathcal{A}$ an *issue*. We denote the set of n individuals by $N := \{1, ..., n\}$ and assume that $n \geq 2$. For each $i \in N$, P_i denotes an individual i's *belief* that is a function from $\mathcal{A}$ to $[0, 1]$ and $\vec{P} := (P_1, ..., P_n) =: (P_i)_i$ denotes a profile of n individuals' beliefs. We call a function F taking $\vec{P}$ into $F(\vec{P}) : \mathcal{A} \to [0, 1]$ an *aggregator*, where $F(\vec{P})$ represents a collective belief. For any individual or collective belief $P : \mathcal{A} \to [0, 1]$, if the codomain is restricted to $\{0, 1\}$, then it is called a *binary belief*, and for any issue $A \in \mathcal{A}$, $P(A) = 1$ means that A is believed and $P(A) = 0$ means that A is not believed. We call the set of all believed issues — i.e., $\{A \in \mathcal{A} | \ P(A) = 1\}$ — the *belief set* of P and denote it by $P^{-1}(1)$. In contrast to binary beliefs, $P : \mathcal{A} \to [0, 1]$ is called a *probabilistically coherent belief*, or simply, a *probabilistic belief*, if P is extendable to a finitely-additive probability on the algebra[1] generated by $\mathcal{A}$(i.e., the smallest algebra that includes $\mathcal{A}$). Even though we adopt a somewhat generalized definition of probabilistic beliefs, finitely-additive probabilities on an algebra $\mathcal{A}$ are thought of as the most basic and typical probabilistic beliefs throughout this chapter, except in Section 2.6.

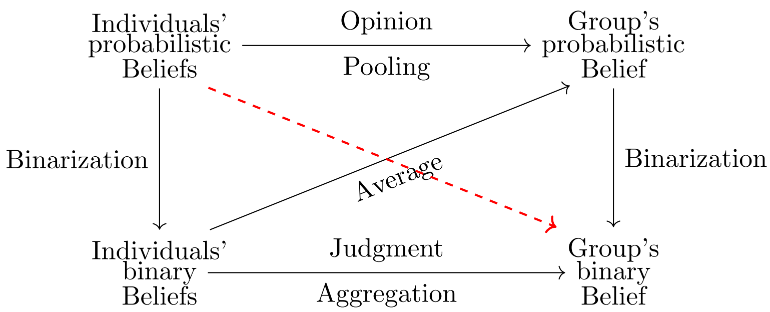


Figure 2.1: Heterogeneous belief aggregation is depicted by the red dotted line.

Heterogeneous belief aggregation deals with individuals' probabilistic beliefs and the group's binary belief (See the figure above). Formally, it is defined as follows:

Definition 2.1 (Heterogeneous Aggregator (HA)). *A heterogeneous aggregator (HA) F is a function that takes each profile $\vec{P}$ of n probabilistic beliefs on $\mathcal{A}$ in a given domain and returns a binary belief, which is a function $F(\vec{P}) : \mathcal{A} \to \{0, 1\}$.*

[1]An algebra $\mathcal{A}$ on a set W is a set of subsets of W such that (1) $W \in \mathcal{A}$, (2) if $A \in \mathcal{A}$, then $\overline{A} \in \mathcal{A}$, and (3) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$. A finitely-additive probability is a function P from an algebra $\mathcal{A}$ to $[0, 1]$ such that (1) $P(W) = 1$, and (2) if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

Studying heterogeneous belief aggregation, we will also compare this with opinion pooling, judgment aggregation, and belief binarization problems. For a start, compare the above definition with the following formal definition of an opinion pooling function and that of a judgment aggregator: A *judgment aggregator (JA)* is a function that takes each profile $\vec{P}$ of n binary beliefs on $\mathcal{A}$ in a given domain and returns a binary belief on it. To define an opinion pooling function we assume that $\mathcal{A}$ is an algebra.[2] An *opinion pooling function (OP)* is a function that takes each profile $\vec{P}$ of n finitely-additive probabilities on the algebra $\mathcal{A}$ in a given domain and returns a finitely-additive probability on it.

[2]In most of the opinion pooling literature, the underlying agenda is assumed to be a σ-algebra. However, whenever we talk about opinion pooling, we assume that $\mathcal{A}$ is an algebra, since we do not need σ-additivity to prove the relation between certainty preservation, independence, and systematicity in opinion pooling, which we will later compare with our result (Lemma 2.4) regarding heterogeneous belief aggregation. Note that exceptionally, in Dietrich & List (2017a, 2017b), they explore generalized opinion pooling where $\mathcal{A}$ need not be even an algebra but just closed under complement. In our research, we investigate heterogeneous belief aggregation where the definition of an agenda is the same as the one in Dietrich & List (2017a, 2017b).

2.3 The Properties of Heterogeneous Aggregators

In this section, we gather axiomatic requirements on heterogeneous belief aggregation. A HA consists of inputs, outputs, and a rule, and accordingly, the requirements can be divided into two groups. The first group concerns the inputs and outputs and includes rationality norms that we impose on individual and collective beliefs. The second group is related to conditions imposed on the rule.

Individual and Collective Rationality We focus on obtaining a rational collective belief given rational individual beliefs. The kinds of rational requirements that we should demand can be different according to whether they are imposed on probabilistic beliefs or on binary beliefs. For probabilistic beliefs, we require the rationality condition that says that they should be extendable to finitely-additive probabilities, which is already satisfied by the definition of probabilistic beliefs. For binary beliefs, rational requirements are related to consistency, deductive closure, and completeness. To define them, we need to first define an entailment relation in the agenda. Even though the agenda in our setting does not need to be finite or generated from valuations in some standard propositional logic, we want it to fit well with compact logic, which can be thought of as a standard model of 'logic'. Thus the definition of an entailment relation is designed to include the notion of compactness.

Definition 2.2 (Entailment)**.** *Let $\mathcal{A}$ be an agenda on W, $\mathcal{Y} \subseteq \mathcal{A}$ and $A \in \mathcal{A}$. $\mathcal{Y}$ entails A($\mathcal{Y} \vDash A$) iff there is a finite subset $\mathcal{X} \subseteq \mathcal{Y}$ such that $\bigcap \mathcal{X} \subseteq A$.*

In this definition, we do not exclude $\mathcal{X} = \emptyset$ and adopt the convention that $\bigcap \emptyset = W$. Now let us define three rationality norms imposed on binary beliefs.

Definition 2.3 (Rationality of Binary Belief)**.** *Let $Bel : \mathcal{A} \to \{0, 1\}$ represent a binary belief. (Recall that the belief set of Bel is defined by $Bel^{-1}(1) := \{A \in \mathcal{A} | Bel(A) = 1\}$.)*

(1) Bel is complete iff $Bel(A) = 1$ or $Bel(\overline{A}) = 1$ for all $A \in \mathcal{A}$.

(2) Bel is consistent iff $Bel^{-1}(1) \nvDash \emptyset$.

(3) Bel is deductively closed iff for all $A \in \mathcal{A}$ such that $Bel^{-1}(1) \vDash A$, $Bel(A) = 1$.

Completeness means that for every issue, the issue is believed or its negation is believed. It is quite demanding in the sense that suspension on an issue — believing neither the issue nor its negation — is not allowed.

Consistency means that the belief set should not entail a contradiction. It says, in accordance with the definition of entailment, that every finite intersection of the belief set $Bel^{-1}(1)$ should be non-empty. It is weaker than the requirement that $\bigcap Bel^{-1}(1) \neq \emptyset$, when $\mathcal{A}$ is infinite. There is another seemingly alternative definition of consistency: a contradiction should not be believed — i.e., $Bel(\emptyset) = 0$. Under the assumption of $\emptyset \in \mathcal{A}$, it holds that our consistency of a binary belief Bel implies $Bel(\emptyset) = 0$. The opposite direction holds under the additional assumption that Bel is deductively closed. Since we do not always require that $\emptyset \in \mathcal{A}$, we will use the definition of part (2). In summary, our definition of consistency is weaker than the consistency allowing infinitary reasoning and stronger than not believing a contradiction.

Now consider the definition of deductive closure. If the agenda is an algebra, then deductive closure of a binary belief means that its belief set contains W, and it is closed under finite intersections and supersets. If we do not require that $\mathcal{A}$ is an algebra, then, part (3) says that $Bel^{-1}(1)$ contains W if $W \in \mathcal{A}$, and it is closed under finite intersections and their supersets that are contained in the agenda $\mathcal{A}$. Notice that even if the finite intersection of some sets in the belief set is not contained in the agenda and thus not in the belief set, a superset of it in the agenda should be contained in the belief set to satisfy deductive closure.

Now we introduce two kinds of rationality of binary beliefs. When a binary belief should be consistent and deductively closed, we call this requirement *rationality*, and when a belief should be consistent and complete, we call it *complete-rationality*. Some of our results require rationality and some others require complete-rationality. Let us compare rationality and complete-rationality. First, a binary belief with complete-rationality can be viewed as a probabilistic belief of which values are only in $\{0, 1\}$. However, this does not hold for binary beliefs with rationality. Second, it is easy to see that 'complete-rationality' is literally stronger than 'rationality', as shown in the following.

Lemma 2.1. *Deductive closure follows from consistency and completeness.*

Proof. Suppose, contrary to deductive closure, that $Bel^{-1}(1) \vDash B$ but $Bel(B) = 0$. By the completeness of Bel, we have $Bel(\overline{B}) = 1$. By the definition of entailment, there is a finite subset $\mathcal{B} \subseteq Bel^{-1}(1)$ such that $\bigcap \mathcal{B} \subseteq B$, and thus $\bigcap \mathcal{B} \cap \overline{B} = \emptyset$, which contradicts that Bel is consistent. $\square$

Now, we are ready to state properties of HA that are related to individual and collective rationality. First of all, the requirement of individual rationality is related to the domain of an aggregation function. It can be alluded to by the requirement of universal domain, which says that the domain should include all profiles of rational individual beliefs.

Definition 2.4 (Universal Domain (UD))**.** *A HA F satisfies universal domain (UD) iff the domain of F is the set of all profiles of n probabilistic beliefs.*

In opinion pooling, an OP F satisfies universal domain, given that $\mathcal{A}$ is an algebra, iff the domain of F is the set of all profiles of n finitely-additive probabilities. The same holds for heterogeneous belief aggregation when we assume $\mathcal{A}$ to be an algebra. A JA F satisfies universal domain iff the domain of F is the set of all profiles of n binary beliefs that are consistent and complete. The reason why we require complete-rationality in the definition of universal domain of a JA is that in most of the literature regarding impossibility results, they assume completeness and consistency of individual beliefs and we would like to compare our results with their results.

Moreover, consistent and complete binary beliefs can be viewed as probabilistic beliefs with values of 0 and 1, which can be inputs of a HA.

Next, we will define a property that is related to outputs of aggregators, i.e., collective beliefs. In contrast to UD, which regulates the domain of aggregators, collective rationality regulates their codomain.

Definition 2.5 (Collective Rationality)**.**

(1) A HA F satisfies collective completeness(CCP) iff $F(\vec{P})$ is complete for all $\vec{P}$ in the domain of F.

(2) A HA F satisfies collective consistency(CCS) iff $F(\vec{P})$ is consistent for all $\vec{P}$ in the domain of F.

(3) A HA F satisfies collective deductive closure(CDC) iff $F(\vec{P})$ is deductively closed for all $\vec{P}$ in the domain of F.

In opinion pooling, collective rationality corresponds to the requirement that $F(\vec{P})$ is a finitely-additive probability on a given algebra. In judgment aggregation, we have the same definition with the domain of F understood as the universal domain of an JA. Some research investigates judgment aggregation under the assumption of complete-rationality[3] and some under rationality.[4] In our study, we will address both assumptions (our triviality and oligarchy results hold under the assumption of collective rationality; our impossibility result under collective complete-rationality). To summarize, we will investigate heterogeneous belief aggregation from the set of all profiles of individual probabilistic beliefs to the set of collective binary beliefs with collective rationality or complete-rationality.

Unanimity, Anonymity, and Independence Now let us consider the second group properties of HAs. Irrespective of whether F is a HA, an OP, or a JA, these norms can be defined in the same ways, although the domain of F is interpreted differently.

Let $\vec{P}(A)$ denote the vector $(P_1(A), ..., P_n(A))(=: (P_i(A))_{i \in N})$ for any $A \in \mathcal{A}$. We will also write it simply $(P_i(A))_i$. We define a kind of unanimity such that unanimous beliefs in an issue are respected whenever everyone has a probabilistic belief of 1 in it or everyone has a probabilistic belief of 0 in it.

Definition 2.6 (Unanimity)**.** *Let F be an aggregator.*

(1) F satisfies certainty preservation(CP) iff for all $A \in \mathcal{A}$, if $\vec{P}(A) = \vec{1}(:= (1, ..., 1) \in [0,1]^n)$, then $F(\vec{P})(A) = 1$ for all $\vec{P}$ in the domain of F.

(2) F satisfies zero preservation(ZP) iff for all $A \in \mathcal{A}$, if $\vec{P}(A) = \vec{0}(:= (0, ..., 0) \in [0,1]^n)$, then $F(\vec{P})(A) = 0$ for all $\vec{P}$ in the domain of F.

[3] See Dokow & Holzman (2010a), Nehring & Puppe (2010).

[4] See Gärdenfors (2006), Dietrich & List (2008).

CP says that if everyone is certain of an issue being true, then it should be collectively believed; ZP says that if everyone has a probabilistic belief of 0 in an issue, then it should not be collectively believed. In opinion pooling, CP and ZP are equivalent. In heterogeneous belief aggregation, we have the following lemma, which is also applicable for judgment aggregation.

Lemma 2.2. *In heterogeneous belief aggregation,*

(1) given CCS(collective consistency), CP implies ZP;

(2) given CCP(collective completeness), ZP implies CP;

(3) $F(\vec{P})(W) = 1$ by CP and $F(\vec{P})(\emptyset) = 0$ by ZP or CCS, when $W, \emptyset \in \mathcal{A}$.

Proof. (1) Assume $\vec{P}(A) = \vec{0}$, which is equivalent to $\vec{P}(\overline{A}) = \vec{1}$ since $\vec{P}$ is a profile of probabilistic beliefs. By CP, we have $F(\vec{P})(\overline{A}) = 1$, from which follows $F(\vec{P})(A) = 0$ by CCS.
(2) Assume $\vec{P}(A) = \vec{1}$, which is equivalent to $\vec{P}(\overline{A}) = \vec{0}$ since $\vec{P}$ is a profile of probabilistic beliefs. By ZP, we have $F(\vec{P})(\overline{A}) = 0$, from which follows $F(\vec{P})(A) = 1$ by CCP.
(3) It holds since $\vec{P}(W) = \vec{1}$ and $\vec{P}(\emptyset) = \vec{0}$. $F(\vec{P})(\emptyset) = 0$ holds by CCS as well. □

This lemma shows that if we demand complete-rationality of collective beliefs, then CP is equivalent to ZP as in opinion pooling. Under the assumption of rationality, CP implies ZP. However, if we only have the assumption of CDC, then from CP does not follow ZP.

Next, let us introduce the anonymity norm of a HA, which requires that collective beliefs should not be inclined to some particular agent's opinion. Although the anonymity norm has been extensively studied in the social choice theory and judgment aggregation literature, it is questionable that this kind of fairness norm should be required in the contexts of epistemic collective decisions. Indeed, in many epistemic collective decision contexts, it would be better to respect and prioritize some agents' opinions who are experts on the issue. However, there may be different situations where the anonymity norm is also required even in epistemic collective decision contexts. E.g., consider situations where it is not known whom the submitted opinions belong to, it is not known which of the agents are experts on the issue, or the group consists of epistemic peers. In the next section, we will proceed with and without this norm to deal with a variety of situations.

Definition 2.7 (Anonymity)**.** *An aggregator F satisfies anonymity(AN) iff $F((P_{\pi(i)})_i) = F((P_i)_i)$ for all $\vec{P}$ in the domain of F and all permutation π on N.*

The last property is the most controversial. Independence between issues ensures that the result of an issue depends only on the individual probabilistic beliefs in the issue regardless of the ones in other issues. It means that an aggregation should be performed issue-wise. It is of practical use that we do not need to consider all values of a profile if we want to focus on the result of one issue. However, this norm can create tension when combined with the requirement of consistency, deductive closure and the logical interconnections among issues. The tension has been pointed out as one of the main culprits of the impossibility results in the judgment aggregation literature. In this chapter, we assume the independence norm and study impossibility results for heterogeneous belief aggregation, and in the succeeding chapters, we will study heterogeneous belief aggregation without the independence norm to avoid our impossibility results. Let us define the independence norm, and a stronger norm, the systematicity norm.

Definition 2.8 (Independence and Systematicity)**.** *Let F be an aggregator.*

(1) F satisfies propositionwise independence(IND) iff for all $A \in \mathcal{A}$, there is a function G_A such that $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P}$ in the domain of F.

(2) F satisfies systematicity(SYS) iff there is a function G such that $F(\vec{P})(A) = G(\vec{P}(A))$ for all $\vec{P}$ in the domain of F and $A \in \mathcal{A}$

Systematicity can be thought of as the independence norm plus the neutrality norm where a HA F is neutral iff for all $A, B \in \mathcal{A}$ if $\vec{P}(A) = \vec{P}(B)$, then $F(\vec{P})(A) = F(\vec{P})(B)$ for all $\vec{P}$ in the domain of F. Neutrality means that each issue is believed or not by the same rule. Combined with independence, it yields the norm of systematicity — i.e., the collective belief on each issue is determined by the same function of individual probabilistic beliefs in that issue.

In opinion pooling, SYS implies CP (certainty preservation) and ZP (zero preservation). In judgment aggregation, we have a lemma in the same form.

Lemma 2.3. *Assume that $\emptyset, W \in \mathcal{A}$. In heterogeneous belief aggregation,*

(1) given CCS(collective consistency), SYS implies ZP.

(2) given CCS and CCP(collective completeness), SYS implies CP.

Proof. (1) By CCS, $0 = F(\vec{P})(\emptyset) = G(\vec{P}(\emptyset)) = G(\vec{0})$.
(2) By CCS, $F(\vec{P})(\emptyset) = 0$ and thus, by CCP, $1 = F(\vec{P})(W) = G(\vec{P}(W)) = G(\vec{1})$. Alternatively, Lemma 2.1 and (1) combined give (2). □

2.4 The Triviality Results

Now we are ready to formulate the first main results. We will prove that an issue-wise independent HA with collective deductive closure and anonymity satisfying universal domain and certainty- and zero-preservation yields the trivial aggregation function — where an issue is believed iff every individual has a probabilistic belief of 1 in the issue. Furthermore, we will drop anonymity and see that it is not so helpful in avoiding degenerate procedures. In addition, we will prove that if we add the assumption of collective consistency and collective completeness, then there is no aggregation function satisfying the above-mentioned properties.

In this section, we assume the following complexity of the logical interconnections in the agenda: it should be a *non-trivial* algebra, defined as follows.

Definition 2.9 (Non-trivial Algebra). *An algebra $\mathcal{A}$ is non-trivial iff it has at least 3 non-empty non-intersecting elements.*

The examples of a trivial algebra on a set W of possible worlds have the form of $\{\emptyset, A, \overline{A}, W\}$ for some $A \subseteq W$. They have only one pair of a contingent issue — i.e., an issue that is neither $\emptyset$ nor W — and its negation. In this case, the logical connection is so minimal that collective deductive closure is not demanding enough for an independent HA to yield the triviality results. Thus, to prove the triviality results, we require that the agenda is not a trivial algebra.

Algebras have quite complex logical connections. Later, we will relax this assumption of being an algebra and prove all the same results in more general settings. The reason why we first provide our main results under a restricted assumption is to compare our results with similar results in opinion pooling where typically, finitely additive probabilities on an algebra are dealt with. Moreover, we consider the aggregation problems where individual beliefs are finitely additive probabilities on a given algebra as the most basic ones in heterogeneous belief aggregation. Therefore, in this section, we will provide direct proofs regarding this restricted, but the most basic form of heterogeneous belief aggregation first.

Triviality Result To prove our main results, we will need a lemma that states that under the assumption of universal domain (UD) and collective deductive closure (CDC), certainty- and zero- preserving (CP, ZP) independent (IND) heterogeneous belief aggregation satisfies systematicity (SYS). We will compare this with the corresponding result in opinion pooling, which says that given UD of opinion pooling, an OP satisfies CP and IND iff it satisfies SYS.[5] In judgment aggregation, there is a lemma that can be stated in the same way as ours except that UD of heterogeneous belief aggregation is replaced by UD of judgment aggregation.[6]

Lemma 2.4 (IND and SYS). *Let $\mathcal{A}$ be a non-trivial algebra and a HA F satisfies UD. If F satisfies CDC, ZP, CP and IND,then it satisfies SYS.*

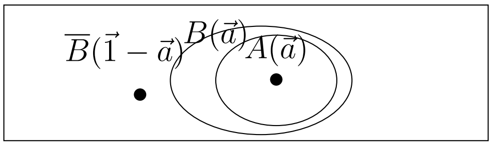
[5] See McConway(1981).

[6] See Dokow & Holzman (2010a).

Proof. By IND, we can let $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P}$. We need to show that $G_A = G_B$ for all $A, B \in \mathcal{A}$.

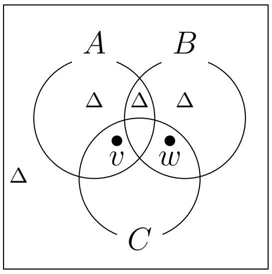
(Case 1) $\emptyset \neq A \subseteq B \neq W$

By UD, $\vec{P}$ given as in the figure below can be an argument of F: since there exist worlds in A and $\overline{B}$, which are represented by dots in the figure, we can assign the probabilities $\vec{a}$ and $\vec{1} - \vec{a}$ to A and $\overline{B}$, respectively. Then B has the probabilities $\vec{a}$.



This gives the following:
(i) If $G_A(\vec{a}) = F(\vec{P})(A) = 1$, then $F(\vec{P})(B) = G_B(\vec{a}) = 1$ by CDC (closure under superset).
(ii) Since $F(\vec{P})(A \cup \bar{B}) = 1$ by CP, if $G_B(\vec{a}) = F(\vec{P})(B) = 1$, then $F(\vec{P})(A) = G_A(\vec{a}) = 1$ by CDC (closure under intersection), since $(A \cup \bar{B}) \cap B = A$.

(Case 2) $A - B \neq \emptyset$ and $B - A \neq \emptyset$



Δ indicates all possible locations at one of which a world is ensured to exist so that $C \neq W$.

Let $v \in A - B$ and $w \in B - A$ as in the above figure. We can use $C \in \mathcal{A}$ such that $\{v, w\} \subseteq C \neq W$ since $\mathcal{A}$ is a non-trivial algebra. (Take the union of two elements in $\mathcal{A}$ one of which includes v and one of which includes w. We can find such two elements of which the union is not W because $\mathcal{A}$ is non-trivial. If there were no such two elements, it means that $(A - B) \cup (B - A) = W$ and thereby $\mathcal{A} = \{\emptyset, A, B, W\}$ where $B = \overline{A}$, which contradicts the non-triviality of $\mathcal{A}$.) By the result of (Case 1) we have

$$G_A = G_{A \cap C} = G_C = G_{C \cap B} = G_B$$

.

(Case 3) We can let $G_\emptyset = G_A$ and $G_W = G_A$, for any $A (\neq \emptyset, W) \in \mathcal{A}$, since $G_A(\vec{0}) = 0$ by ZP and $G_A(\vec{1}) = 1$ by CP.

$\square$

This lemma shows that to obtain SYS from CP and IND, we need CDC and ZP, neither of which are needed to get the corresponding result in opinion pooling. First, let us focus on CDC. Recall that the requirement of collective rationality on a probabilistic belief is that it is extendable to a finitely-additive probability. In opinion pooling, this is not satisfied by a separate rationality condition but rather by the definition of an opinion pooling function. Accordingly, one may expect that we should add some rationality conditions in heterogeneous belief aggregation. This lemma shows that we only need to add CDC without requiring CCS. Second, ZP is used only in (Case 3) for $\emptyset$. If our agenda is not an algebra and $\emptyset \notin \mathcal{A}$ (to be discussed later), ZP is not required.[7] Provided $\emptyset \in \mathcal{A}$ as in this section, we need ZP, which does not follow from CP — differently from opinion pooling — since we do not assume CCS for this lemma.

Now we prove our main results that the following two theorems state. We first assume anonymity (AN) and prove that the conditions for obtaining SYS in the previous lemma yield only the trivial aggregation function. Further, we will drop AN and prove that those conditions lead to oligarchic aggregation functions such that there are oligarchs whose unanimous certain beliefs are the necessary and sufficient condition for the collective belief in the issue.

It is worth comparing our theorems with the corresponding ones in judgment aggregation: if a JA F satisfies UD of JA, ZP, CP, IND and CDC, then F is oligarchic and if AN is added, then F is trivial.[8] In opinion pooling, given UD of OP, an OP satisfies CP and IND iff it is a linear pooling function under the assumption that $\mathcal{A}$ is a σ-algebra.[9] In heterogeneous belief aggregation, we have the following:

Theorem 2.5 (Triviality Result)**.** *Let $\mathcal{A}$ be a non-trivial algebra. The only HA satisfying UD, ZP, CP, IND, CDC and AN is the following trivial function:*

$$F(\vec{P})(A) = \begin{cases} 1 & \text{if } \vec{P}(A) = \vec{1} \\ 0 & \text{otherwise} \end{cases}$$

for all $A \in \mathcal{A}$ and all profiles $\vec{P}$ in the domain.

Proof. It is easily seen that F satisfies all mentioned properties. For the other direction, by Lemma 2.4 we have SYS and thus, we can let $F(\vec{P})(A) = G(\vec{P}(A))$ where $G(\vec{1}) = 1$ by CP. Now suppose that $G(\vec{a}) = 1$ for some $\vec{a} \neq \vec{1}$ and pick up any $a_i \neq 1$ in $\vec{a}$. To derive a contradiction, we take the following three steps.

[7] See Lemma 2.8 in Section 2.6

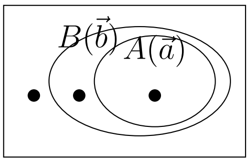
[8] See Dietrich & List (2008). They assumed, instead of ZP and CP, the weaker condition of unanimity preservation such that if $P_i = P_j (:= P)$ for all i and j, then $F((P_i)_i) = P$. We call it weak unanimity. From ZP and CP, follows weak unanimity but the converse does not hold without any further assumptions. However, it is easily shown that if we have IND, the converse also holds. Therefore, their result is equivalent to the above statement.

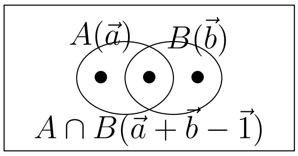
[9] See McConway (1981)

[**Step 1**] To show the following:

(Fact 1) if $\vec{a} \leq \vec{b}$ and if $G(\vec{a}) = 1$, then $G(\vec{b}) = 1$

(Fact 2) if $\vec{a} + \vec{b} - \vec{1} \geq \vec{0}$ and if $G(\vec{a}) = 1$ and $G(\vec{b}) = 1$, then $G(\vec{a} + \vec{b} - \vec{1}) = 1$





Since $\mathcal{A}$ is non-trivial, we have at least 3 non-empty elements of $\mathcal{A}$ that have no intersections with each other. We represent a world of each element as a dot in the above figures. Since $\mathcal{A}$ is an algebra there are A and B in $\mathcal{A}$ as in the left figure and A, B and $A \cap B$ as in the right figure. Moreover, we can, by UD, assign the probabilities $\vec{a}$ and $\vec{b}$ such that $\vec{a} \leq \vec{b}$ to A and B, respectively, in the left figure. In the right figure, by UD we can assign the probabilities $\vec{a}$, $\vec{b}$ and $\vec{a} + \vec{b} - \vec{1}$ to A, B and $A \cap B$, respectively where $\vec{a} + \vec{b} - \vec{1} \geq \vec{0}$. The left figure gives us (Fact 1) by the closure under superset from CDC and the right figure gives us (Fact 2) by the closure under intersection from CDC.

[**Step 2**] To show that $G(\vec{a}[a_i \mapsto 0, a_l \mapsto 1 \text{for all} l \neq i]]) = 1$

By (Fact 1), we can substitute a_i and a_l with any higher values and by mixed applications of (Fact 1) and (Fact 2), we can substitute a_i with any lower value using that

$$\text{if } G(\vec{a}) = 1 \text{ then } G(\vec{a} + \vec{a} - 1) = 1 \tag{2.1}$$

as the following shows:

e.g., $a_i = 0.8$; $G(\vec{a}[a_l \mapsto 1 \text{ for all } l \neq i]) = 1$, $0.8+0.8-1 = 0.6$, $0.6 \leq 0.75$, $0.75 + 0.75 - 1 = 0.5$, $0.5 + 0.5 - 1 = 0$

For example, let $G(\vec{a}) = 1$ where $a_i = 0.8$ as in the above figure. By (Fact 1), we can substitute all other components a_l (i.e., $l \neq i$) that is not 1 with 1 and so we have $G((1, ..., 1, 0.8, 1, ..., 1)) = 1$ where 0.8 is the i-th component of the vector. This process enable us to focus on i-th component of vectors in $\{\vec{a}' | G(\vec{a}') = 1\}$ when we apply (2.1), because $1 + 1 - 1 = 1$. Now apply (Fact 1) and (2.1). The above calculation shows $G((1, ..., 1, 0.6, 1, ..., 1)) = G((1, ..., 1, 0.75, 1, ..., 1)) = G((1, ..., 1, 0.5, 1, ..., 1)) = G((1, ..., 1, 0, 1, ..., 1)) = 1$. This example can be generalized so that we have $G((1, ..., 1, 0, 1, ..., 1)) = 1$ for any value of a_i: if we start with $a_i \leq 0.75$, then we can apply the last steps in the example; otherwise, repeated application of (2.1) will lead to $G((1, ..., 1, a'_i, 1, ..., 1)) = 1$ for some $a'_i \leq 0.75$. It is because after applying (2.1) k times, we have $G((1, ..., 1, a_i^{(k)}, 1, ..., 1)) = 1$ where $a_i^{(k)} = 1 - 2^k(1 - a_i)$ and there must be k such that $a_i^{(k)} \leq 0.75$.

[Step 3] To show, by induction, $G((0, ..., 0)) = 1$ that contradicts ZP.

$G((0, 1, ..., 1)) = G((1, 0, 1, ..., 1)) = ... = G((1, ..., 1, 0)) = 1$ by UD, AN and (Step 2). Let $\vec{a}_k$ be a vector $(0, ..., 0, 1, ..., 1)$ where the first k entries are zero and one elsewhere. For $k = 1$, we have $G(\vec{a}_1) = 1$. Assume $G(\vec{a}_k) = 1$. Since $G((1, ..., 1, 0, 1, ..., 1)) = 1$ where all entries of the input are one except for the (k+1)th one that is zero, by (Fact 2), we have $G(\vec{a}_{k+1}) = 1$. The following figure depicts this process.

$$\boxed{\begin{array}{l} G(0, ..., 0, 1, 1, ..., 1) = 1 \\ G(1, ..., 1, 0, 1, ..., 1) = 1 \\ \hline G(0, ..., 0, 0, 1, ..., 1) = 1 \end{array}}$$

□

This theorem shows that under certain mild conditions (UD, ZP, and CP) it is impossible that an anonymous (AN) issue-wise (IND) heterogeneous belief aggregation generates deductively closed (CDC) collective beliefs, except for the trivial function. To obtain this result, we made use of neither complete-rationality nor collective consistency (CCS). CDC alone is a sufficient rationality condition on collective beliefs that causes tension with IND.

It is of interest to see that the conditions that lead to the trivial function are the same as the ones in the corresponding result in judgment aggregation, which we mentioned before. One might wonder whether our result can be obtained directly from that result, but this is not the case. We will compare our results with the ones in judgment aggregation in the next section in detail and provide the reason why not.

It is worth understanding the structure of our proof because it will be repeatedly applied to prove other theorems and discussed in several places throughout this chapter. [10] First, from UD, CDC, ZP, CP, and IND follows SYS by Lemma 2.4, so we have a function G with CP that assigns 0 or 1 to a vector of probabilities on any issue. [Step 1] is to prove (Fact 1) and (Fact 2) about G using CDC and the agenda condition of being a non-trivial algebra. (Fact 1) represents, under the assumption of SYS, a kind of *monotonicity*, defined by

$$\text{(MON) If } \vec{P}(A) \leq \vec{P}'(A) \text{ and } F(\vec{P})(A) = 1\text{, then } F(\vec{P}')(A) = 1$$

where $\leq$ is applied to each component of two vectors. (Fact 1) means that if G assigns 1 to a vector, then it does the same to every greater vector than that, which we call *upward closure of* $G^{-1}(1)$. In contrast, (Fact 2) shows that certain smaller vectors than a vector in $G^{-1}(1)$ are also contained in it, which we call *restricted downward closure of* $G^{-1}(1)$. Next, in [Step 2], we prove that any non-trivial HA —i.e., 1 is assigned to a vector $\vec{a}$ with $a_i \neq 1$ — yields $\vec{a}' := (1, ..., 1, 0, 1, ..., 1) \in G^{-1}(1)$ where 0 is i-th component.(i.e., $a'_i = 0$): Upward closure of G leads to $a'_l = 1$ for all $l \neq i$;

[10] We will apply the same method as in the proof of Theorem 2.6 to Theorem 2.10 (1) and Theorem 2.13 (1).

both upward- and restricted downward- closure of G give $a'_i = 0$. Finally, using AN and (Fact 2), [Step 3] shows $G((0, ..., 0)) = 1$, a contradiction to ZP.

One may ask whether AN is the main culprit in the difficulty of an issue-wise HA to get deductively closed collective beliefs. The following theorem shows that this is not the case and dropping AN is not a sufficient mean to avoid the difficulty.

Without Anonymity: Oligarchy Now, we drop the assumption of anonymity and show that it leads to a degenerate HA as well. To do that, we define an oligarchy first. Recall that N denotes the set of the individuals and $\mathcal{A}$ denotes the agenda.

Definition 2.10 (Oligarchy)**.** *An aggregator F is an oligarchy if there is a non-empty subset M of N such that*

$$F(\vec{P})(A) = \begin{cases} 1 & \textit{if } P_i(A) = 1 \textit{ for all } i \in M \\ 0 & \textit{otherwise} \end{cases}$$

for all $A \in \mathcal{A}$ and all profiles $\vec{P}$ in the domain. When $|M| = 1$, we call F a dictatorship.

So an oligarchy means that there are oligarchs whose unanimous certain belief on an issue is the necessary and sufficient condition for the collective belief in that issue. This is also problematic partly because other individuals than the oligarchs are excluded in the decision process. However, we can think of oligarchs as experts, and to obtain true collective beliefs, relying on experts would not be irrational. Nevertheless, non-oligarchy is a rational requirement even in epistemic contexts where beliefs are dealt with, because after excluding non-experts, the decision process among the oligarchs can be viewed as the trivial aggregation among the oligarchs.

With this type of degenerate HA, we have another impossibility theorem, which says that the same conditions as the triviality result except for AN lead to oligarchies.

Theorem 2.6 (Oligarchy Result)**.** *Let $\mathcal{A}$ be a non-trivial algebra. The only HAs satisfying UD, ZP, CP, IND and CDC are oligarchies.*

Proof. It is obvious that an oligarchy satisfies the properties.
For the other direction, to construct the set M of oligarchs in the Definition 2.10, we employ [Step 1] and [Step 2] in the proof of Theorem 2.5. (By UD, ZP, CP, IND and CDC, we have SYS by Lemma 2.4 and from UD and CDC follows [Step 1] and [Step 2]. Note that in the proof of Theorem 2.5, we did not use AN except [Step 3].) Consider the set $G^{-1}(1) := \{\vec{a}|G(\vec{a}) = 1\}$ where G is a function satisfying $F(\vec{P})(A) = G(\vec{P}(A))$. We collect individuals i such that $a_i = 1$ for all $\vec{a} \in G^{-1}(1)$ and define the set M of such individuals:

$$M := \{i \in N | a_i = 1 \text{ for all } \vec{a} \text{ such that } G(\vec{a}) = 1\}$$

We will show (i) and (ii) in the following:

(i) M is non-empty.

Suppose M is empty. Then $G(0,1,...,1) = G(1,0,1,...,1) = ... = G(1,...,1,0) = 1$ by [Step 1] and [Step 2] and we have $G(0,...,0) = 1$ using the same way of [Step 3], which contradicts ZP.

(ii) $a_i = 1$ for all $i \in M$ iff $G(\vec{a}) = 1$

($\leftarrow$) It is obvious by the construction of M. ($\rightarrow$) Since we have (Fact 1) in [Step 1], it is enough to show that

$$G((\delta_{i\in M})_i) = 1$$

where $\delta_{i\in M} = 1$ if $i \in M$, otherwise $\delta_{i\in M} = 0$. For any $j \notin M$, there is $\vec{a}$ such that $G(\vec{a}) = 1$ and $a_j \neq 1$, by definition of M. By [Step 2],

$$G(\vec{a}[a_j \mapsto 0, a_l \mapsto 1 \text{ for all } l \neq j]) = 1 \tag{2.2}$$

Now we proceed by induction analogously to [Step 3]. Enumerate individuals who are not in M, like $j_1, j_2, ..., j_{|N|-|M|}$ and let $\vec{a}_k$ be a profile where $a_{j_1} = 0, ..., a_{j_k} = 0$ and other components are all 1. For $k = 1$, we have $G(\vec{a}_1) = 1$. Assume $G(\vec{a}_k) = 1$. Since by (2.2) we have $G(1,...,1,0,1,...,1) = 1$ where 0 is j_{k+1}'th component,by (Fact 2) in [Step 1], we have $G(\vec{a}_{k+1}) = 1$. Therefore, we have

$$G(\vec{a}_{|N|-|M|}) = G((\delta_{i\in M})_i) = 1$$

□

This theorem generalizes the triviality result and shows that dropping AN leads to the oligarchy result. In the proof, we adopted [Step 1]— thus (Fact 1) and (Fact 2) — and [Step 2] in the proof of the triviality result, since we have not used AN to prove them. The two proofs are similar in spirit; but instead of deducing $G((0,...,0)) = 1$ from $G(\vec{a}) = 1$ where $a_i \neq 1$ for some $i \in N$, we derived $G((\delta_{i\in M})_i) = 1$ from $G(\vec{b}) = 1$ where $b_j \neq 1$ for some $j \in N - M$. In this process, we used [Step 2] and applied induction analogously to [Step 3] — except that induction is not on all individuals but only on non-oligarchs — to prove that even if all non-oligarchs certainly believe that an issue is false, the oligarchs' unanimous certain beliefs in the issue yield the collective belief in it.

With CCP and CCS: Impossibility Until now we assumed only CDC. If we impose stronger collective rationality, we would get a stronger result, which we obtain as a corollary of the oligarchy result.

In judgment aggregation, if F satisfies UD of judgment aggregation, CP, IND, CCP, and CCS, then F is a dictatorship,[11] and if we add AN then there is no such JA. We now present a similar result in heterogeneous belief aggregation.

Corollary 2.6.1 (Impossibility Result). *Let $\mathcal{A}$ be a non-trivial algebra. There is no HA satisfying UD, CP, IND, CCP, and CCS.*

[11] See Dokow & Holzman (2010a)

Proof. Since CDC follows from CCP and CCS and ZP follows from CCS and CP, the only possible HAs satisfying the above conditions would be oligarchies by Theorem 2.6, which do not satisfy collective completeness. □

Note that to use Lemma 2.4 and the triviality result of Theorem 2.6, we need to assume CCS since we need to get CDC from CCP and CCS. Under the assumption of CCS and CCP, it holds that ZP iff CP. (See Lemma 2.2.)

In judgment aggregation, from the same assumptions follow dictatorships where a dictator's binary belief is complete. By contrast, in heterogeneous belief aggregation, a dictator's binary belief — an issue is believed iff a dictator gives it a probability of 1 — does not satisfy completeness. Therefore, we need neither non-dictatorship nor anonymity to obtain the impossibility result in heterogeneous belief aggregation.

To summarize this section, we proved that given that $\mathcal{A}$ is a non-trivial algebra, there is no HA satisfying one of the following:
(1) UD, ZP, CP and IND + CDC + Non-oligarchy
(2) UD, ZP, CP and IND + CDC + AN + Non-triviality
(3) UD, CP and IND + CCS and CCP
Note that from the impossibility of (1) follow that of (2) and that of (3). It is because the only anonymous oligarchy is the trivial HA and there is no way that the resulting opinion of the oligarchy is complete. Nevertheless, we first provided a direct proof of the triviality result separately and then modified it to obtain the oligarchy result. It is because the fact that the triviality result follows from the oligarchy result holds only under certain agenda conditions[12] (to be described in Section 2.6), which the agenda $\mathcal{A}$ in this section, being a non-trivial algebra, satisfies. In more general settings (to be discussed later[13]), the triviality result might not follow from the oligarchy result, and to prove the triviality result, we need to apply [STEP 3] in the proof of Theorem 2.5, which is a step only for the triviality result.

[12] In Section 2.6, we will prove that path-connectedness and even-negatability are the necessary and sufficient agenda conditions for the oligarchy result and thus the triviality result follows from the oligarchy result only under that agenda condition.

[13] For example, if the agenda is negation connected as in Theorem 2.13 (1), the triviality result does not follow from the oligarchy result and to prove the triviality result, we need to apply [STEP 3] in the proof of Theorem 2.5. See the proof of Theorem 2.13 (1).

2.5 Comparison with JA, OP and Belief Binarization

HA and JA Let us compare our results with some impossibility results in judgment aggregation. The impossibility results with (1),(2), and (3) are our results in the last section, the ones with (1′) and(2′) from Dietrich & List (2008). The impossibility results with (3′) and (4′) are from Dokow & Holzman (2010a) and Nehring & Puppe (2010).

There is no HA with UD of heterogeneous belief aggregation satisfying
(1) ZP, CP and IND + CDC + Non-oligarchy
(2) ZP, CP and IND + CDC + AN + Non-triviality
(3) CP and IND + CCS and CCP

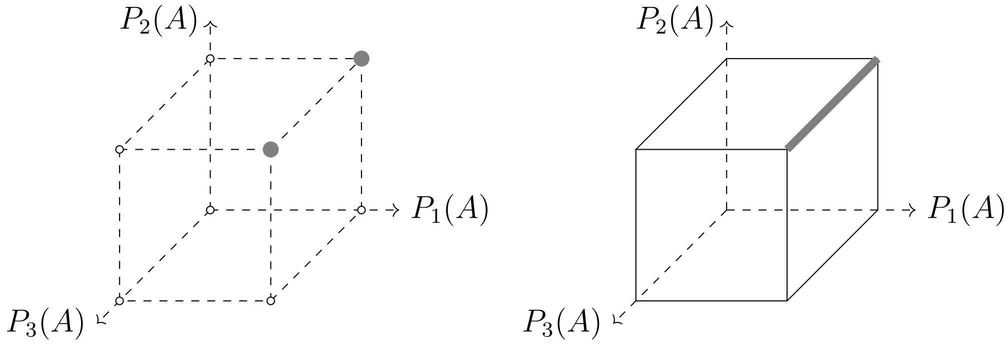
There is no JA with UD of judgment aggregation satisfying
(1′) ZP, CP and IND + CDC + Non-oligarchy
(2′) ZP, CP and IND + CDC + AN + Non-triviality
(3′) CP and IND + CCS and CCP + Non-dictatorship
(4′) CP and IND + CCS and CCP + AN

The universal domain of judgment aggregation can be viewed as a subset of the one of heterogeneous belief aggregation since consistent and complete binary beliefs can be viewed as probabilistic beliefs. Therefore, the restriction of a HA F with UD to the universal domain of judgment aggregation can be regarded as a JA, denoted by $F\upharpoonright$, and an extension of a JA F' to the universal domain of heterogeneous belief aggregation can be regarded as a HA, denoted by $F'\upharpoonleft$. So it is of interest to know whether our results can be obtained directly from the corresponding results in judgment aggregation or the other way around, through some restrictions or expansions. In what follows, (I) and (II) will argue that these are not the case. In addition, (III) will show the similarity and difference of the proofs in heterogeneous belief and judgment aggregation. In (III), we will also explain in detail where the originality of our proofs lies. From now on, HAs and JAs are assumed to satisfy UD of heterogeneous belief aggregation and UD of judgment aggregation respectively, which will not affect our argument. Our arguments will be based on the following observations.

(Observation 1) If a HA F satisfies ZP/CP/IND/CDC/AN/CCP/CCS, then so does the JA $F\upharpoonright$.

(Observation 2) If a JA F' satisfies non-oligarchiy/non-triviality, then so does any HA $F'\upharpoonleft$.

(Observation 1) holds because each property mentioned above is stated with "for all $\vec{P}$ in the domain" and the universal domain of heterogeneous belief aggregation



The left and right figures describe an oligarchy in JA and HA, respectively where $N = \{1, 2, 3\}$ and $M = \{1, 2\}$. Grey points represent vectors assigned to 1.

includes the one of judgment aggregation. (Observation 2) holds because if a HA F is an oligarchic/trivial function, then the JA $F \upharpoonright$ is an oligarchic/trivial function.(See the above figure.)

(I) First, consider whether our results follow directly from the results in judgment aggregation. In (Observation 1) regarding ZP/CP/IND/CDC/AN/CCP/CCS, each property of a HA F alone leads to that of $F \upharpoonright$ in an obvious way without combining other properties or agenda conditions. As for non-oligarchy/non-triviality, this is not so: without combining any other properties — e.g., CP, IND and CDC — and some agenda conditions, it does not hold that if a HA F satisfies non-oligarchy/non-triviality, then so does the JA $F \upharpoonright$, in contrast to (Observation 2). If that did hold, then we could, together with (Observation 1), argue as follows: if there were a HA F with UD satisfying (1)/(2)/(3), then $F \upharpoonright$ would satisfy (1′)/(2′)/[(3′) and (4′)], a contradiction, which would show that there is no such HA; thus we could obtain the impossibility of (1)/(2)/(3) directly from the ones of (1′)/(2′)/[(3′) or (4′)]. Hence, in this sense, we can conclude that our results do not directly follow from the result in judgment aggregation. Indeed, in our proof of Theorem 2.5, [Step 2] is a step to show that for any non-trivial HA F, $F \upharpoonright$ is a non-trivial JA, using (Fact 1) and (Fact 2) of [Step 1], which can be proved under the assumption of SYS and CDC and under the agenda condition of being non-trivial algebra.

(II) Next, consider the other way — whether the results in judgment aggregation follow directly from our results. For the sake of argument, suppose that there was a direct and typical way to extend a JA F' satisfying ZP, CP, and IND together with CDC/[AN and CDC]/[CCS and CCP] to a HA $F' \upharpoonleft$ satisfying the same properties of heterogeneous belief aggregation. Then $F' \upharpoonleft$ would [be an oligarchy]/[be the trivial function]/[not exist], that would lead to an oligarchy/ the trivial function/nonexistence of F' by (Observation 2). However, there is no such direct and typical way. We will see this in the next section.[14]

[14]Two natural ways to extend a JA to a HA would be the following:(i) assign 0 to all profiles of probabilities that are not 0-1 probabilities — such as the counterexample of Theorem 2.10 (3) — (ii) extend to keep monotonicity but minimally extend — such as the counterexample of Theorem

(III) Nevertheless, following the same methods as our proofs, we can obtain the same statement as Lemma 2.4, Theorem 2.5 and Theorem 2.6 for judgment aggregation. It is because our reasoning is not affected if we restrict every $\vec{P}$ in our proofs to profiles of 0-1-valued probabilities — e.g., it is obvious that (Fact 1) and (Fact 2) of [Step 1] in the proofs of our triviality and oligarchy results imply the restrictions of (Fact 1) and (Fact 2) to the vectors of which components are 0 or 1, denoted by (Fact 1)$\restriction$ and (Fact 2)$\restriction$.[15]

In contrast, adopting the corresponding proofs in judgment aggregation is not enough to get our results. First of all, even though the proofs in judgment aggregation might give (Fact 1)$\restriction$ and (Fact 2)$\restriction$, we need to prove that it is extended to our domain so that we can obtain (Fact 1) and (Fact 2). On top of that, in the proofs of our triviality and oligarchy results, $a_i \neq 1$ (where i is any individual)/$a_j \neq 1$(where j is not an oligarch) does not mean that $a_i = 0/a_j = 0$ contrary to judgment aggregation. Thus, we need to prove that $G(\vec{a}[a_i \mapsto 0, a_l \mapsto 1 \text{ for all } l \neq i]) = 1/\ G(\vec{a}[a_j \mapsto 0, a_l \mapsto 1 \text{ for all } l \neq j]) = 1$ in [Step 2], which is the step we mentioned in (I). More precisely, this step can be divided into two sub-steps: (i) substitute every $a_l (l \neq i/l \neq j)$ with 1 and prove that G still assigns 1 by (Fact 1); (ii) substitute a_i/a_j with 0 and prove that G still assigns 1 by (Fact 1) and (Fact 2). (i) and thus (Fact 1) are needed not only for heterogeneous belief aggregation but also when we use our proof to get the results in judgment aggregation.[16] (ii) is the step only for heterogeneous belief aggregation. The following table compares the key claims to prove the triviality result in heterogeneous belief aggregation with the ones we need when applying our proofs to prove the triviality result in judgment aggregation.

	[Step 1]	[Step 2]	[Step 3]
HA	(Fact 1), (Fact 2)	$G(\vec{a}[a_i \mapsto 0, a_l \mapsto 1 \text{ for all } l \neq i]) = 1$ (using (Fact 1) and (Fact 2))	$G(0, ..., 0) = 1$ (using (Fact 2))
JA	(Fact 1)$\restriction$, (Fact 2)$\restriction$	$G(\vec{a}[a_l \mapsto 1 \text{ for all } l \neq i]) = 1$ (using (Fact 1)$\restriction$)	$G(0, ..., 0) = 1$ (using (Fact 2)$\restriction$)

HA and OP (1) We cannot use the proofs in opinion pooling to get the Lemma and the triviality/oligarchy result, because we do not assume CCS and CCP to get them, thus collective beliefs cannot be regarded as 0-1-valued probabilistic beliefs. Indeed, the crucial step in the corresponding proofs in opinion pooling is to use the additivity

2.10 (2). However, each of them does not work in a direct and typical way: the latter example shows that (i) does not work and the former example shows that (ii) does not work. Therefore, the counterexamples in judgment aggregation do not extend to heterogeneous belief aggregation in an obvious way.

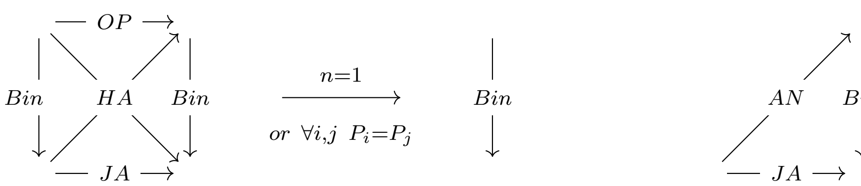
[15] (Fact 1)$\restriction$ corresponds to closure under superset of winning coalitions(sets of agents whose beliefs and the other agents' non-beliefs are the necessary and sufficient condition for the collective belief) in judgment aggregation.

[16] It does not mean that (Fact 2)$\restriction$ is not needed for the results in judgment aggregation, because (Fact 2)$\restriction$ is needed in [Step 3] that we need when we use our proof for judgment aggregation.

axiom: $F(\vec{P})(A\dot{\cup}B) = F(\vec{P})(A) + F(\vec{P})(B)$ given $A \cap B = \emptyset$. In heterogeneous belief aggregation, this does not hold under the assumption of collective weak rationality (i.e., CCS and CCD).

(2) On the other hand, provided $\mathcal{A}$ is a σ-algebra, we can employ the proofs of McConway (1981) to get the impossibility result in heterogeneous belief aggregation under the assumption of complete-rationality (i.e., CCP and CCS). Under this assumption, our outputs - collective binary beliefs - can be thought of as 0-1-valued probability measures, thus heterogeneous belief aggregation as opinion pooling with the restricted co-domain. Recall that according to McConway(1981), given UD of OP, an OP satisfies CP and IND iff it is a linear pooling. This gives the impossibility result since a linear average might not be in the co-domain of a HA.

HA and Belief Binarization Recall that we assumed the number n of individuals is more than 1. It is because we are investigating collective decisions dealing with multiple people. However, in any proofs so far, we have not used the assumption that $n \geq 2$, thus our results so far can be applied to the case where $n = 1$ that is the same as the problem of belief binarization. Another way to equate heterogeneous belief aggregation with belief binarization is to restrict the domain of heterogeneous belief aggregation to the set of profiles of probabilistic beliefs where $P_i = P_j$ for all i and j. Note that, as Dietrich & List (2018) indicated, issue-wise(IND) judgment aggregation with AN is the same problem as the problem of belief binarization, to which our results can also be applied.



The left figure illustrates the relation between heterogeneous belief aggregation and belief binarization, while the right figure depicts the relation between judgment aggregation and belief binarization.

Thus, it is not surprising that our triviality and impossibility results modified for $n = 1$ recover the theorems that state that there is neither JA with AN satisfying (2′) or (4′), nor belief binarization satisfying (2′′) or (4′′), which are the results from Dietrich & List (2021) and Dietrich & List (2018).

There is no belief binarization rule satisfying ...
(2′′) UD, CCS[17], CP and IND + CDC + Non-triviality[18]
(4′′) UD, CP and IND + CCS and CCP

[17]In Dietrich & List (2021), the stronger assumption of CCS is used rather than ZP. Later (See the comments after Theorem 2.13.) we will show that this can be weakened to ZP.

[18]Note that Non-triviality here corresponds to non-looseness in Dietrich & List (2021)

2.6 The Agenda Conditions

This section is based on joint work with Chisu Kim. In the last two sections, we assumed the agenda to be a non-trivial algebra, which is the most typical when dealing with probabilistic beliefs. However, when we make a collective decision, there are more general cases where the agenda does not have the structure of an algebra. For example, we might want to obtain a collective belief on two issues, but not on their conjunction or disjunction. Thus, in this section, we relax the assumption about the richness of the logical interconnections in our agenda and prove that the results in the last sections hold in more general settings as well. Moreover, we examine under what minimal agenda conditions we obtain the results of the last section.

While it is typical in opinion pooling and belief binarization to deal with finitely additive probabilities on an algebra, in judgment aggregation, the agenda being an algebra is not typically required and the minimal agenda conditions that lead to impossibility results have been extensively explored. Recently, this approach has been applied to opinion pooling problems (Dietrich & List (2017a, 2017b), Herzberg (2017)) and belief binarization problems (Dietrich & List (2018, 2021)). In heterogeneous belief aggregation, it is also worthwhile to study how we can expand the classes of agendas to ensure our impossibility results, and characterize them.

We emphasize that, as seen in Section 2.1, in our framework for heterogeneous belief aggregation, the agenda of an HA is not assumed to be an algebra, inputs are profiles of individual probabilistically coherent beliefs, and the properties of an HA are all defined in this general framework. Thus, we can address the above questions in our framework without fixing the setup. In this framework, let us reformulate our question: what are the sufficient and necessary agenda conditions for the oligarchy, triviality, and impossibility results?

The answer (to be proved in this section) is stated in Table 2.1: (1) path- connectedness and even-negatability constitute the exact agenda condition for the oligarchy result; (2) negation-connectedness is for the triviality result; and (3) blockedness is for the impossibility result. These new findings can be compared to the existing characterization theorems in judgment aggregation and belief binarization. (1) has the same agenda condition as (1′) and (3′) in judgment aggregation. (2) is similar to (2″) in belief binarization, except that ZP is used for (2) in the place of CCS for (2″). We will argue that applying our proofs can weaken CCS to ZP and thus, the agenda condition for (2′), which has not been discussed in the literature, is also negation-connected because an anonymous and independent JA can be viewed as a belief binarization function. (3) is similar to (4′) in judgment aggregation and (4″) in belief binarization.

There is no HA satisfying ...	Sufficient and Necessary Agenda Condition
(1) UD, ZP, CP and IND + CDC + Non-oligarchy	path-connected, even-negatable
(2) UD, ZP, CP and IND + CDC + AN + Non-triviality	negation-connected
(3) UD, CP and IND + CCS and CCP	blocked

There is no JA satisfying ...	Sufficient and Necessary Agenda Condition [19]
(1′) UD, ZP, CP and IND + CDC + Non-oligarchy	path-connected, even-negatable (Dietrich & List (2008))
(2′) UD, ZP, CP and IND + CDC + AN + Non-triviality	negation-connected
(3′) UD, CP and IND + CCS and CCP + non-dictatorship	path-connected, even-negatable (Dokow & Holzman (2010a))
(4′) UD, CP and IND + CCS and CCP + AN	blocked (Nehring & Puppe (2010))

There is no belief binarization rule satisfying ...	Sufficient and Necessary Agenda Condition
(2″) UD, CCS , CP and IND + CDC + Non-triviality	negation-connected (Dietrich & List (2021))
(4″) UD, CCS, CP and IND + CCP	blocked (Dietrich & List (2018))

Table 2.1: Classification of agendas generating impossibility results

2.6.1 The Agenda Condition for the Oligarchy Result

We can prove the same results in the previous sections even though we reduce the agenda richness of a nontrivial algebra to some agenda conditions, e.g., path-connectedness and even-negatability. We now introduce them. Recall that W denotes a non-empty set of worlds and $\mathcal{A}$ is an agenda on W, which is a complement-closed non-empty set of some subsets of W. In this section, we assume that $\mathcal{A}$ is finite.[20] By Definition 2.2, a subset $\mathcal{Y} \subseteq \mathcal{A}$ entails $A(\in \mathcal{A})(\mathcal{Y} \vDash A)$ iff there is a finite subset $\mathcal{X} \subseteq \mathcal{Y}$ such that $\bigcap \mathcal{X} \subseteq A$. Since we are dealing with finite agendas, $\mathcal{Y} \vDash A$ iff $\bigcap \mathcal{Y} \subseteq A$. Note that we do not exclude $\mathcal{Y} = \emptyset$ and adopt the convention that $\bigcap \emptyset = W$. When $\mathcal{Y}$ is a singleton set, say $\{A\}$, and $B \in \mathcal{A}$, we write $A \vDash B$ for $\{A\} \vDash B$.

How do we capture the type of complexity of the logical interconnections in the agenda? In the judgment aggregation literature, conditional entailment has been one of the preliminary concepts to do it. The notion of entailment only captures logical interdependences between any antecedents and their consequents. However, two issues that seem logically unrelated at first glance might become logically related when some other issues are combined. Conditional entailment includes this kind of indirect entailment as well as direct one. Furthermore, conditional entailment relations can be thought of as a bridge from one issue to another, and we can start

[19] Note that in Dietrich & List (2008) they do not assume a finite agenda while in Dietrich & List (2021), Dokow & Holzman (2010a), and Nehring & Puppe (2010) they assume a finite agenda. In Dietrich & List (2018) for the sufficient condition they do not assume a finite agenda and for the necessary condition, they do assume a finite agenda.

[20] In this section, we will use some results of previous researches where they assume finite agenda.(e.g., Dokow & Holzman (2010a), Nehring & Puppe (2010) and Dietrich & List (2021)) Some of our results do not hold with infinite agenda and some do. To avoid confusion, we assume finite agenda. Even when the assumption is redundant, assuming finite agenda makes our proofs simple.

from one issue and reach other issues via several conditional entailment relations in a row. The notion of a path from one issue to another issue is devised to capture this kind of even more indirect relations.

Definition 2.11 (Conditional Entailment). *Let $A, B \in \mathcal{A}$ and $\mathcal{Y} \subseteq \mathcal{A}$ where $\mathcal{Y}$ is consistent with A and $\overline{B}$. (That is, $\mathcal{Y} \cup \{A\} \nvDash \emptyset$ and $\mathcal{Y} \cup \{\overline{B}\} \nvDash \emptyset$) A entails B conditional on $\mathcal{Y}$($A \vDash^*_{\mathcal{Y}} B$) iff $\{A\} \cup \mathcal{Y} \vDash B$. If $\mathcal{Y}$ is a singleton set, say $\{C\}$, we write $\vDash^*_C$ instead of $\vDash^*_{\{C\}}$. Moreover, A conditionally entails B($A \vDash^* B$) iff there is a subset $\mathcal{Y} \subseteq \mathcal{A}$ such that A entails B conditional on Y. We denote the transitive closure of $\vDash^*$ by $\vDash^{**}$. We read $A \vDash^{**} B$ as: there is a path from A to B.*

Note that $\mathcal{Y}$ can be an empty set where $A \vDash^*_{\mathcal{Y}} B$ becomes $A \vDash B$ as far as $A \neq \emptyset$ and $\overline{B} \neq \emptyset$, which holds if A and B are contingent, which means that they are neither W nor $\emptyset$. Let us mention a useful fact. If $A \vDash^*_{\mathcal{Y}} B$, it also holds that $\overline{B} \vDash^*_{\mathcal{Y}} \overline{A}$, and thus if $A \vDash^{**} B$, then $\overline{B} \vDash^{**} \overline{A}$. The following examples illustrate the above notions.

Example 2.1. *Let $A, B \subseteq W$ be logically independent, i.e., $A - B, B - A, A \cap B, \overline{A \cup B} \neq \emptyset$. Consider the following agendas:*
$\mathcal{A}_1 = \{A, B, \overline{A}, \overline{B}\}$;
$\mathcal{A}_2 = \{A, B, A \cap B, \overline{A}, \overline{B}, \overline{A \cap B}\}$;
$\mathcal{A}_3 = \{A, B, (A \cap B) \cup (\overline{A \cup B}), \overline{A}, \overline{B}, (A - B) \cup (B - A)\}$;
$\mathcal{A}_4 = \{A, B, A \cap B, A \cup B, \overline{A}, \overline{B}, \overline{A \cap B}, \overline{A \cup B}\}$.
It can be easily seen that $\mathcal{A}_1$ does not include any path. In the case of $\mathcal{A}_2$, for some issues, we have a conditional entailment relation such as $A \vDash^*_{\overline{A \cap B}} \overline{B}$ whereas for some other issues, say A and $\overline{A}$, we do not have a path, let alone a conditional entailment relation. In the case of $\mathcal{A}_3$, for some issues we have a conditional entailment relation such as $A \vDash^*_{(A \cap B) \cup (\overline{A \cup B})} B$ and $A \vDash^*_{(A-B) \cup (B-A)} \overline{B}$, which eventually lead to the fact that there is a path between every two issues. In the case of $\mathcal{A}_4$, the situation is similar to the case of $\mathcal{A}_3$.

The following two agenda conditions, namely path-connectedness and even- negatability have been studied a lot because they can characterize the most famous impossibility agendas in judgment aggregation. Let us first define path-connectedness.

Definition 2.12 (Path-connected Agenda). *An agenda $\mathcal{A}$ is path-connected(PC) iff for all contingent issues $A, B \in \mathcal{A}$ it holds that $A \vDash^{**} B$.*

Path-connectedness ensures that there is a path between every two issues. Next, to define even-negatable agenda, we need to introduce one more basic notion. A subset $\mathcal{Y} \subseteq \mathcal{A}$ is called minimally inconsistent when it is inconsistent and every proper subset of it is consistent. For any $\mathcal{Y}, \mathcal{Z} \subseteq \mathcal{A}$ with $\mathcal{Z} \subseteq \mathcal{Y}$, $\mathcal{Y}_{\neg \mathcal{Z}}$ is defined by $(\mathcal{Y} \backslash \mathcal{Z}) \cup \{\overline{A} |\ A \in \mathcal{Z}\}$ where $\overline{A}$ is the complement of A. Let us define an even-negatable agenda.

Definition 2.13 (Even-negatable Agenda). *An agenda $\mathcal{A}$ is even-negatable(EN) iff there is a minimally inconsistent set $\mathcal{Y} \subseteq \mathcal{A}$ such that $\mathcal{Y}_{\neg \mathcal{Z}}$ is consistent for some subset $\mathcal{Z} \subseteq \mathcal{Y}$ of even size.*

So even-negatability means that there is a minimally inconsistent set that can be made consistent by negating some even number of its element. Note that $\mathcal{Z} = \mathcal{Y}$ is allowed where $\mathcal{Y}_{\neg\mathcal{Z}} = \{\overline{A} | \ A \in \mathcal{Z}\}$. The definition of even-negatable agenda is a little bit involved. There is an algebraic characterization of even-negatability, which helps us decide whether a given agenda is even-negatable.[21]

Example 2.2 (Continued)**.** *The following facts can be easily proved: $\mathcal{A}_1$ is neither path-connected nor even-negatable; $\mathcal{A}_2$ is not path-connected but even-negatable. Indeed, $\{A, B, \overline{A \cap B}\}$ is the minimally inconsistent set, but $\{\overline{A}, \overline{B}, \overline{A \cap B}\}$ is consistent. (For another abstract method, see the last footnote.); $\mathcal{A}_3$ is path-connected but not even-negatable; $\mathcal{A}_4$ is path-connected and even-negatable.*

The following lemma shows that $\mathcal{A}$ being a non-trivial algebra — the underlying agenda condition for our previous results — is stronger than $\mathcal{A}$ being path-connected and even-negatable.

Lemma 2.7.

(1) Every non-trivial algebra is path-connected.

(2) Every non-trivial algebra is even-negatable.

Proof. (1) Let A, B be two contingent issues in a non-trivial algebra.

(Case 1) $\emptyset \neq A \subseteq B \neq W$.
Since (i) $A \neq \emptyset$, (ii) $B \neq W$ and (iii) $A \models B$ we have $A \models^* B$. Moreover, we have $B \models^*_{\{A \cup \overline{B}\}} A$, for (i) $(A \cup \overline{B}) \cap B = A \neq \emptyset$, (ii) $(A \cup \overline{B}) \cap \overline{A} = \overline{B} \neq \emptyset$ and (iii) $B \cap (A \cap \overline{B}) = A \subseteq A$. Thus, $A \models^* B$ and $B \models^* A$.

(Case 2) $A - B \neq \emptyset$ and $B - A \neq \emptyset$.
As in the proof of Lemma 2.4, let $v \in A - B$ and $w \in B - A$. We can use $C \in \mathcal{A}$ such that $\{v, w\} \subseteq C \neq W$ since $\mathcal{A}$ is a non-trivial algebra.[22] By the result of (Case 1), $A \models^* A \cap C \models^* C \models^* C \cap B \models^* B$. Thus, $A \models^{**} B$

[21] See Dokow & Holzman (2010a). For a given agenda $\mathcal{A}$, the set of all possible binary valuations on the agenda can be represented by a set of 0/1-vectors in $\{0, 1\}^{\frac{|\mathcal{A}|}{2}}$, which can be seen as a vector space over the field $\{0, 1\}$ on which componentwise addition is modulo 2. According to Proposition 2.1 and 4.2 in Dokow & Holzman (2010a), an agenda $\mathcal{A}$ is not even-negatable iff the set of all possible binary valuations on the agenda is an affine subspace of $\{0, 1\}^{\frac{|\mathcal{A}|}{2}}$ iff it is closed under addition of odd-tuples. For example, consider the agenda $\mathcal{A}_2$ in our example. For $(A, B, A \cap B)$, all possible valuations are $(1, 1, 1), (1, 0, 0), (0, 1, 0)$, and $(0, 0, 0)$. Since $(0, 1, 0) + (1, 0, 0) + (0, 0, 0) = (1, 1, 0)$, and $(1, 1, 0)$ is not a possible valuation, $\mathcal{A}_2$ is even-negatable. In contrast, for the agenda $\mathcal{A}_3$ with $(A, B, (A \cap B) \cup \overline{(A \cup B)})$, the set of possible valuations on it consists of $(1, 1, 1), (1, 0, 0), (0, 1, 0)$, and $(0, 0, 1)$. We can easily see that it is closed under addition of odd-tuples, thus $\mathcal{A}_3$ is not even-negatable.

[22] Recall that we can take as C the union of two elements in $\mathcal{A}$ one of which includes v and one of which includes w. We can find such two elements of which the union is not W because $\mathcal{A}$ is non-trivial. If there were no such two elements, it means that $(A - B) \cup (B - A) = W$ and thereby $\mathcal{A} = \{\emptyset, A, B, W\}$ where $B = \overline{A}$, which contradicts the non-triviality of $\mathcal{A}$.

(2) Let $\mathcal{A}$ be a non-trivial algebra. Then there are three non-empty issues $A, B, C \in \mathcal{A}$ that have no intersections with each other. Set $\mathcal{Y} = \{A, B\}$. Then $\mathcal{Y}$ is minimally inconsistent set. Set $\mathcal{Z} = \mathcal{Y}$. Then $\mathcal{Y}_{\neg\mathcal{Z}} = \{\overline{A}, \overline{B}\}$. Since $\overline{A} \cap \overline{B} \supseteq C \neq \emptyset$, $\mathcal{Y}_{\neg\mathcal{Z}}$ is consistent. $\square$

From now on, we add one more assumption on $\mathcal{A}$ that $\emptyset \notin \mathcal{A}$(and thereby $W \notin \mathcal{A}$).[23] Thus, our agenda $\mathcal{A}$ is a complement-closed finite non-empty set of some contingent subsets of the underlying set W of worlds.

The following lemma shows that the agenda condition of Lemma 2.4(IND and SYS) can be weakened to path-connectedness.

Lemma 2.8 (Path-connectedness and SYS)**.** *Let $\mathcal{A}$ be path-connected and a HA F satisfies UD. If F satisfies CDC, CP and IND,then it satisfies SYS.*

Proof. [24] By IND, we can let $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P}$ and $A \in \mathcal{A}$. We need to show that $G_A = G_B$ for all $A, B \in \mathcal{A}$. $\mathcal{A}$ being path-connected, it is enough to show that if $A \vDash^* B$, then for all $\vec{a}$, it holds that if $G_A(\vec{a}) = 1$, then $G_B(\vec{a}) = 1$.

Assume that $\{A\} \cup \mathcal{Y} \vDash B$ where $\mathcal{Y}$ is consistent with A and $\overline{B}$, i.e.,$A \cap \bigcap \mathcal{Y} \subseteq B$, $\bigcap \mathcal{Y} \cap A \neq \emptyset$ and $\bigcap \mathcal{Y} \cap \overline{B} \neq \emptyset$. By UD, we can take as an input a profile of probabilistic beliefs $\vec{P}$ that can be extended to a profile of probabilities $\vec{P'}$ such that $\vec{P'}(\bigcap \mathcal{Y} \cap A) = \vec{a}$ and $\vec{P'}(\bigcap \mathcal{Y} \cap \overline{B}) = \vec{1} - \vec{a}$, as illustrated in the following figure.[25]

---

[23]In the following, especially in Theorem 2.13(2) and Theorem 2.15(2), we will use some results of Nehring & Puppe(2010), where the agenda consists of contingent issues. To describe our proof more simply, we adopt that assumption.

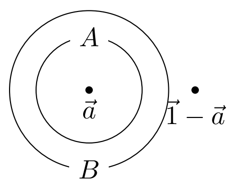
[24]In opinion pooling, collective beliefs should satisfy the additivity axiom $F(\vec{P})(A\dot{\cup}B) = F(\vec{P})(A) + F(\vec{P})(B)$, which does not hold under the assumption of collective rationality, CCS and CDC. Thus usually we cannot use proofs in OP or generalized OP of Dietrich & List (2017a). However consider the proof of Theorem 3(a) in Dietrich & List (2017a), which corresponds to this lemma. We can prove it using the following weaker properties than the additivity axiom:

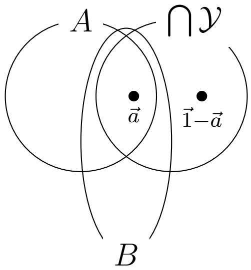
(i) if $A \subseteq B$, then $F(\vec{P})(A) \leq F(\vec{P})(B)$

(ii) if $F(\vec{P})(A) = 1$ and $F(\vec{P})(B) = 1$, then $F(\vec{P})(A \cap B) = 1$

Since HAs with CDC also satisfy them, we can apply the proof in heterogeneous belief aggregation as well.

[25]Recall $\bigcap \emptyset = W$ and we include the case where $\mathcal{Y} = \emptyset$. In this case, we have $A \subseteq B$, $A \neq \emptyset$ and $\overline{B} \neq \emptyset$, as illustrated in the following figure.





Notice that $\vec{P}(Y) = \vec{1}$ for all $Y \in \mathcal{Y}$, $\vec{P}(A) = \vec{a}$ and $\vec{P}(B) = \vec{a}$. By CP, $F(\vec{P})(Y) = 1$ for all $Y \in \mathcal{Y}$. Thus, if $G_A(\vec{a}) = F(\vec{P})(A) = 1$, then $F(\vec{P})(B) = G_B(\vec{a}) = 1$ by CDC, for $\{A\} \cup \mathcal{Y} \vDash B$. $\square$

Compared to Lemma 2.4, ZP is not required in this lemma. It is just because we assume $\emptyset, W \notin \mathcal{A}$ throughout this section. Recall that in Lemma 2.4 ZP is used only for $\emptyset$.

This lemma parallels the one in generalized opinion pooling of Dietrich & List (2017a): path-connectedness characterizes that if generalized OP satisfies CP and IND, then it satisfies SYS. In our lemma as well, its converse — if $\mathcal{A}$ is not path-connected, then there is a HA F on $\mathcal{A}$ satisfying CDC, CP, and IND but not SYS — also holds. The counterexample will be indicated in Theorem 2.10(2).

The following definitions and lemma will be needed to prove our succeeding main theorem.

Definition 2.14 (Non-simple Agenda and Pair-negatable Agenda)**.**

(1) An agenda $\mathcal{A}$ is non-simple(NS) iff there is a minimally inconsistent subset $\mathcal{Y} \subseteq \mathcal{A}$ with $|\mathcal{Y}| \geq 3$.

(2) An agenda $\mathcal{A}$ is pair-negatable iff there is a minimally inconsistent set $\mathcal{Y} \subseteq \mathcal{A}$ such that $\mathcal{Y}_{\neg Z}$ is consistent for some subset $\mathcal{Z} \subseteq \mathcal{Y}$ with $|\mathcal{Z}| = 2$.

Non-simple agenda can be used as a criterion for determining whether a given agenda has minimal complexity. Pair-negatable agenda is a special case of even-negatable agenda. The following lemma shows that a pair-negatable agenda is sufficient to be an even-negatable agenda, and a path-connected agenda already has a fairly complex structure.

Lemma 2.9.

(1) An agenda $\mathcal{A}$ is even-negatable iff $\mathcal{A}$ is pair-negatable.

(2) If an agenda $\mathcal{A}$ is path-connected, then it is non-simple.

Proof. (1) Dietrich & List (2013) Remark 1
(2) Dokow & Holzman (2010a) Claim 3.2 in Theorem 2.2 $\square$

Example 2.3 (Continued)**.** *The agendas $\mathcal{A}_1$ is simple while the agendas $\mathcal{A}_2$, $\mathcal{A}_3$ and $\mathcal{A}_4$ are non-simple. In $\mathcal{A}_3$, $\{A, B, \overline{A \cap B}\}$ is the minimally inconsistent set including at least three elements.*

Now we prove that the agenda being path-connected and even-negatable is the sufficient((1)) and necessary condition((2) and (3)) for the oligarchy result.

Theorem 2.10 (Agenda Condition for the Oligarchy Result)**.** *Let $\mathcal{A}$ be an agenda.*

(1) If $\mathcal{A}$ is path-connected and even-negatable then the only HAs on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, and IND are oligarchies.

(2) If $\mathcal{A}$ is not path-connected, then there is a HA on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, IND, and non-oligarchy. It also satisfies CCS.

(3) Let $|N| \geq 3$. If $\mathcal{A}$ is not even-negatable, then there is a HA on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, IND, and non-oligarchy. It also satisfies CCS.

Proof. (1) First of all, by Lemma 2.8 path-connectedness gives SYS and we can let $F(\vec{P})(A) = G(\vec{P}(A))$. Recall the following two facts about G in Theorem 2.5:

(Fact 1) if $\vec{a} \leq \vec{b}$ and if $G(\vec{a}) = 1$, then $G(\vec{b}) = 1$

(Fact 2) if $\vec{a} + \vec{b} - \vec{1} \geq \vec{0}$ and if $G(\vec{a}) = 1$ and $G(\vec{b}) = 1$, then $G(\vec{a} + \vec{b} - \vec{1}) = 1$

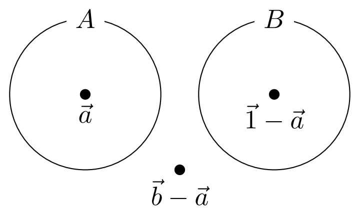
It is sufficient to show that if $\mathcal{A}$ is path-connected and even-negatable then (Fact 1) and (Fact 2) hold. It is because once we establish (Fact 1) and (Fact 2), we can prove the oligarchy result by exactly the same method as in Theorem 2.6.

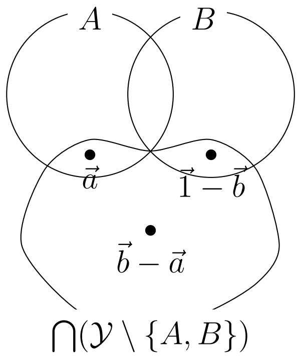
Lemma 2.9 enables us to prove (Fact 1′) and (Fact 2′) in the following.

(Fact 1′) If $\mathcal{A}$ is pair-negatable, then (Fact 1) holds.

Assume that $\mathcal{Y}(\subseteq \mathcal{A})$ is minimally inconsistent and $\mathcal{Y}_{\neg\{A,B\}}$ is consistent. Since $\bigcap \mathcal{Y}_{\neg\{B\}}$, $\bigcap \mathcal{Y}_{\neg\{A\}}$ and $\bigcap \mathcal{Y}_{\neg\{A,B\}}$ are not empty, they can have profiles of probabilities $\vec{a}$, $\vec{1} - \vec{b}$ and $\vec{b} - \vec{a}$, respectively. Thus, the profile $\vec{P}$ of probabilistic beliefs such that $\vec{P}(A) = \vec{a}$, $\vec{P}(B) = \vec{1} - \vec{b}$ and $\vec{P}(Y) = \vec{1}$ for all $Y \in \mathcal{Y} \setminus \{A, B\}$ can be an input by UD where $\vec{a} \leq \vec{b}$.(See the figure below.[26])

[26] Recall $\bigcap \emptyset = W$, and we include the case where $\mathcal{Y} \setminus \{A, B\} = \emptyset$. In this case, $A \cap B = \emptyset$, and it holds that A, B and $\overline{A} \cap \overline{B}$ are not empty.



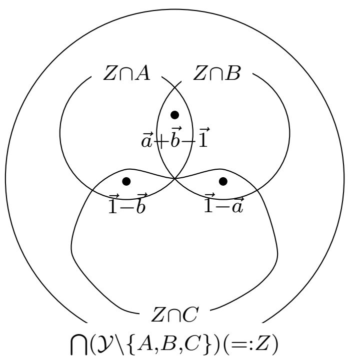


Notice that, by CP, $F(\vec{P})(Y) = 1$ for all $Y \in \mathcal{Y} \setminus \{A, B\}$ and $\mathcal{Y} \setminus \{B\} \vDash \overline{B}$, because $\mathcal{Y}$ is inconsistent. Therefore, from $F(\vec{P})(A) = G(\vec{a}) = 1$ we obtain $F(\vec{P})(\overline{B}) = G(\vec{b}) = 1$ by CDC.

(Fact 2′) If $\mathcal{A}$ is non-simple, then (Fact 2) holds.
Let $\mathcal{Y}$ be minimally inconsistent with $|\mathcal{Y}| \geq 3$, say $A, B, C \in \mathcal{Y}$. Since $\bigcap \mathcal{Y}_{\neg\{A\}}$, $\bigcap \mathcal{Y}_{\neg\{B\}}$ and $\bigcap \mathcal{Y}_{\neg\{C\}}$ are not empty, there can be a profile $\vec{P'}$ of probabilities on the algebra generated by $\mathcal{A}$ satisfying

$$\vec{P'}(\bigcap \mathcal{Y}_{\neg\{A\}}) = \vec{1} - \vec{a}, \vec{P'}(\bigcap \mathcal{Y}_{\neg\{B\}}) = \vec{1} - \vec{b} \text{ and } \vec{P'}(\bigcap \mathcal{Y}_{\neg\{C\}}) = \vec{a} + \vec{b} - \vec{1}$$

where $\vec{a} + \vec{b} - \vec{1} \geq \vec{0}$.



From this follows $\vec{P'}(\bigcap(\mathcal{Y} \setminus \{A, B, C\}) = 1$ and we have, by UD, the profile $\vec{P}$ of probabilistic beliefs on our domain such that

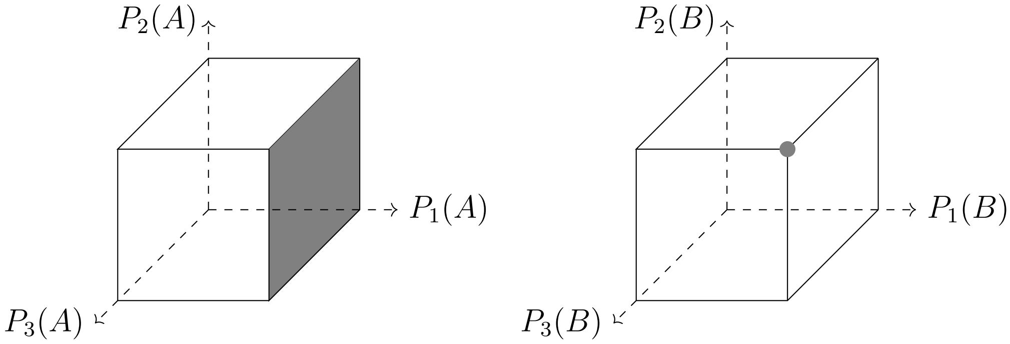
$$\vec{P}(A) = \vec{a}, \vec{P}(B) = \vec{b}, \vec{P}(\overline{C}) = \vec{a} + \vec{b} - \vec{1} \text{ and } \vec{P}(Y) = \vec{1} \text{ for all } Y \in \mathcal{Y} \setminus \{A, B, C\}$$

and hence, by CP, it follows that $F(\vec{P})(Y) = 1$ for all $Y \in \mathcal{Y}\setminus\{A, B, C\}$. Notice that we have $\mathcal{Y} \setminus \{C\} \vDash \overline{C}$, for $\mathcal{Y}$ is inconsistent. From this we conclude that if $G(\vec{a}) = F(\vec{P})(A) = 1$ and $G(\vec{b}) = F(\vec{P})(B) = 1$, then $F(\vec{P})(\overline{C}) = G(\vec{a}+\vec{b}-\vec{1}) = 1$ by CDC.

(2) Suppose that $\mathcal{A}$ is not path-connected. Then there are issues, say P, Q, such that there is no path from P to Q. Now we can partition $\mathcal{A}$ into two subsets $\mathcal{X}_1, \mathcal{X}_2 \subseteq \mathcal{A}$ such that $\mathcal{X}_1 = \{A \in \mathcal{A} |\ P \vDash^{**} A\}$ and $\mathcal{X}_2 = \{B \in \mathcal{A} |\ P \nvDash^{**} B\}$. Note that there is no path from an issue in $\mathcal{X}_1$ to an issue in $\mathcal{X}_2$. Let us define a HA F as follows:

For every issue $A \in \mathcal{X}_1$, $F(\vec{P})(A) = 1$ iff $P_1(A) = 1$.

For every issue $B \in \mathcal{X}_2$, $F(\vec{P})(B) = 1$ iff $\vec{P}(B) = \vec{1}$.



This represents the HA in (2), provided $N = \{1, 2, 3\}$. Grey points represents that 1 is assigned. The left figure is for any $A \in \mathcal{X}_1$ and the right one for any $B \in \mathcal{X}_2$.

It is easy to check that the above F satisfies UD, ZP, CP, IND, and non-oligarchy. Recall that we denote by $F(\vec{P})^{-1}(1)$ the set $\{C \in \mathcal{A} |\ F(\vec{P})(C) = 1\}$. Let $P_1^{-1}(1)$ denote the set $\{C \in \mathcal{A} |\ P_1(C) = 1\}$. Since $F(\vec{P})^{-1}(1) \subseteq P_1^{-1}(1)$ and $P_1^{-1}(1)$ is consistent, it follows that F satisfies CCS.

Now let us prove that F satisfies CDC. Suppose towards contradiction that there is an issue, say C, such that $F(\vec{P})(C) = 0$ but $F(\vec{P})^{-1}(1) \vDash C$. Then there is a minimal subset $\mathcal{T} \subseteq F(\vec{P})^{-1}(1)$ such that $\bigcap \mathcal{T} \subseteq C$. The remaining proof will be divided into three steps:
(i) there is an issue, say D, in $\mathcal{T} \cap \mathcal{X}_1$,
(ii) $C \in \mathcal{X}_2$,
(iii) there is a path from D to C
, which contradicts our assumption that there is no path from $\mathcal{X}_1$ to $\mathcal{X}_2$.

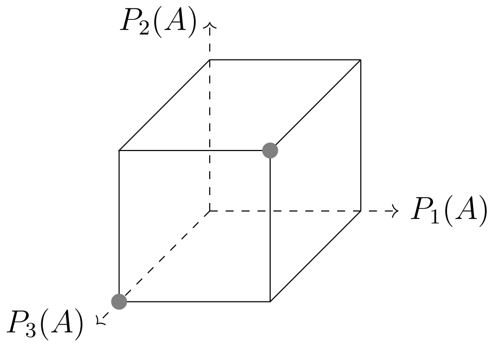
We first prove (i). The set $\mathcal{T} \cap \mathcal{X}_1$ is not empty since otherwise $\mathcal{T}$ would be included in $F(\vec{P})^{-1}(1) \cap \mathcal{X}_2 (= \{B \in \mathcal{X}_2 | \vec{P}(B) = \vec{1}\})$, which is actually deductively closed, and consequently $F(\vec{P})(C) = 1$, a contradiction.

We now turn to (ii). It suffices to show that $P_1(C) = 1$ for it would force $C \notin \mathcal{X}_1$, since otherwise $F(\vec{P})(C) = 1$. As $P_1^{-1}(1)$ includes $F(\vec{P})^{-1}(1)$, we have $P_1^{-1}(1) \vDash C$. As $P_1^{-1}(1)$ is deductively closed we obtain $C \in P_1^{-1}(1)$.

It remains to show (iii). Since $\mathcal{T} \cup \{\overline{C}\}$ is minimally inconsistent, we see that $\mathcal{T} \setminus \{D\}$ is consistent with D and $\overline{C}$. As $\mathcal{T} \vDash C$ we get $D \vDash^*_{\mathcal{T} \setminus \{D\}} C$, and hence $D \vDash^{**} C$.

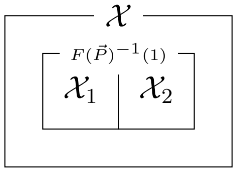
(3) Suppose that $\mathcal{A}$ is not even-negatable. Let us define a HA F as the following:

For every $A \in \mathcal{A}$, $F(\vec{P})(A) = 1$ iff $\vec{P}(A) = \vec{1}$ or $\vec{P}(A) = (0, 0, 1, ..., 1)$.



This illustrates the HA in (3) provided $N = 1, 2, 3$. Grey points represent that 1 is assigned.

It is easily seen that F satisfies UD, ZP, CP, IND, and non-oligarchy. Note that the above F will fail to satisfy ZP if we drop the assumption that $|N| \geq 3$. We will denote by $\mathcal{X}$ the set $\{A \in \mathcal{A} |\ \forall i \in N \setminus \{1, 2\}\ P_i(A) = 1\}$. Since $F(\vec{P})^{-1}(1) \subseteq \mathcal{X}$ and $\mathcal{X}$ is consistent, it follows that F satisfies CCS.



Now let us prove that F satisfies CDC. Suppose that, contrary to our claim, F does not satisfy CDC. We will prove that $\mathcal{A}$ is even-negatable. The remaining proof will be divided into two steps:
(i) we will find a minimally inconsistent subset $\mathcal{Y} \subseteq \mathcal{A}$ that has a subset $\mathcal{Z} \subseteq \mathcal{Y}$ and an element $X \in \mathcal{Y}$ such that $\mathcal{Z} \neq \emptyset$, and $X \notin \mathcal{Z}$,
(ii) we will prove that both $\mathcal{Y}_{\neg\mathcal{Z}}$ and $\mathcal{Y}_{\neg(\mathcal{Z} \cup \{X\})}$ are consistent.
Then we would get the desired result since either $\mathcal{Z}$ or $\mathcal{Z} \cup \{X\}$ must have even number elements.
We first construct $\mathcal{Y}$. Since F does not satisfy CDC, there is an issue, say B, such that $F(\vec{P})(B) = 0$ but $F(\vec{P})^{-1}(1) \vDash B$. From this we see that there is a minimal subset $\mathcal{T} \subseteq F(\vec{P})^{-1}(1)$ such that $\bigcap \mathcal{T} \subseteq B$. Set

$$\mathcal{Y} = \mathcal{T} \cup \{\overline{B}\}$$

, which is minimally inconsistent. Second, let

$$\mathcal{Z} = \mathcal{T} \cap (\mathcal{X}_2 :=) \{A \in \mathcal{A} | \vec{P}(A) = (0, 0, 1, ..., 1)\}$$

It is not empty set since otherwise $\mathcal{T}$ would be included in the set $\{A \in \mathcal{A} | \vec{P}(A) = \vec{1}\} (=: \mathcal{X}_1)$ that is deductively closed, and hence includes B, a contradiction. Third, we set

$$X = \overline{B}.$$

As $\mathcal{Z} \cup \{B\}$ is consistent, we obtain $\overline{B} \notin \mathcal{Z}$.
We now turn to (ii). To do this, we claim that $\forall i \in N \setminus \{1,2\}$ $P_i(B) = 1$. Since $\mathcal{X}$ includes $F(\vec{P})^{-1}(1)$ we have $\mathcal{X} \vDash B$. As $\mathcal{X}$ is deductively closed, we obtain the claim. From this and the fact that $F(\vec{P})(B) = 0$, it follows that

$$P_1(B) \neq 0 \text{ or } P_2(B) \neq 0 \tag{2.3}$$

and

$$P_1(\overline{B}) \neq 0 \text{ or } P_2(\overline{B}) \neq 0 \tag{2.4}$$

Now consider the set

$$\mathcal{Y}_{\neg\mathcal{Z}} = \{\overline{B}\} \cup \{\overline{A} |\ A \in \mathcal{T} \cap \mathcal{X}_2\} \cup (\mathcal{T} \cap \mathcal{X}_1)$$

$$\mathcal{Y}_{\neg(\mathcal{Z}\cup\{X\})} = \{B\} \cup \{\overline{A} |\ A \in \mathcal{T} \cap \mathcal{X}_2\} \cup (\mathcal{T} \cap \mathcal{X}_1)$$

Observe that $P_1(\overline{A}) = 1 = P_2(\overline{A})$ for every $A \in \mathcal{T} \cap \mathcal{X}_2$ and $P_1(C) = 1 = P_2(C)$ for every $C \in \mathcal{T} \cap \mathcal{X}_1$. From (2.4) we can assert that $P'_1(\bigcap Y_{\neg\mathcal{Z}}) > 0$ or $P'_2(\bigcap Y_{\neg\mathcal{Z}}) > 0$ where P'_1 and P'_2 are probability measures that are extensions of P_1 and P_2, respectively. Furthermore, from (2.3) we see that $P'_1(\bigcap \mathcal{Y}_{\neg(\mathcal{Z}\cup\{\overline{B}\})}) > 0$ or $P'_2(\bigcap \mathcal{Y}_{\neg(\mathcal{Z}\cup\{\overline{B}\})}) > 0$, and consequently $\bigcap \mathcal{Y}_{\neg\mathcal{Z}} \neq \emptyset$, and $\bigcap \mathcal{Y}_{\neg(\mathcal{Z}\cup\{X\})} \neq \emptyset$, which is the desired conclusion. □

Part (1) of the theorem generalizes Theorem 2.6 and shows that even if the agenda satisfies a weaker condition — path-connectedness and even-negatability — than a non-trivial algebra, the oligarchy result holds. If we look at the proof of Theorem 2.6 in detail, then we can observe that the agenda condition was used only to show (Fact 1) — upward closure of $G^{-1}(1)$ — and (Fact 2) — restricted downward closure of $G^{-1}(1)$. Thus, to prove (1), it is enough to derive (Fact 1) from even-negatability and (Fact 2) from path-connectedness. Agenda conditions are related only to (Fact 1) and (Fact 2) and once we see that they hold then we can carry over the proof of Theorem 2.6.

To obtain our result, we assumed IND, from which follows SYS under the assumption of path-connectedness (PC) by Lemma 2.8. Our proof also reveals that if we assume not IND but the stronger property of SYS, then non-simplicity(NS) is enough to get the oligarchy result. This shows that stronger properties of a HA lead to weaker agenda conditions for the oligarchy result. One more example is that if we assume MON, defined in Section 2.4, then we need not prove (Fact 1) thus, we need not require the agenda to be even-negatable (EN). It is because (Fact 1) is implied by SYS and MON. The following table demonstrates which agenda conditions are sufficient to obtain the oligarchy result with what properties of an HA.

	IND	SYS
without MON	PC, EN	NS, EN
with MON	PC	NS

It is of interest that the sufficient condition for our oligarchy result is the same as the one for the dictatorship and oligarchy results in judgment aggregation(Dokow & Holzman (2010a) and Dietrich & List (2008)). Concerning the relation between our proof and the proofs for judgment aggregation, the same comparison can be made as the comments in Section 2.5.

Part (2) suggests a counterexample of the oligarchy result when the agenda is not path-connected. Notice that the counterexample does not satisfy SYS and thus, it also serves as a counterexample of Lemma 2.8 as mentioned earlier.

In the proof, some methods in the proofs of Theorem 3 (c) in Dietrich & List (2008) and Claim 3.6 in Dokow & Holzman (2010a) are carried over to our domain. We extend the counterexample in Dietrich & List (2008) to our domain so that UD, ZP, CP, IND, CDC, and CCS are satisfied. Notice that the non-oligarchy of an extension follows from the non-oligarchy of the JA. (See (Observation 2) in Section 2.5) Our extension is a minimal extension satisfying monotonicity (MON). It is not that any extension works. For example, if we were to extend the JA so that 0 were assigned every input outside the domain of the JA, which violates MON, the extension would not satisfy CDC. (Concerning (ii) in the proof of part (2), from $P_1(C) = 1$ would not follow $C \notin \mathcal{X}_1$.) Indeed, since we do not exclude the case where $\mathcal{A}$ is even-negatable, the extension should satisfy MON.

Part (3) gives a counterexample of the oligarchy result when the agenda is not even-negatable. The counterexample is an extension of the counterexample discussed in Theorem 3 (b) in Dietrich & List (2008). In contrary to (2), it is not an extension satisfying MON, which is not imposed because the agenda is not even-negatable. In contrast, we do not exclude the agenda being path-connected, which imposes the counterexample to satisfy SYS.

The proof of part (3) is similar to the one in Dietrich & List (2008). But the last steps in our proof include novel ideas that are needed because of the difference between binary and probabilistic beliefs. In particular, step (ii) in the proof of part (3) utilized our own methodology: to show the consistency of a set of issues we use the fact that there is an agent who would assign a positive probability to the intersection of the issues.

To summarize, we proved that our oligarchy result can be obtained in more general settings than non-trivial algebras. Furthermore, we showed that the necessary and sufficient agenda condition for the oligarchy result with CDC and the dictatorship result with CCS and CCP in judgment aggregation — path-connectedness and even-negatability — is also the one for the oligarchy result with CDC in heterogeneous belief aggregation.

2.6.2 The Agenda Condition for the Triviality Result

As we said before, stronger properties of an HA yield weak agenda conditions. Thus, one might ask whether the agenda condition for the oligarchy result can be weakened, if we add AN and seek the agenda condition for the triviality result. In this section we will show that the agendas that yield the triviality result can be characterized by negation-connectedness, which is introduced in Dietrich & List (2018) and defined by the following.

Definition 2.15 (Negation-connected Agenda)**.** *An agenda $\mathcal{A}$ is negation-connected (NC) iff for every contingent issue $A \in \mathcal{A}$ it holds that $A \vDash^{**} \overline{A}$.*

So negation-connectedness means that every issue has a path to and from its complement. According to Proposition 1 in Dietrich & List (2018), the agenda being negation-connected is equivalent to the agenda being partitioned into subagendas each of which is path-connected where a non-empty subset of the agenda is called subagenda when it is closed under complementation.[27]

Example 2.4 (Continued)**.** *Consider the following agenda as well as $\mathcal{A}_1$-$\mathcal{A}_4$:*

$\mathcal{A}_5 = \{A, B, A \cap B, A \cup B, \overline{A}, \overline{B}, \overline{A \cap B}, \overline{A \cup B}, C, D, C \cap D, C \cup D, \overline{C}, \overline{D}, \overline{C \cap D}, \overline{C \cup D}\}$

where every two propositions among $A, B, C,$ and D are logically independent. The agenda $\mathcal{A}_1$ and $\mathcal{A}_2$ are not negation-connected while $\mathcal{A}_3$,$\mathcal{A}_4$,and $\mathcal{A}_5$ are negation-connected. Note that $\mathcal{A}_5$ is partitioned into two path-connected subagendas, namely $\{A, B, A \cap B, A \cup B, \overline{A}, \overline{B}, \overline{A \cap B}, \overline{A \cup B}\}$ and $\{C, D, C \cap D, C \cup D, \overline{C}, \overline{D}, \overline{C \cap D}, \overline{C \cup D}\}$.

The following lemma will be needed for the proof of the first part of the succeeding theorem. Part (1) allows us to consider the stronger condition, namely path-connectedness, than negation-connectedness to prove the triviality result. Part (2) will be used when the agenda is path-connected and not even-negatable.

Lemma 2.11. *(1) If the triviality results hold — i.e., the only HA on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, IND, and AN is the trivial one — for any path-connected agenda $\mathcal{A}$, then the same holds for any negation-connected agenda.*

(2) If an agenda $\mathcal{A}$ is not even-negatable, then for any minimally inconsistent subset $\mathcal{Y} \subseteq \mathcal{A}$ and any even-sized subset $\mathcal{Z} \subseteq \mathcal{Y}$ it holds that $\mathcal{Y}_{\neg \mathcal{Z}}$ is also minimally inconsistent.

Proof. (1) Assume that the triviality results hold for any path-connected agenda. Let $\mathcal{A}$ be a negation-connected agenda. Then $\mathcal{A}$ can be partitioned into path-connected subagendas, say, $\mathcal{A}_1, ..., \mathcal{A}_m$. Further, assume that a non-trivial HA F on $\mathcal{A}$ satisfies UD, CDC, ZP, CP, IND, and AN. Then, it can be easily seen that $F \restriction \mathcal{A}_k$, treated as a HA on $\mathcal{A}_k$, satisfies UD, ZP, CP, IND, and AN for any $k \leq m$. Thus, if we prove that $F \restriction \mathcal{A}_k$ satisfies CDC for any $k \leq m$, then we see that $F \restriction \mathcal{A}_k$ is the trivial function for all $k \leq m$, since $\mathcal{A}_k$ is path-connected. This implies that F is a trivial function.

[27]To see the reason why a negation-connected agenda is partitioned into path-connected subagendas, it is instructive to see that a path-relation in negation-connected agenda is actually an equivalence relation. Indeed, for any issues A, B in a negation-connected agenda $\mathcal{A}$ if $A \vDash^{**} B$, then $B \vDash^{**} \overline{B} \vDash^{**} \overline{A} \vDash^{**} A$, which implies that a path-relation in $\mathcal{A}$ satisfies symmetry; reflexivity and transitivity are trivial.

Thus, it is enough to show $F \restriction \mathcal{A}_k$ satisfy CDC for any $k \leq m$. Indeed, if $F \restriction \mathcal{A}_k(\vec{P})^{-1}(1) \vDash B$ for any issue $B \in \mathcal{A}$, then by CDC of F, $F(\vec{P'})(B) = 1$, $\vec{P'}$ being any extension of $\vec{P}$ to $\mathcal{A}$. Thus, we only need to show that $B \in \mathcal{A}_k$. Note that there is a finite subset $\mathcal{B} \subseteq F \restriction \mathcal{A}_k(\vec{P})^{-1}(1)(\subseteq \mathcal{A}_k)$ such that $\mathcal{B} \vDash B$. If we choose an issue $B_1 \in \mathcal{B}(\subseteq \mathcal{A}_k)$, we have $B_1 \vDash^*_{\mathcal{B}\setminus\{B_1\}} B$, which implies $B \in \mathcal{A}_k$ since $\mathcal{A}_k$ is path-connected and $B_1 \in \mathcal{A}_k$.

(2) Dokow & Holzman (2010a) Prop. 4.2., Dietrich & List (2021) Lem. 14. □

The following lemma will be needed for the proof of the second part of the succeeding theorem. This lemma looks technical but it is closely related to the notion of median point in the next section. Indeed, if $\mathcal{H}_0$ is the empty set, then $\bigcap \mathcal{M}$ is the set of all median points where $\mathcal{H}_0$ and $\mathcal{M}$ are defined in the following lemma.

Lemma 2.12. *Let $\mathcal{H}_0$ be the set $\{A \in \mathcal{A} | \, A \vDash^{**} \overline{A}$ and $\overline{A} \vDash^{**} A\}$. If $\mathcal{A}$ is not negation-connected, then there is a non empty subset $\mathcal{M} \subseteq \mathcal{A} \setminus \mathcal{H}_0$ such that for any minimally inconsistent set $\mathcal{Y} \subseteq \mathcal{A}$ it holds that $|\mathcal{Y} \cap \mathcal{M}| \leq 1$. Furthermore, for any minimally inconsistent set $\mathcal{Y} \subseteq \mathcal{A}$ intersecting $\mathcal{H}_0$ it holds that $|\mathcal{Y} \cap \mathcal{M}| = 0$. In addition, for $B \in \mathcal{A} \setminus \mathcal{H}_0$, it holds that $B \in \mathcal{M}$ iff $\overline{B} \notin \mathcal{M}$.*

Proof. Nehring & Puppe (2010) Proposition 3.1. □

Now let us prove the theorem. The first one states that negation-connectedness is the sufficient condition for the triviality result and the second one asserts that it is also the necessary condition.

Theorem 2.13 (Agenda Condition for the Triviality Result)**.**

(1) If $\mathcal{A}$ is negation-connected, then the only HA on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, IND, and AN is the trivial one.

(2) If $\mathcal{A}$ is not negation-connected, then there is a HA on $\mathcal{A}$ satisfying UD, CDC, ZP, CP, IND, AN, and non-triviality. It also satisfies CCS.

Proof. (1) It suffices to show the claim under the assumption of $\mathcal{A}$ being path-connected and not even-negatable. It is because we have Lemma 2.11 (1) and we can apply Theorem 2.10, if the agenda is even-negatable. Since $\mathcal{A}$ is path-connencted, by Lemma 2.8 we can set $F(\vec{P})(A) = G((\vec{P})(A))$. Moreover, by Lemma 2.9 (2), $\mathcal{A}$ is non-simple and thus we have (Fact 2) by (Fact 2′) in Theorem 2.10.

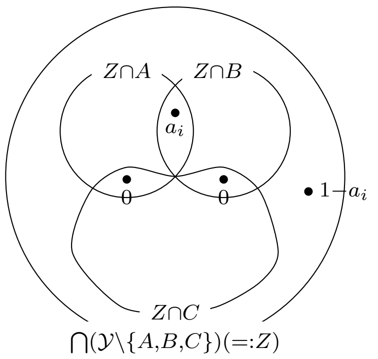
Now we will prove the following using $\mathcal{A}$ being non-simple and not even-negatable:

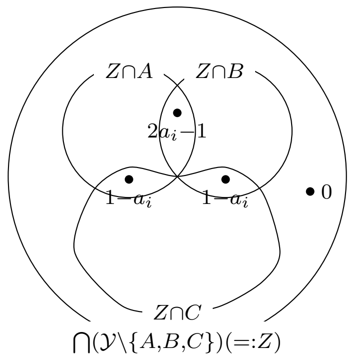
(Fact 1″) If $G(\vec{a}) = 1$, then $G(\vec{c}) = 1$ for all $\vec{c} \geq |2\vec{a} - \vec{1}|$.

where $|\vec{x}|$ is defined to be $(|x_i|)_i$. By non-simplicity and Lemma 2.11 (2), it follows that there is minimally inconsistent $\mathcal{Y}$ that has more than three elements, say $A, B, C \in \mathcal{Y}$ such that $\bigcap \mathcal{Y}_{\neg\{A\}}$, $\bigcap \mathcal{Y}_{\neg\{B\}}$, $\bigcap \mathcal{Y}_{\neg\{C\}}$ and $\bigcap \mathcal{Y}_{\neg\{A,B,C\}}$ are not empty. Consider individual probabilities P'_i on the algebra generated by $\mathcal{A}$ satisfying $P'_i(\bigcap(\mathcal{Y} \setminus \{A, B, C\})) = 1$ and $P'_i(A) = P'_i(B) = a_i$.

(Case 1) $2a_i \geq 1$

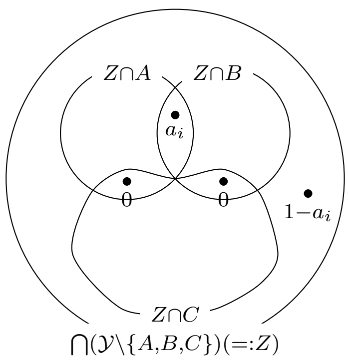
$P'_i(\bigcap \mathcal{Y}_{\neg\{A\}})$ and $P'_i(\bigcap \mathcal{Y}_{\neg\{B\}})$ might have the value between 0 and $1 - a_i$, from which follows that $P_i(\overline{C}) = P'_i(\bigcap \mathcal{Y}_{\neg\{C\}}) + P'_i(\bigcap \mathcal{Y}_{\neg\{A,B,C\}}) = 1 - P'_i(\bigcap \mathcal{Y}_{\neg\{A\}}) - P'_i(\bigcap \mathcal{Y}_{\neg\{B\}}) \in [2a_i - 1, 1]$ where P_i is the probabilistic belief that can be extended to P'_i. The left/right figure illustrates P'_i when $P'_i(\bigcap \mathcal{Y}_{\neg\{A\}})$ and $P'_i(\bigcap \mathcal{Y}_{\neg\{B\}})$ have the minimun/maximum value.

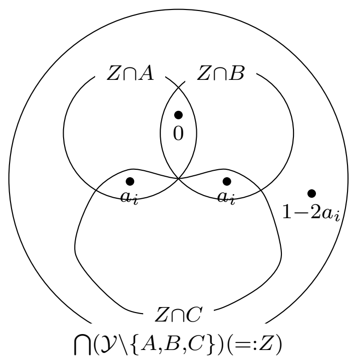




(Case 2) $2a_i < 1$

$P'_i(\bigcap \mathcal{Y}_{\neg\{A\}})$ and $P'_i(\bigcap \mathcal{Y}_{\neg\{B\}})$ might have the value between 0 and a_i, from which follows that $P_i(\overline{C}) \in [1 - 2a_i, 1]$.





Combining two cases we can assert that there can be probabilistic beliefs P_i satisfying

(i) $P_i(A) = P_i(B) = a_i$

(ii) $P_i(\overline{C}) \in [|2a_i - 1|, 1]$

(iii) $P_i(Y) = 1$ for all $Y \in \mathcal{Y} \setminus \{A, B, C\}$

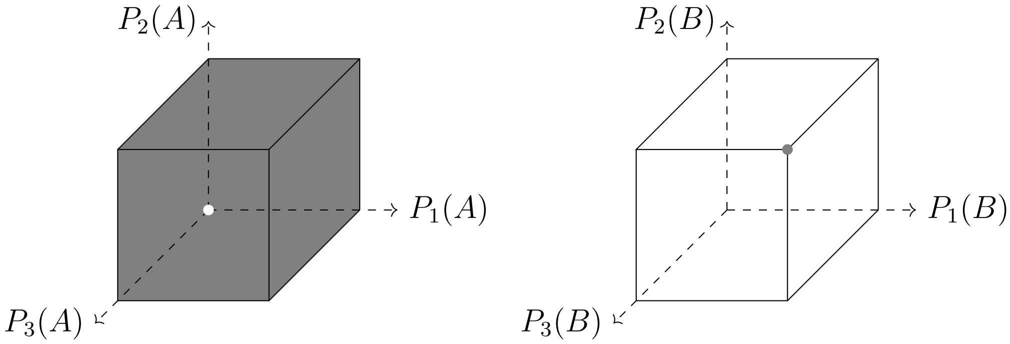
By CP, we have $F(\vec{P})(Y) = 1$ for all $Y \in \mathcal{Y} \setminus \{A, B, C\}$. Note that we have $\mathcal{Y} \setminus \{C\} \vDash \overline{C}$, for $\mathcal{Y}$ is inconsistent. From this we conclude that if $G(\vec{a}) = F(\vec{P})(A) = F(\vec{P})(B) = 1$, then $F(\vec{P})(\overline{C}) = G(\vec{c}) = 1$ for all $\vec{c} \in [|2\vec{a} - \vec{1}|, \vec{1}]$ by CDC, which completes the proof of (Fact 1$''$).

We now apply (Step 2) and (Step 3) in Theorem 2.5 again, with (Fact 1) replaced by (Fact 1$''$) to obtain the triviality result.

(2) Suppose that $\mathcal{A}$ is not negation-connected. Then there is a subset $\mathcal{M} \subseteq \mathcal{A}$ satisfying properties in the above lemma. Let us define a HA F as the following:

For every issue $A \in \mathcal{M}$, $F(\vec{P})(A) = 1$ iff $\vec{P}(A) \neq \vec{0}$

For every issue $B \notin \mathcal{M}$, $F(\vec{P})(B) = 1$ iff $\vec{P}(A) = \vec{1}$



This represents the HA in (2), provided $N = \{1, 2, 3\}$. Grey points represents that 1 is assigned. The left figure is for any $A \in \mathcal{M}$ and the right one for any $B \notin \mathcal{M}$.

It is easy to check that F satisfies UD, ZP, CP, IND, AN and non-triviality.

Let us prove that F satisfies CCS. Suppose, contrary to our claim, that $F(\vec{P})^{-1}(1)$ is inconsistent. Then there is a minimally inconsistent subset, say $\mathcal{Y} \subseteq F(\vec{P})^{-1}(1)$. By lemma 2.12 we have $|\mathcal{Y} \cap \mathcal{M}| \leq 1$ and hence we can find an issue, say C, such that $\mathcal{Y} \setminus \{C\}$ has no intersection with $\mathcal{M}$. This leads to $\vec{P}(Y) = \vec{1}$ for all issues $Y \in \mathcal{Y} \setminus \{C\}$ by the definition of F. Note that $\mathcal{Y}$ is inconsistent and so $\mathcal{Y} \setminus \{C\} \vDash \overline{C}$. From this we conclude that $\vec{P}(\overline{C}) = \vec{1}$ and thus $\vec{P}(C) = \vec{0}$, which contradicts $C \in \mathcal{Y} \subseteq F(\vec{P})^{-1}(1)$.

It remains to show that F satisfies CDC. Suppose that $F(\vec{P})^{-1}(1) \vDash D$. Then there is a subset $\mathcal{Y} \subseteq F(\vec{P})^{-1}(1)$ such that $\mathcal{Y} \cup \{\overline{D}\} (=: \mathcal{Z})$ is minimally inconsistent. By lemma 2.12 it holds that $|\mathcal{Z} \cap \mathcal{M}| \leq 1$. First, consider the case where $|\mathcal{Z} \cap \mathcal{M}| = 0$. Then for any $Y \in \mathcal{Y}$ it holds that $\vec{P}(Y) = \vec{1}$, and hence $\vec{P}(D) = \vec{1}$, which implies that $F(\vec{P})(D) = 1$. Now consider the case where $|\mathcal{Z} \cap \mathcal{M}| = 1$. If $\mathcal{Z} \cap \mathcal{M} = \{\overline{D}\}$, then similar arguments to the former case can be applied. In the case where $\overline{D} \notin \mathcal{M}$, we see that $D \in \mathcal{M}$ from the last part of lemma 2.12, because $|\mathcal{Z} \cap \mathcal{M}| = 1$ implies $\overline{D} \notin \mathcal{H}_0$ by the second part of lemma 2.12. Thus, we need to show that $\vec{P}(D) \neq \vec{0}$. Since $\vec{P}(\bigcap \mathcal{Y}) \leq \vec{P}(D)$, it is enough to show that $\vec{0} < \vec{P}(\bigcap \mathcal{Y})$. Denote by E the unique element in $\mathcal{Y} \cap \mathcal{M}$. By the definition of F, we have $\vec{P}(E) > \vec{0}$ and $\vec{P}(Y) = \vec{1}$ for all $Y \in \mathcal{Y} \setminus \{E\}$. Thus we have $\vec{P}(\bigcap \mathcal{Y}) > \vec{0}$. □

Part (1) shows that the triviality result holds if the agenda is negation-connected, which is a generalization of Theorem 2.5. The proof also reveals that if we assume SYS, non-simplicity (NS) is the sufficient condition to get the triviality result and we need neither even-negatability nor monotonicity (MON) unlike Theorem 2.10 (1).

	IND	SYS
with or without MON	NC	NS

Compared to the generalization of the oligarchy result, adding AN, we obtain the triviality result even under a weaker agenda condition: (i) instead of path-connectedness (PC), negation-connectedness (NC) is enough, and (ii) we have the triviality result even when the agenda is not even-negatable (EN). The difference indicated in (i) fulfils no role in finding the sufficient condition by Lemma 2.11, but the necessary condition is not path-connectedness but negation-connectedness, which will be discussed below. When the agenda is PC and EN, we can apply Theorem 2.10 since the oligarchy satisfying AN is the trivial one. Thus, we only need to focus on the cases where the agenda is PC and not EN.

When the agenda is assumed to be not EN, we run into the following difficulty: to show the triviality result, we used (Fact 1) — upward closure of $G^{-1}(1)$ —, which could be proved when the agenda is assumed to be EN. Our strategy here is to prove the weaker claim of (Fact 1″) than (Fact 1), because (Fact 1″) is enough to prove the triviality result. (Fact 1″) is weaker than (Fact 1) since not all vectors greater than $\vec{a}$ but only vectors greater than $|2\vec{a} - \vec{1}|$ are mapped to 1, if $G(\vec{a}) = 1$. In this sense, we call (Fact 1″) restricted upward closure of $G^{-1}(1)$.

One may ask why (Fact 1), which follows from even-negatability, is required for the oligarchy result, whereas (Fact 1″) is enough for the triviality result to hold. (Fact 1) and (Fact 1″) differ in the following way: for example, if $(0, 1, ..., 1)$ is mapped to 1 then $(r, 1, ..., 1)$ with any $r > 0$ is mapped to 1 by (Fact 1), but by (Fact 1″) only $(1, 1, ..., 1)$ is mapped to 1, which is enough to prove the triviality result. For the oligarchy result, we need (Fact 1), because without it we cannot deduce the fact that if $a_i = 1$ for all $i \in M$, M being the set of oligarchs, then $G(\vec{a}) = 1$ from $G((\delta_{i\in M})_i) = 1$, which can be proved not by (Fact 1″) but by (Fact 1).

As indicated in Table 2.1 (2″) at the beginning of this section, negation-connectedness is the same agenda condition for the triviality result on belief binarization, which is shown in Theorem 2* in Dietrich & List (2021). Their result can be restated using our terminology as follows: there is no belief binarization rule satisfying UD, CCS, CP, IND, CDC, and non-triviality iff the agenda is negation-connected. One might ask whether we can follow their proof for our theorem or the other way around. On the one hand, we cannot use their proof because whereas they deal with probabilistic beliefs, we are dealing with profiles of probabilistic beliefs and so their reasoning cannot be applied. In particular, for PC and not EN agendas we proved (Fact 1″), which is a new idea and what has done in our own way.

On the other hand, since we have not used the fact that $|N| \geq 2$ in the proof of part (1), our proof can be applied for the case where $|N| = 1$ as well, which is the same problem as belief binarization, as indicated in Section 2.5. Notice that Dietrich & List (2021) used CCS (stronger assumption) for belief binarization while our result assumes only ZP (weaker assumption). Therefore, if we apply our proof to the triviality result in belief binarization, then we can prove it using only ZP without requiring CCS. Thus we have the stronger claim(the triviality result with ZP) than

Dietrich & List (2021)'s claim (the triviality result with CCS). On top of that, we can use this argument for judgment aggregation, since anonymous independent judgment aggregation can be thought of as belief binarization, as indicated in Section 2.5. Thus, we also have (2′) in Table 2.1 there is no judgment aggregation satisfying UD, ZP, CP, IND, CDC, AN, and non-triviality iff the agenda is negation-connected. This is a stronger claim than the one in Dietrich & List (2008). They dealt with the agenda conditions — PC and EN — for the oligarchy result of judgment aggregation and derive the triviality result as a corollary under those agenda conditions. Our argument shows that the agenda condition can be weakened to negation-connectedness.

Part (2) gives the counterexample of the triviality result when the agenda is not negation-connected, which implies the agenda being not path-connected. The counterexample of the latter in Theorem 2.10 (2) does not work for the former, because it does not satisfy AN. Moreover, there will be no counterexample if we assume the agenda just not to be path-connected. This is the reason why we need to weaken path-connectedness to negation-connectedness even though they fulfill the same role concerning the sufficient agenda condition for the triviality result.

Our counterexample is an extension of the belief binarization rule in Theorem 2* in Dietrich & List (2021), which can be viewed as an anonymous JA. As said in Section 2.5, extending a counterexample in judgment aggregation to satisfy UD, CDC, ZP, CP, IND, AN is the key for heterogeneous belief aggregation, since non-triviality is directly satisfied. Notice that we have not excluded even-negatable agenda, thus MON must be satisfied and one can see that our example satisfies MON. However, it is not minimal among such extensions, which differs from the way of the extension — minimal extension with MON — in Theorem 2.10 (2). On the other hand, SYS is not forced since we do not have path-connectedness and our example does not satisfy SYS.

The upshot is that negation-connectedness is the necessary and sufficient condition not only for the triviality result in belief binarization and in judgment aggregation but also for our triviality result in heterogeneous belief aggregation.

2.6.3 The Agenda Condition for the Impossibility Result

In this section, we will show that agendas for the previous impossibility result can be characterized by blocked agendas. We begin by introducing this agenda condition.

Definition 2.16 (Blocked Agenda)**.** *An agenda $\mathcal{A}$ is blocked iff there is an issue $A \in \mathcal{A}$ such that $A \vDash^{**} \overline{A}$ and $\overline{A} \vDash^{**} A$.*

So a blocked agenda contains an issue that has a path to and from its complement. Recall that $\mathcal{H}_0$ is defined by the set $\{A \in \mathcal{A} | \; A \vDash^{**} \overline{A}$ and $\overline{A} \vDash^{**} A\}$. Then $\mathcal{A}$ is negation-connected iff $\mathcal{H}_0 = \mathcal{A}$, and $\mathcal{A}$ is blocked iff $\mathcal{H}_0 \neq \emptyset$. If $\mathcal{A}$ is negation-connected, then it is blocked.

Example 2.5 (Continued)**.** *The agenda $\mathcal{A}_1$ and $\mathcal{A}_2$ are unblocked while the agendas $\mathcal{A}_3, \mathcal{A}_4$, and $\mathcal{A}_5$ are blocked.*

The following definition and lemma will be needed for the succeeding theorem.

Definition 2.17 (Median Point). *Let $\mathcal{A}$ be an agenda on the set of worlds W. A world $m \in W$ is a median point iff for any minimally inconsistent subset $\mathcal{Y} \subseteq \mathcal{A}$, it holds that $|\{A \in \mathcal{Y} | \ m \in A\}| \leq 1$.*

So a median point is a world that is contained in at most one issue in every minimally inconsistent set.[28]

Example 2.6 (Continued). *In the agenda $\mathcal{A}_1$, every world is a median point; in $\mathcal{A}_2$, any world in $\overline{A} \cap \overline{B}$ is a median point.*

It is well-known in judgment aggregation that if a median point is guaranteed to exist, then we can easily construct an anonymous, complete, and consistent JA where a median point is thought of as a default collective judgment unless everybody believes the issue being true/false at the median point to be false/true.[29] The following lemma indicates that the agenda being unblocked is the necessary and sufficient condition for a median point to exist.

Lemma 2.14. *An agenda $\mathcal{A}$ is unblocked iff there is a median point.*

Proof. Nehring & Puppe (2010) Proposition 3.1 □

Now let us formulate and prove our last theorem.

Theorem 2.15 (Agenda Condition for Impossibility Result).

(1) If $\mathcal{A}$ is a blocked then there is no HA on $\mathcal{A}$ satisfying UD, CCP, CCS, CP, and IND.

(2) If $\mathcal{A}$ is not blocked, then there is a HA on $\mathcal{A}$ satisfying UD, CCP, CCS, CP, and IND. It also satisfies AN and non-dictatorship.

Proof. (1) Assume that $A \vDash^{**} \overline{A}$ and $\overline{A} \vDash^{**} A$. To obtain a contradiction, suppose that there is a HA F satisfying UD, CCP, CCS, CP, and IND. By IND, we can let $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P}$ and $A \in \mathcal{A}$. In Lemma 2.8, we made use of UD and CCD, which can be followed from CCP and CCS, and proved that for all $A, B \in \mathcal{A}$ if $A \vDash^{*} B$, then for all $\vec{a}$ it holds that if $G_A(\vec{a}) = 1$, then $G_B(\vec{a}) = 1$. Therefore, we have $G_A = G_{\overline{A}}$. Since every issue is contingent, $A, \overline{A} \neq \emptyset$. Hence there is a profile $\vec{P}$ of probabilistic belief satisfying $\vec{P}(A) = \vec{P}(\overline{A}) = 0.\vec{5}$ in our domain by UD, which yields $F(\vec{P})(A) = F(\vec{P})(\overline{A})$. This contradicts our assumption that F satisfies CCP

[28]Although this definition is a little bit involved, it has a geometrical meaning (see Nehring & Puppe (2007, 2005)): first of all, let us define a betweenness relation on worlds and introduce a useful notation. A world c is between a and b iff for all issues A such that $a, b \in A$, $c \in A$; $[a, b]$ denotes the set of all worlds between a and b. Then the geometrical definition of a median point says that a world m is a median point iff for any worlds v, w the three worlds m, v, w admit a median where a world m' is a median of m, v, w iff $m' \in [m, v] \cap [v, w] \cap [w, m]$. According to Lemma 5 in Nehring & Puppe (2005), the geometrical definition of a median point is equivalent to definition 2.17.

[29]See Nehring & Puppe (2007, 2005, 2010)

and CCS.

(2) Suppose that $\mathcal{A}$ is unblocked. By the above lemma, there is a median point m. Let us define a HA F as the following:

For every issue A with $m \in A$, $F(\vec{P})(A) = 1$ iff $\vec{P}(A) \neq \vec{0}$

For every issue B with $m \notin B$, $F(\vec{P})(B) = 1$ iff $\vec{P}(B) = \vec{1}$

It is easily seen that F satisfies UD, CP, IND, AN, and non-dictatorship.

Let us prove that F satisfies CCP. For any issue C we have $m \in C$ or $m \in \overline{C}$. W.l.o.g. we can assume that $m \in C$. If $\vec{P}(C) = \vec{0}$, then $\vec{P}(\overline{C}) = \vec{1}$ and hence $F(\vec{P})(\overline{C}) = 1$, and if otherwise, then $F(\vec{P})(C) = 1$.

It remains to show that F satisfies CCS. Suppose, contrary to our claim, that $F(\vec{P})^{-1}(1)$ is inconsistent. Then there is a minimally inconsistent subset $\mathcal{Y} \subseteq F(\vec{P})^{-1}(1)$. Since m is a median point, we see that $|\{A \in \mathcal{Y}|\ m \in A\}| \leq 1$. Hence we can find an issue, say D, such that none of the sets Y in $\mathcal{Y} \setminus \{D\}$ contains m. From the construction of F, it follows that $\vec{P}(Y) = 1$ for any issue $Y \in \mathcal{Y} \setminus \{D\}$, which implies that $\vec{P}(\bigcap(\mathcal{Y} \setminus \{D\})) = 1$. By the inconsistency of $\mathcal{Y}$, we obtain $D \cap [\bigcap(\mathcal{Y} \setminus \{D\})] = \emptyset$. From this we conclude that $\vec{P}(D) = 0$, which implies that $F(\vec{P})(D) = 0$, by the construction of F. This contradicts $F(\vec{P})(D) = 1$. □

Part (1) asserts that the impossibility result holds even when the agenda is blocked. Since CCS and CCP together are stronger than CDC, we obtain the impossibility result more easily — without AN and non-dictatorship and with more relaxed agenda condition. The proof shows that if we add SYS, the impossibility result holds even without CP and even when no agenda condition is assumed — e.g., even when $\mathcal{A} = \{A, \overline{A}\}$.

As indicated in Table 2.1 (4′) and (4″) at the beginning of this section, blocked agenda is also the agenda condition for the impossibility results on judgment aggregation with AN (Nehring & Puppe (2010)) and belief binarization (Dietrich & List (2018)). Concerning the relation, the same comments in Section 2.5 can be made.

Part (2) gives the counterexample that is an extension of the counterexample of Theorem 6 in Dietrich & List (2018). It is an extension that satisfies (MON), but not minimally so. This is the same way as the extension in Theorem 2.13 (2). The median point m in the proof of this theorem plays the same role as $\mathcal{M}$ in the proof of Theorem 2.13 (2). The only difference is that m is a world and $\mathcal{M}$ is a set of issues, which comes from the difference between assuming CDC and assuming CCS and CCP.

2.7 Conclusion

In this chapter we first established three impossibility results when the agenda is a non-trivial algebra, and then found the necessary and sufficient agenda condition for each result, i.e., provided three characterizations of impossibility agendas. The following table summarizes the results.

	Properties	Agenda Condition
Oligarchy Result	UD, CP, ZP, IND, CDC	PC, EN
Triviality Result	UD, CP, ZP, IND, CDC, AN	NC
Impossibility Result	UD, CP, IND, CCP, CCS	Blocked

We proved that when the agenda is as rich as a non-trivial algebra, independent(IND) heterogeneous belief aggregation satisfying collective deductive closure (CDC) yields oligarchies under certain conditions (UD, CP and ZP). Moreover, it is shown that adding anonymity (AN) leads to the trivial aggregation and adding collective completeness (CCP) and collective consistency (CCS) makes the aggregation impossible even without requiring AN. On top of that, we proved that these three impossibility results arise under different agenda conditions, as shown in the above table, and compared them with some impossibilities in judgment aggregation and belief binarization. We analyzed similarities and differences between our proofs and other related proofs and concluded that the problem of heterogeneous belief aggregation is not reduced to the other related problems. Moreover, we showed that our methods can be applied to other similar impossibilities.

Chapter 3

Threshold-based Heterogeneous Belief Aggregation

Heterogeneous belief aggregation is the problem about how to aggregate individuals' probabilistic beliefs on the logically connected issues into the group's binary belief. In the last chapter, we introduced the properties like collective deductive-closure and independence, and showed that they yield trivial results under certain conditions when the agenda is complex. Now we move toward studying specific aggregation rules and their properties. In particular, we will mainly address rules satisfying collective deductive-closure. We divide heterogeneous belief aggregation rules into two categories: (1) *direct rules* and (2) *collective belief binarization* given an opinion pooling procedure. Both will be explored in this chapter, where we address *threshold-based rules*, and in the next chapter, where we investigate rules based on distance and epistemic utility.

$$\begin{tikzcd} \text{Individuals' probabilistic Beliefs } \vec{P} \arrow[rr, "\text{Opinion}", "\text{Pooling } f"'] \arrow[drr, "(1) \text{ Direct Threshold Rules}"'] & & \text{Group's probabilistic Belief } f(\vec{P}) \arrow[d, "(2) \text{ Binarization (Threshold Rules)}"] \\ & & \text{Group's binary Belief } F(\vec{P}) \end{tikzcd}$$

3.1 Introduction

As mentioned in the last chapter, the problem of heterogeneous belief aggregation can be seen, on the one hand, as a generalization of judgment aggregation, but on the other hand, as a generalization of belief binarization. Indeed, there have been many articles indicating the structural parallel between judgment aggregation and belief binarization (Douven & Romeijn (2007), Dietrich & List (2018, 2021), Chandler (2013), Cariani (2016)). Their core consists in the fact that a quota of individuals believing an issue corresponds to the group's probability of the issue. We advance beyond this parallel: collective belief binarization can also be interpreted as heterogeneous belief aggregation as well, if we presuppose an opinion pooling procedure. This is why we treat belief binarization as the second category of heterogeneous belief aggregation mentioned above.

In judgment aggregation (e.g., Dietrich & List (2007)) or belief binarization, many rules associate the resulting binary belief on an issue with a high quota of individuals believing the issue or high probability of the issue. These rules are based on some sorts of thresholds to identify a high probability and called *threshold-based approaches*. The most typical one is given by the well-known *Lockean thesis* (LT^t), which suggests that an agent (in heterogeneous belief aggregation, the group) should believe an issue iff its probability exceeds a given threshold t. However, the lottery paradox, as illustrated in the introductory chapter,[1] shows that unless $t = 1$, LT^t does not ensure rationality (defined in the last chapter by consistency and deductive closure). There have been the following suggestions to resolve the paradox: relaxing closure under conjunction (Kyburg (1961), Leitgeb (2021)); relaxing probabilism (Spohn (2009)); relaxing the Lockean thesis (Leitgeb (2017a), Lin & Kelly (2012b)).

This is the same problem as the discursive dilemma in quota rules and as the triviality results in independent heterogeneous belief aggregation rules demonstrated in the last chapter. Confronting this problem, we will preserve probabilism and closure under conjunction, and relax LT^t.

To find a way to relax LT^t, our classification of thresholds in this chapter will be helpful. We will classify thresholds into four groups according to two criteria. The first criterion is whether thresholds are applied to probabilities of events (sets of worlds) to determine the belief set (the set of believed events) — we call them *event thresholds* (event ths.) —, or they are applied to probabilities of worlds to obtain the belief core (the smallest believed event of which supersets constitute the belief set) — we call them *world thresholds* (world ths.). The second criterion is whether thresholds depend on the input — we call them *local* — or not — we call them *global*. The thresholds t in LT^t correspond to global event thresholds. Among the various attempts to weaken LT^t, we note two different threshold-based approaches. One is to

[1]The lottery paradox shows: Consider a fair 1,000-ticket lottery that has only one winning ticket. An agent believes that ticket i will not win for each i since her probabilistic belief is a uniform distribution and her Lockean threshold 0.99. She also believes that one ticket will win. Suppose her belief is deductively closed. Then it is deduced that no ticket will win, so her belief is inconsistent. Unless $t = 1$, there is an example where the Lockean thesis does not generate consistent and deductively closed binary beliefs.

use local thresholds — e.g., the rules that generate belief states (pairs of probability and binary belief) satisfying the *Humean thesis* with a parameter r (HT^r, $r \in [\frac{1}{2}, 1)$) in Leitgeb's *stability theory of belief*, which says that an issue is believed iff its conditional probability on every issue of which complement is not believed is above r (Leitgeb (2013, 2014, 2017a)). And the other is to use world thresholds — e.g., Lin & Kelly's *Camera Shutter rules* with a parameter s (CS^s, $s > 1$), which collect the worlds whose probability ratio to the maximal probability is above $\frac{1}{s}$ as the elements of the belief core (Lin & Kelly (2012a, 2012b).

Our study on threshold-based heterogeneous belief aggregation is found on this basis. As mentioned at the beginning of this chapter, we create two categories of heterogeneous belief aggregation: *collective belief binarization* that is combined with a given opinion pooling procedure, and *direct rules* that do not go through a procedure to form the group's probabilistic belief. Accordingly, the classification of thresholds is applied not only to collective belief binarization but also to the direct rules. This will be the first part of this chapter. We will introduce criteria to classify types of thresholds, and according to these criteria, we will classify relevant threshold-based rules in both categories. And then, we will show that they are characterized by certain properties, e.g., various forms of independence and monotonicity — *global or local, and event-wise or world-wise monotonicity* that correspond to global or local, and event- or world-thresholds.

In the second part of this chapter, we will narrow our scope of research and focus only on *collective belief binarization* and *local monotonicity.* The key properties of binarization rules here will be local event-wise monotonicity and local world-wise monotonicity. We add non-emptiness of every belief set to the former and non-emptiness of every belief core to the latter, and call them *Lockean* and *coherent*, respectively. We relate these properties to other properties of binarization rules, being stable and rational-likely (r-likely). We say that a rule is stable if every belief state generated by the rule satisfies $\mathrm{HT}^{\frac{1}{2}}$ and every resulting binary belief is consistent. According to Leitgeb's stability theory of belief, being stable can be shown to be equivalent to the conjunction of being Lockean and coherent. Being r-likely means that every resulting binary belief has a non-empty belief core with probability above $\frac{1}{2}$. This property is weaker than being stable but neither weaker nor stronger than being coherent. We will review the inclusion relationship between those properties. We not only give an overview in this particularly structured way but also deliver novel results: we will provide geometrical characterizations of some properties — especially coherence — using the Voronoi diagram, which enables one to check whether a rule satisfies the properties easily.

After addressing properties of belief binarization, we will study specific rules and identify which rules satisfy the above properties. Our focus will be on rational binarization rules, of which outputs have non-empty belief cores and thereby can be represented in a geometrical space according to our method, which will be proposed. The rules to be examined here are: (i) the rule called $\mathrm{HT}^r(\mathrm{S})$, which picks the smallest non-empty belief core inducing a belief state satisfying HT^r (Leitgeb (2013), Cariani (2016), Thorn (2018), Wright (2018)), (ii) the Camera shutter rule (CS^s) (Lin & Kelly (2012a, 2012b)), and (iii) the coherent core-threshold rule (CCT^g) that picks the smallest coherent non-empty belief core of which probability is above a given threshold g (Cantwell & Rott (2019)). Note that all of these rules are not only rational but also coherent, and thus can be seen as local world threshold rules. We add one more rational binarization rule, the *distance minimization function* with the squared Euclidean distance ($\mathrm{DM(SE)}^+$), which will be discussed in full detail in the next chapter, in order to examine whether the rule satisfies the above properties and to compare the results with other threshold-based rules. We will study these rules in terms of the properties, and the main question will be whether each of these rules satisfies each of the above properties.

Last but not least, we will investigate one more property and whether each rule in the above satisfies it: we suggest studying the notion of convexity concerning belief binarization. Furthermore, we devise various forms of convexity and extend binarization methods beyond functions to relations, correspondences and ordinalizations. Which binarization methods satisfy which forms of convexity will also be shown.

To conclude, we will classify and characterize threshold-based heterogeneous belief aggregation, which is closely related to belief binarization. We will provide a structured framework that brings together various belief binarization properties and rules so that they can be evaluated and compared together. In particular, we provide geometrical characterizations of some properties and utilize these to answer the above questions. Moreover, we propose to study the property of convexity and devise a new belief binarization rule $\mathrm{DM(SE)}^+$. In our framework, properties and rules are derived from various existing binarization methods. For instance, the stability theory of belief provides the property of being stable that is distinguished from the rule of $\mathrm{HT}^r(\mathrm{S})$. From CS^s and CCT^g, we abstract coherence, from CCT^g r-likeliness, and from $\mathrm{DM(SE)}^+$ convexity. This approach gives insights into compatibility between binarization methods and provides interesting questions:

- Are CS^s, CCT^g and $\mathrm{DM(SE)}^+$ stable?
- Are CS^s and $\mathrm{DM(SE)}^+$ r-likely?
- Is $\mathrm{DM(SE)}^+$ coherent?
- Are $\mathrm{HT}^r(\mathrm{S})$, CS^s and CCT^g convex?

We will answer all of these questions in this chapter.

The rest of this chapter is organized as follows: In Section 3.2, we classify relevant direct rules and collective belief binarization rules based on thresholds, and characterize them by various forms of monotonicity and some other properties. In Section 3.3, we review local monotonicity and related properties of belief binarization, and provide geometrical characterization. In Section 3.4, we introduce various binarization rules and investigate whether each rule satisfies each property. We devote a separate section to the property of convexity in Section 3.5, based on joint work with Chisu Kim. We formulate various kinds of convexity norms and examine whether the binarization functions and some other binarization methods satisfy certain kinds of convexity requirements. Finally, we conclude with some topics for future work in Section 3.6

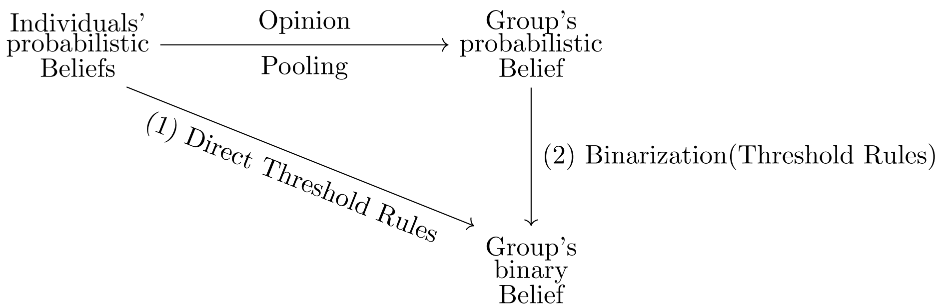
3.2 Classification and Characterizations of Threshold-based HAs

Threshold-based rules are generally characterized by some kinds of monotonicity. For example, quota rules for judgment aggregation can be characterized using a kind of monotonicity.[2] In heterogeneous belief aggregation as well, we can consider rules based on some thresholds. In this section we systematically introduce and classify various kinds of threshold-based heterogeneous aggregators (HAs), and investigate exactly which kinds of monotonicity and what other properties of HAs are needed to characterize them. It will help evaluate and compare various kinds of threshold-based rules in one theoretical frame.

3.2.1 Classification of Threshold-based HAs

We first set the notation and terminology that will be needed in this chapter. We follow the ones in the last chapter except for the fact that we assume, from now on, the set W of possible worlds to be finite, and the agenda $\mathcal{A}$ (the set of issues, which we call events in this chapter) to be the *powerset* $\mathcal{P}(W)$ of W so that the probabilities of the singleton set of each world are well-defined, which will be used for world-threshold rules.

Recall that $N := \{1, ..., n\}$ is a set of individuals($n \geq 2$); for each individual $i \in N$, P_i denotes i's probability function on $(W, \mathcal{P}(W))$, and we write $\vec{P}$ or $(P_i)_i$ for a profile $(P_1, ..., P_n)$ of individual probability functions; an *opinion pooling function (OP)* f is defined to be a function taking $\vec{P}$ and returning a probability function $f(\vec{P})$ on $(W, \mathcal{P}(W))$; a function $F : \vec{P} \mapsto F(\vec{P})$ is called a *heterogeneous aggregator (HA)* on $(W, \mathcal{P}(W))$, with $F(\vec{P})$ being a binary belief on $(W, \mathcal{P}(W))$, i.e., a function from $\mathcal{P}(W)$ to $\{0, 1\}$.



Now we introduce and classify some relevant *threshold-based heterogeneous aggregators*. First of all, we can categorize heterogeneous aggregators (not only threshold-based ones but also in general) into two groups according to whether they can be represented by a combination of an opinion pooling function (OP) and a *binarization rule*, which is defined as follows.

[2]See Dietrich & List (2007).

Definition 3.1 (Binarization Rule). *Let W be a finite non-empty set of possible worlds and $\mathcal{A}$ be an algebra on W. A function G mapping a probability function P on $(W, \mathcal{A})$ in a given domain to a binary belief $G(P)$ on $(W, \mathcal{A})$ is called a binarization rule(BR).*

The HAs in the first group that do not go through an opinion pooling procedure and thus do not form a group's probability are called *direct threshold rules*. The other group is called *pooling + threshold-based binarization*. If a HA F belongs to the second group, then we have $F = G \circ f$ for some OP f and some BR G, and we write $F = f + G$.

The next two criteria we are suggesting to classify threshold-based HAs pertain to types of threshold. The first is whether the threshold is applied to probabilities of events (in this case we call it an *event-th.*) or probabilities of worlds (in this case we call it a *world-th.*); we might have to believe all and only the events whose probability exceeds the event-th. or we might have to believe the set of all worlds with probability above the world-th. and its supersets. The second is whether the threshold depends on inputs, i.e., profiles of individual probabilities. If it depends on that, it is called a *local* (event- or world-) th., and if not, a *global* (event- or world-) th.

On the basis of these three criteria, we introduce eight classes of threshold-based HAs. Let us formulate the first four classes of rules precisely.

Definition 3.2 (Direct Threshold Rules). *Let F be a HA on $(W, \mathcal{P}(W))$ with the universal domain $\mathbb{P}(W)$, which denotes the set of all probability functions on $(W, \mathcal{P}(W))$.*

(i) F is called a direct threshold rule with global event-th. if for each $A \in \mathcal{P}(W)$ there exist $(\rhd_{A,i})_{i \in N} \in \{>, \geq\}^N$ and $(t_{A,i})_i \in [0,1]^N$ such that for all $\vec{P} \in \mathbb{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \text{ iff } P_i(A) \rhd_{A,i} \ t_{A,i} \text{ for all } i \in N;$$

(ii) F is called the one with global world-th. if there exist $(\rhd_{w,i})_{(w,i) \in W \times N} \in \{>, \geq\}^{W \times N}$ and $(s_{w,i})_{(w,i) \in W \times N} \in [0,1]^{W \times N}$ such that for all $\vec{P} \in \mathbb{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \text{ iff } A \supseteq \{w \in W | P_i(w) \rhd_{w,i} \ s_{w,i} \text{ for all } i \in N\}$$

for all $A \in \mathcal{P}(W)$;

(iii) F is called the one with local event-th. if for each $\vec{P} \in \mathbb{P}(W)$ there exist $(\rhd_{\vec{P},i})_{i \in N} \in \{>, \geq\}^N$ and $(t_{\vec{P},i})_{i \in N} \in [0,1]^N$ such that for all $A \in \mathcal{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \text{ iff } P_i(A) \rhd_{\vec{P},i} \ t_{\vec{P},i} \text{ for all } i \in N;$$

(iv) F is called the one with local world-th. if for each $\vec{P} \in \mathbb{P}(W)$ there exist $(\rhd_{\vec{P},i})_{i \in N} \in \{>, \geq\}^N$ and $(s_{\vec{P},i})_{i \in N} \in [0,1]^N$ such that

$$F(\vec{P})(A) = 1 \textit{ iff } A \supseteq \{w \in W | P_i(w) \rhd_{\vec{P},i} \ s_{\vec{P},i} \textit{ for all } i \in N\}$$

for all $A \in \mathcal{P}(W)$.

Notice that in the definition of the rules with *event-ths.*($t_{A,i}$ and $t_{\vec{P},i}$), thresholds are applied to probabilities of events and used to determine the *belief set* $F(\vec{P})^{-1}(1)(:= \{A \in \mathcal{P}(W) | F(\vec{P})(A) = 1\})$. By contrast, *world-thresholds*($s_{A,i}$ and $s_{\vec{P},i}$) are applied to obtain the *belief core* of the binary belief $F(\vec{P})$, defined as usual as follows:

Definition 3.3 (Belief Core)**.** *Let* W *be a finite non-empty set and* $Bel : \mathcal{P}(W) \to \{0, 1\}$ *be a binary belief. A subset* $B \subseteq W$ *is called the belief core of* Bel *if for all* $A \in \mathcal{P}(W)$

$$Bel(A) = 1 \textit{ iff } A \supseteq B$$

If there exists a belief core B of a binary belief Bel, then it is unique ($B = \bigcap Bel^{-1}(1)$), and we say that Bel has a (unique) belief core B or equivalently, that B induces Bel.

Simply put, by the event-th. rules, the events with probability being above (either greater than or not less than) the event-ths. form the belief set. And by the world-th. rules, the worlds with probability being above the world-ths. constitute the belief core.

Event- and world- th. can both be global or local. *Local* thresholds ($t_{\vec{P},i}$ or $s_{\vec{P},i}$) may vary with probability profiles $\vec{P}$. On the contrary, *global* ones ($t_{A,i}$ or $s_{w,i}$) do not depend on probability profiles. Global ones may differ according to events A/worlds w. We call the thresholds *uniform*, if all events/worlds have the same threshold. Notice that in our definition, local thresholds are all uniform by design, since the notion of the *local* threshold rule would otherwise be empty in the sense that every rule would be seen as a local non-uniform threshold rule. It is important to notice one more dependency of thresholds in direct threshold rules. We include the general cases where individuals may have different values of thresholds, which is why we add the subscript i to all kinds of thresholds for direct rules.

Now consider the inequality symbol $\rhd$ in each definition, which designates either $\geq$ or $>$. The distinction between the two might not so relevant when it comes to local thresholds: consider a HA with local event-ths. and fix $\vec{P}$ and i. It is clear that for each $t_{\vec{P},i} \neq 0$ there exists $t'_{\vec{P},i}(\neq 1)$, and for each $t'_{\vec{P},i} \neq 1$ there exists $t_{\vec{P},i}(\neq 0)$ such that

$$\{A \in \mathcal{P}(W) | P_i(A) \geq t_{\vec{P},i}\} = \{A \in \mathcal{P}(W) | P_i(A) > t'_{\vec{P},i}\}$$

because $\mathcal{P}(W)$ is finite. A similar reasoning applies to the local world-th. rules because W is finite: for each $s_{\vec{P},i} \neq 0$ there exists $s'_{\vec{P},i}$, and for each $s'_{\vec{P},i} \neq 1$ there exists $s_{\vec{P},i}$ such that

$$\{w \in W | P_i(w) \geq s_{\vec{P},i}\} = \{w \in W | P_i(w) > s'_{\vec{P},i}\}$$

On the contrary, as far as global threshold rules are concerned, an inequality with $\geq$ and one with $>$ cannot be represented by each other. For example, there exists no $t'_{A,i}$ satisfying $\{P_i \in \mathbb{P}(W) | P_i(A) \geq t_{A,i}\} = \{P_i \in \mathbb{P}(W) | P_i(A) > t'_{A,i}\}$, with A, i and $t_{A,i}$ being fixed, because $\mathbb{P}(W)$ is not discrete. Therefore, to deal with global thresholds, the distinction is not superfluous.

One more important feature regarding (strict) inequalities is that each individual i, each event A/world w (in the case of global event-/world- th.) and each profile $\vec{P}$ (in the case of local th.) can have a different kind of inequality — either strict or not —, just as each of them can have a different value of threshold. (The subscripts represent the dependency.) This enables us to study the most general cases.

Lastly, concerning direct threshold rules, note that the inequalities in each definition should be satisfied for all individuals. This can be generalized by relaxing "all individuals" to certain proportion of individuals, but in this research we will focus on the basic case of unanimously exceeding each one's threshold, which is easy to characterize.

Summarizing, we define direct threshold-based rules using the following inequalities to determine the belief set in the case of event-th. and the belief core in the case of world-th.

	event-th.	world-th.
global	$P_i(A) \rhd_{A,i} \; t_{A,i}$ for all i	$P_i(w) \rhd_{w,i} \; s_{w,i}$ for all i
local	$P_i(A) \rhd_{\vec{P},i} \; t_{\vec{P},i}$ for all i	$P_i(w) \rhd_{\vec{P},i} \; s_{\vec{P},i}$ for all i

We next turn to *pooling + threshold-based binarization*. Let f be an opinion pooling function (OP) with the universal domain $\mathbb{P}(W)$. Now we first form the group's probability $f(\vec{P}) \in \mathbb{P}(W)$ and then use it as an input of a binarization rule (BR) G, which outputs a binary belief. In this way, the composition of an OP and a BR can be used as a method of heterogeneous belief aggregation. Since the individuals' opinions are collected into a group's probability, we do not need to evaluate whether each individual's probability is above some threshold. Instead, we evaluate the group's probability $f(\vec{P}) \in \mathbb{P}(W)$. Thus, on substituting $P_i(A)$ and $P_i(w)$ with $f(\vec{P})(A)$ and $f(\vec{P})(w)$, respectively, as shown in the table, we obtain the definition of the four classes of *pooling + threshold-based Binarization*.

	event-th.	world-th.
global	$f(\vec{P})(A) \rhd_A \; t_A$	$f(\vec{P})(w) \rhd_w \; s_w$
local	$f(\vec{P})(A) \rhd_{\vec{P}} \; t_{\vec{P}}$	$f(\vec{P})(w) \rhd_{\vec{P}} \; s_{\vec{P}}$

The definition can be stated in full detail as follows.

Definition 3.4 (Pooling(f) + Threshold-based Binarization)**.** *Let f be an OP on $(W, \mathcal{P}(W))$ with the universal domain $\mathbb{P}(W)$ and F be a HA on $(W, \mathcal{P}(W))$ with the universal domain $\mathbb{P}(W)$.*

(i) F is called a pooling(f) + threshold-based binarization with global event-th. if for each $A \in \mathcal{P}(W)$ there exist $\rhd_A \in \{>, \geq\}$ and $t_A \in [0, 1]$ such that for all $\vec{P} \in \mathbb{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \textit{ iff } f(\vec{P})(A) \rhd_A \; t_A;$$

(ii) F is called the one with global world-th. if there exist $(\rhd_w)_{w\in W} \in \{>, \geq\}^W$ and $(s_w)_{w\in W} \in [0,1]^W$ such that for all $\vec{P} \in \mathbb{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \textit{ iff } A \supseteq \{w \in W | f(\vec{P})(w) \rhd_w \; s_w\}$$

for all $A \in \mathcal{P}(W)$;

(iii) F is called the one with local event-th. if for each $\vec{P} \in \mathbb{P}(W)$ there exist $\rhd_{\vec{P}} \in \{>, \geq\}$ and $t_{\vec{P}} \in [0,1]$ such that for all $A \in \mathcal{P}(W)$ it holds that

$$F(\vec{P})(A) = 1 \textit{ iff } f(\vec{P})(A) \rhd_{\vec{P}} \; t_{\vec{P}};$$

(iv) F is called the one with local world-th. if for each $\vec{P} \in \mathbb{P}(W)$ there exist $\rhd_{\vec{P}} \in \{>, \geq\}$ and $s_{\vec{P}} \in [0,1]$ such that

$$F(\vec{P})(A) = 1 \textit{ iff } A \supseteq \{w \in W | f(\vec{P})(w) \rhd_{\vec{P}} \; s_{\vec{P}}\}$$

for all $A \in \mathcal{P}(W)$.

The contrast between event-ths. and world-ths., and the one between global and local thresholds can be made in the same way as in the direct threshold rules: (i) and (iii) utilize event-ths. so that the group believes the events with high group probability, and (ii) and (iv) apply world ths. so that the belief core consists of the worlds with high group probability. The thresholds and the types of the inequalities might be different for each event and for each world in (i) and (ii), whereas they might be different for each probability profile in (iii) and (iv). The point that allowing two types of inequality is not redundant can be applied here as well.

If we focus on the relation between the group's probability $f(\vec{P})$ and the resulting binary belief, we can see that this definition embraces many binarization methods in the literature on belief binarization. The famous *Lockean thesis* combined with an OP f is no more than the rules with global even-th. The *Camera Shutter rules (CS^s rules)* of Lin & Kelly (Lin & Kelly (2012b)) can be seen as a special kind of binarization rule with local world-ths. Last but not least, the *Humean Thesis(HT^r)* in Leitgeb's stability theory of belief (Leitgeb (2013, 2014, 2017a)) can generate binarization rules with local event-ths. which can be proven to be also seen as rules with local world-ths. We will discuss them in detail in Section 3.4.

3.2.2 Characterizations of Threshold-based HAs

Now let us characterize the eight classes of threshold-based HAs. We begin by introducing properties of HAs that we will need to characterize the classes.

Nonskeptism and Deductive Closure The first property concerns the notion of rationality. Recall that, in the last chapter, we introduced *rationality* of binary belief, i.e., *consistency* and *deductive closure*, which can be defined, as W is finite, by the following:

(i) Bel is *consistent* iff $\bigcap Bel^{-1}(1) \neq \emptyset$ (the belief set should not entail a contradiction).

(ii) Bel is *deductively-closed* iff $Bel^{-1}(1)$ contains W and it is closed under intersection (any intersection of two sets in the belief set should also be in the belief set) and closed under superset (any superset of a set in the belief set should also be in the belief set).

In this chapter we add one more, the notion of *non-skepticism.* A *non-skeptical* binary belief means that the belief set (the set of all believed events) is non-empty. It is a plausible requirement because believing nothing implies not believing even W (a tautology), which is absurd. Note that deductive closure entails non-skepticism. However, considering them separately will be needed in this chapter.

Definition 3.5 (Rationality). *Let W be a finite non-empty set and $Bel : \mathcal{P}(W) \to \{0, 1\}$ be a binary belief.*

(1) Bel is non-skeptical iff $Bel^{-1}(1) \neq \emptyset$;

(2) Bel is rational iff Bel is consistent and deductively-closed.

(3) A HA F is non-skeptical(NSK)/collectively consistent(CCS)/collectively deductively-closed (CDC)/rational iff $F(\vec{P})$ is non-skeptical/consistent/deductively-closed/rational for all $\vec{P}$ in the domain of F.

CDC will be used especially to characterize the rules with world-ths. because these rules presuppose that the resulting binary belief $F(\vec{P})$ should have a belief core, which is closely related to CDC, as seen in the following well-known statement:

$$Bel \text{ is } \textit{deductively-closed} \text{ iff } Bel \text{ has a } \textit{belief core } B$$

which can be proven by letting $B := \bigcap Bel^{-1}(1)$ (the intersection of all believed events). This gives a 1-1 correspondence between binary beliefs with CDC and subsets of W. To be more precise, for each binary belief Bel with CDC, there exists a unique belief core $B(:= \bigcap Bel^{-1}(1))(\subseteq W)$, and each subset $B \subseteq W$ induces a unique binary belief function Bel with CDC, which assigns 1 to all and only the supersets of B. Thus, assuming CDC allows us to abuse notation and represent a binary belief by its belief core, a subset of W, if it causes no confusion. Let us mention one more relevant point that we will use often. If we add consistency to the left side of the equivalence statement, we obtain the following:

$$Bel \text{ rational iff } Bel \text{ has a non-empty } \textit{belief core}$$

We can extend properties concerning binary beliefs to HAs, as (3) states, by demanding that the output of HAs should always have the properties.

Independence and Neutrality Next we turn to *independence* for global threshold rules, and *neutrality* for local threshold rules. In the last chapter, we investigated *event-wise independence* and *event-neutrality*, whose tension with CDC (collective deductive closure) leads to the triviality results when the agenda is sufficiently complex. In this section, we introduce, in addition, six different kinds of independence and neutrality — two for the direct rules with world ths. and four for pooling + threshold-based binarization. Roughly speaking, *independence* means that to decide whether an event/a world belongs to the belief set/the belief core, the probability assigned to only that event/world matters, regardless of the probabilities of all other events/worlds. By contrast, *neutrality* means that every event/world is determined to be believed or not believed by the same rule, and thereby all events/worlds are considered equally. We can use this notion to characterize *uniform thresholds*.

Before providing the formal definition, let us mention some points needed to understand the definition: recall that world-ths. are used to determine the belief core, and from now on, bear in mind that $F(\vec{P})(\overline{w}) = 0$ means that w is in the belief core of $F(\vec{P})$.

Here are the definitions of independence and neutrality to characterize the direct rules and the ones of *f-independence and f-neutrality* for pooling(f) + binarization.

Definition 3.6 (Independence and Neutrality). *Let F be a HA on $(W, \mathcal{P}(W))$. F is called*

(i) (event-wise) independent(IND) if for every $A \in \mathcal{A}$, there is a function G_A such that $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P}$ in the domain of F;

(ii) world-wise independent(IND^w) if for every $w \in W$, there is a function G_w such that $F(\vec{P})(\overline{w}) = G_w(\vec{P}(w))$ for all $\vec{P}$ in the domain of F;

(iii) event-neutral(eNEU) if for every $\vec{P} \in \mathbb{P}(W)$, there is a function $G_{\vec{P}}$ such that $F(\vec{P})(A) = G_{\vec{P}}(\vec{P}(A))$ for all $A \in \mathcal{A}$;

(iv) world-neutral(wNEU) if for every $\vec{P} \in \mathbb{P}(W)$, there is a function $G_{\vec{P}}$ such that $F(\vec{P})(\overline{w}) = G_{\vec{P}}(\vec{P}(w))$ for all $w \in W$.

Definition 3.7 (f-Independence and f-Neutrality). *Let f be a OP and F be a HA on $(W, \mathcal{P}(W))$. F is called*

(i) f-(event-wise) independent(f-IND) if for every $A \in \mathcal{A}$, there is a function G_A such that $F(\vec{P})(A) = G_A(f(\vec{P})(A))$ for all $\vec{P}$ in the domain;

(ii) f-world-wise independent(f-IND^w) if for every $w \in W$, there is a function G_w such that $F(\vec{P})(\overline{w}) = G_w(f(\vec{P})(w))$ for all $\vec{P}$ in the domain;

(iii) f-event-neutral(eNEU) if for every $f(\vec{P})(\in \mathbb{P}(W))$, there is a function $G_{f(\vec{P})}$ such that $F(\vec{P})(A) = G_{f(\vec{P})}(f(\vec{P})(A))$ for all $A \in \mathcal{A}$;

(iv) *f-world-neutral(wNEU) if for every* $f(\vec{P})(\in \mathbb{P}(W))$*, there is a function* $G_{f(\vec{P})}$ *such that* $F(\vec{P})(\overline{w}) = G_{f(\vec{P})}(f(\vec{P})(w))$ *for all* $w \in W$*.*

Alternatively, F is defined to be IND/IND^w/eNEU/wNEU iff (i′)/(ii′)/(iii′)/(iv′) where

(i′) for every $A \in \mathcal{A}$, if $\vec{P}(A) = \vec{P'}(A)$, then $F(\vec{P})(A) = F(\vec{P'})(A)$ for all $\vec{P}, \vec{P'}$

(ii′) for every $w \in W$, if $\vec{P}(w) = \vec{P'}(w)$, then $F(\vec{P})(\overline{w}) = F(\vec{P'})(\overline{w})$ for all $\vec{P}, \vec{P'}$

(iii′) for every $\vec{P}$, if $\vec{P}(A) = \vec{P}(B)$, then $F(\vec{P})(A) = F(\vec{P})(B)$ for all $A, B \in \mathcal{A}$

(iv′) for every $\vec{P}$, if $\vec{P}(w) = \vec{P}(v)$, then $F(\vec{P})(\overline{w}) = F(\vec{P})(\overline{v})$ for all $w, v \in W$

While (i′) and(ii′) state that the same individual probabilities of an event/a world yield the same collective belief in the event/world, (iii′) and (iv′) assert that if two events/worlds have the same individual probabilities, then the collective belief in them should be the same. Similarly, equivalent definitions can be formulated for being f-IND/f-IND^w/f-eNEU/f-wNEU.

Monotonicity Now we formalize various kinds of *monotonicity* that play a central role to characterize any threshold-based rules. Since we separately defined independence and neutrality in the above, here we define *strict* monotonicity (It will be shown that strict monotonicity taken together with independence and neutrality amounts to *monotonicity*). Informally, strict monotonicity means the following: assume that an event/a world is in the belief set/belief core of the resulting collective binary belief of a probability profile. The first two kinds ((i) and (ii) in Definition 3.8) of strict monotonicity say that other probability profiles with greater probability values of the event/the world should yield the same result. In contrast, the other two ones ((iii) and (iv)) require that other events/worlds with greater probability values in the probability profile should also be in the belief set/belief core. *f-strict-monotonicity* in Definition 3.9 can be explained in the similar way if we replace a probability profile by a collective probability, which is the output of an OP f.

Definition 3.8 (Strict-Monotonicity)**.** *Let* F *be a HA on* $(W, \mathcal{P}(W))$*.* F *is called*

(i) *strict-monotone(SMON) if for every* $A \in \mathcal{P}(W)$*, if for some* $i \in N$*,* $P_i(A) < P'_i(A)$ *and for all* $j \neq i$ $P_j(A) = P'_j(A)$*, and if* $F(\vec{P})(A) = 1$*, then* $F(\vec{P'})(A) = 1$ *for all* $\vec{P}, \vec{P'}$ *in the domain;*

(ii) *worldwise strict-monotone(SMON^w) if for every* $w \in W$*, if for some* $i \in N$*,* $P_i(w) < P'_i(w)$ *and for all* $j \neq i$ $P_j(w) = P'_j(w)$*, and if* $F(\vec{P})(\overline{w}) = 0$*, then* $F(\vec{P'})(\overline{w}) = 0$ *for all* $\vec{P}, \vec{P'}$ *in the domain;*

(iii) *event-strict-monotone(eSMON) if for every* $\vec{P}$ *in the domain, for some* $i \in N$*,* $P_i(A) < P_i(B)$ *and for all* $j \neq i$ $P_j(A) = P_j(B)$*, and if* $F(\vec{P})(A) = 1$*, then* $F(\vec{P})(B) = 1$ *for all* $A, B \in \mathcal{A}$*;*

(iv) world-strict-monotone(wSMON) if for every $\vec{P}$ in the domain, for some $i \in N$, $P_i(v) < P_i(w)$ and for all $j \neq i$ $P_j(v) = P_j(w)$, and if $F(\vec{P})(\overline{v}) = 0$, then $F(\vec{P})(\overline{w}) = 0$ for all $w, v \in W$.

Definition 3.9 (f-Strict-Monotonicity)**.** *Let f be a OP and F be a HA on $(W, \mathcal{P}(W))$. F is called*

(i) f-strict-monotone(f-SMON) if for every $A \in \mathcal{P}(W)$, $f(\vec{P})(A) < f(\vec{P'})(A)$ and if $F(\vec{P})(A) = 1$, then $F(\vec{P'})(A) = 1$ for all $\vec{P}, \vec{P'}$ in the domain;

(ii) f-world-wise strict-monotone(f-SMON^w) if for every $w \in W$, $f(\vec{P})(w) < f(\vec{P'})(w)$ and if $F(\vec{P})(\overline{w}) = 0$, then $F(\vec{P'})(\overline{w}) = 0$ for all $\vec{P}, \vec{P'}$ in the domain;

(iii) f-event-strict-monotone(f-eSMON) if for every $\vec{P}$ in the domain, $f(\vec{P})(A) < f(\vec{P})(B)$ and if $F(\vec{P})(A) = 1$, then $F(\vec{P})(B) = 1$ for all $A, B \in \mathcal{A}$;

(iv) f-world-strict-monotone(f-wSMON) if for every $\vec{P}$ in the domain, $f(\vec{P})(v) < f(\vec{P})(w)$ and if $F(\vec{P})(\overline{v}) = 0$, then $F(\vec{P})(\overline{w}) = 0$ for all $w, v \in W$.

Alternatively, in (i) of Definition 3.8 the condition that "for some $i \in N$, $P_i(A) < P'_i(A)$ and for all $j \neq i$ $P_j(A) = P'_j(A)$" can be replaced by

$$\text{“}\vec{P}(A) \leq \vec{P'}(A) \text{ and } \vec{P}(A) \neq \vec{P'}(A)\text{”}$$

where $\leq$ and $\neq$ between two vectors are understood as component-wise comparison. The same can be said for (ii)-(iv) of Definition 3.8 as well. This indicates that combining independence or neutrality with strict-monotonicity yields monotonicity — e.g., IND plus SMON amounts to the statement that for every $A \in \mathcal{A}$,

$$\text{if } \vec{P}(A) \leq \vec{P'}(A), \text{ and if } F(\vec{P})(A) = 1, \text{ then } F(\vec{P'})(A) = 1$$

for all $\vec{P}, \vec{P'}$ in the domain, which we call *monotonicity*(MON). For other cases, the same reasoning can be applied.

Conjunctiveness Finally, we now introduce *conjunctiveness*, which will be crucial to characterize direct threshold rules. It basically means the following: assume that two probability profiles ($\vec{P}$ and $\vec{P'}$) generate beliefs in a given event (A). And consider any probability profile ($\vec{P''}$) of which values of the event consists of the individual-wise minimum probabilities of the event out of the two profiles. Then it also should generate a belief in the event. For example, if $\vec{P}(A), \vec{P'}(A), \vec{P''}(A)$ are given by Table 3.1 and $F(\vec{P})(A) = F(\vec{P'})(A) = 1$, then $F(\vec{P''})(A) = 1$. In the table, $minR$ denotes the minimum value in R for any $R \subseteq \mathbb{R}$.

Similar notion can be made for a world in a belief core as well. There also can be other kinds of conjunctiveness: if two events/worlds are believed/in the belief core, then so are any events/worlds whose probabilities are individual-wise minimum values out of probabilities of the first two events/worlds.

individual	$\vec{P}(A)$	$\vec{P'}(A)$	$\vec{P''}(A) = (min\{P_i(A), P'_i(A)\})_i$
1	0.8	0.7	0.7
2	0.6	0.9	0.6

Table 3.1

	(1) Direct Threshold Rules	(2) Pooling(f)+Th.Binarization
(i) global event-th	IND, SMON, Conj	f-IND, f-SMON
(ii) global world-th	CDC, IND^w, SMON^w, Conj^w	CDC, f-IND^w, f-SMON^w
(iii) local event-th	eNEU, eSMON, eConj	f-eNEU, f-eSMON
(iv) local world-th	CDC, wNEU, wSMON, wConj	CDC, f-wNEU, f-wSMON

Table 3.2: Characterizations of Threshold Rules

These notions are needed, because in the direct threshold rules, we demand that *not a part of but all of individuals' probabilities* should exceed their thresholds. This can be seen as the requirement that the individual-wise minimum in $\{\vec{P}(A)|F(\vec{P})(A) = 1\}$ should also exceed each individual's threshold. Here is the formal definition.

Definition 3.10 (Conjunctiveness)**.** *Let F be a HA on $(W, \mathcal{P}(W))$. F is called*

(i) conjunctive(Conj) if $F(\vec{P})(A) = 1$ and $F(\vec{P'})(A) = 1$, then $F(\vec{P''})(A) = 1$ for any $\vec{P''}$ such that $P''_i(A) = min\{P_i(A), P'_i(A)\}$ for all i;

(ii) world-wise conjunctive(Conj^w) if $F(\vec{P})(\overline{w}) = 0$ and $F(\vec{P'})(\overline{w}) = 0$, then $F(\vec{P''})(\overline{w}) = 0$ for any $\vec{P''}$ such that $P''_i(w) = min\{P_i(w), P'_i(w)\}$ for all i;

(iii) event-conjunctive(eConj) if $F(\vec{P})(A) = 1$ and $F(\vec{P})(B) = 1$, then $F(\vec{P})(C) = 1$ for any C such that $P_i(C) = min\{P_i(A), P_i(B)\}$ for all i;

(iv) world-conjunctive(wConj) if $F(\vec{P})(\overline{v}) = 0$ and $F(\vec{P})(\overline{w}) = 0$, then $F(\vec{P})(\overline{u}) = 0$ for any u such that $P_i(u) = min\{P_i(v), P_i(w)\}$ for all i.

Characterizations of Threshold Rules We are now ready to characterize eight classes of threshold rules introduced in the last section in terms of properties in this section. Assume that F satisfies UD of HAs and f satisfies UD of OPs, that is, both has the domain $\mathbb{P}(W)$. Our results can be presented by Table 3.2.

As illustrated in the table, every threshold rule satisfies some kind of strict monotonicity combined with independence in the case of global thresholds and with neutrality in the case of local thresholds, which we call monotonicity. For example, in

(2)(i), the rules are characterized by f-MON, which is defined as f-IND plus f-SMON, and in (2)(iii) by f-eMON, which is f-eNEU plus f-eSMON. To characterize the rules with world threshold rules we additionally need CDC because they presuppose the existence of a belief core. Lastly, we need to add Conj to characterize direct threshold rules. The following two theorems make this formally precise.

Theorem 3.1 (Characterization of Direct Threshold Rules). *(i) The direct threshold rules with global event-ths. are fully characterized by UD, IND, SMON and CONJ;*

(ii) so are the ones with global world-ths. by UD, CDC, IND^w, SMON^w and CONJ^w;

(iii) so are the ones with local event-ths. by UD, eNEU, eSMON and eCONJ;

(iv) so are the ones with local world-ths. by UD, CDC, wNEU, wSMON and wCONJ

Proof. (i) It is obvious that the rule satisfies the properties. For the other direction, we need to find $t_{A,i}$ and $\rhd_{A,i}$ for each $i \in N$, with A being fixed. By UD and IND, we can let $F(\vec{P})(A) = G_A(\vec{P}(A))$ for all $\vec{P} \in \mathbb{P}(W)$. In the case of ${G_A}^{-1}(1) = \emptyset$, let $t_{A,i} := 1$ and $\rhd_{A,i} :=>$ for all $i \in N$. Otherwise, let $t_{Ai} := \inf\{a_i | \vec{a} \in {G_A}^{-1}(1)\}$ for each $i \in N$. We divide N into two subgroups N_1 and N_2 where N_1 is the set of individuals j such that the set $\{a_j | \vec{a} \in {G_A}^{-1}(1)\}$ has the infimum and N_2 is the set of the rest individuals. For every individual $j \in N_1$, set $\rhd_{A,j} :=\geq$ and for other individuals $k \in N_2$ define $\rhd_{A,k} :=>$. First observe that if for some $j \in N_1$, $x_j < t_{A,j}$ or for some $k \in N_2$, $x_k \leq t_{A,k}$, then $G_A(\vec{x}) = 0$ by the definition of infimum. What is left is to show that $G_A(\vec{y}) = 1$ for all $\vec{y}$ such that for every $j \in N_1$ and $k \in N_2$, $y_j \geq t_{A,j}$ and $y_k > t_{A,k}$. Since $t_{A,i}$ is the infimum of the i-th components of the vectors in ${G_A}^{-1}(1)$ and we have SMON and IND, it follows that for every i there is a vector $\vec{a}^i$ in ${G_A}^{-1}(1)$ such that the i-th component is y_i. Note that $\vec{a}^i$ has the following form where $n := |N|$:

$$\begin{aligned}
\vec{a}^1 &= (y_1, \quad a^1_2, \quad a^1_3, \quad \dots \quad, \quad a^1_n) \\
\vec{a}^2 &= (a^2_1, \quad y_2, \quad a^2_3, \quad \dots \quad, \quad a^2_n) \\
&\dots \\
\vec{a}^n &= (a^n_1, \quad a^n_2, \quad a^n_3, \quad \dots \quad, \quad y_n)
\end{aligned}$$

By iterated application of CONJ, we have $G_A((\min\{a^i_l | i \in N\})_{l \in N}) = 1$. Since we have $\min\{a^i_l | i \in N\} \leq y_l$ for all $l \in N$, by SMON and IND, we get $G_A(\vec{y}) = 1$, as desired.

(ii) We can prove this in much the same way, the only differences being (a) $F(\vec{P})(\overline{w}) = G_w(\vec{P}(w))$ where $F(\vec{P})(\overline{w}) = 0$ iff $w \in B$ by CDC, with B being the belief core of $F(\vec{P})$, (b) $t_{A,i}/\rhd_{A,i}/G_A/G^{-1}_A(1)/G_A(\vec{x}) = 0/G_A(\vec{y}) = 1$ replaced by $s_{w,i}/\rhd_{w,i}/G_w/G^{-1}_w(0)/G_w(\vec{x}) = 1/G_A(\vec{y}) = 0$ and (c) IND/SMON/CONJ replaced by IND^w/SMON^w/CONJ^w.

(iii) Similarly, (a) let $F(\vec{P})(A) = G_{\vec{P}}(\vec{P}(A))$ and replace (b) $t_{A,i}/\rhd_{A,i}/G_A/G^{-1}_A(1)$ by $t_{\vec{P},i}/\rhd_{\vec{P},i}/G_{\vec{P}}/G^{-1}_{\vec{P}}(1)$ and (c) IND/SMON/CONJ by eIND/eSMON/eCONJ. Note that in this case $N_1 = N$(thereby $N_2 = \emptyset$), because given $\vec{P}$, $G_{\vec{P}}^{-1}(1)(\subseteq \{\vec{P}(A)|A \in \mathcal{P}(W)\})$ is finite since $\mathcal{P}(W)$ is finite.

(iv) Likewise, (a) let $F(\vec{P})(\overline{w}) = G_{\vec{P}}(\vec{P}(w))$ and replace (b) $t_{A,i}/\rhd_{A,i}/G_A/G_A^{-1}(1)$ $/G_A(\vec{x}) = 0/G_A(\vec{y}) = 1$ by $s_{\vec{P},i}/\rhd_{\vec{P},i}/G_{\vec{P}}/G_{\vec{P}}^{-1}(0)/G_{\vec{P}}(\vec{x}) = 1/G_{\vec{P}}(\vec{y}) = 0$ and (c) IND/SMON/CONJ by wIND/wSMON/wCONJ. Note that in this case $N_1 = N$ as in (iii) since W is finite.

□

Theorem 3.2 (Characterization of Pooling(f) + Threshold-based Binarization)**.** *Let f be an OP with the universal domain $\mathbb{P}(W)$.*

(i) The Pooling(f) + Threshold Binarization rules with global event-ths. are fully characterized by UD, f-IND and f-SMON;

(ii) so are the ones with global world-ths. by UD, CDC, f-IND^w and f-SMON^w;

(iii) so are the ones with local event-ths. by UD, f-eNEU and f-eSMON;

(iv) so are the ones with local world-ths. by UD, CDC, f-wNEU and f-wSMON

Proof. (i) It is clear that the rule satisfies the properties. For the other direction, by UD and f-IND we can let $F(\vec{P})(A) = G_A(f(\vec{P})(A))$ for all $\vec{P} \in \mathbb{P}(W)$. In the case of ${G_A}^{-1}(1) = \emptyset$, let $t_A := 1$ and $\rhd_A :=>$. Otherwise, let $t_A := \inf {G_A}^{-1}(1)$. If ${G_A}^{-1}(1)$ has the infimum, then let $\rhd_A :=\geq$, and otherwise let $\rhd_A :=>$. Our claim follows by f-SMON and f-IND.

(ii) This follows in the same manner with (a) $F(\vec{P})(\overline{w}) := G_w(f(\vec{P})(w))$ where $F(\vec{P})(\overline{w}) = 0$ iff $w \in B$ by CDC, where B is the belief core of $F(\vec{P})$. (b) $t_A/\rhd_A/G_A/G_A^{-1}(1)$ replaced by $s_w/\rhd_w/G_w/G_w^{-1}(0)$ and (c) f-IND/f-SMON replaced by f-IND^w/f-SMON^w.

(iii) Similarly, (a) let $F(\vec{P})(A) = G_{f(\vec{P})}(f(\vec{P})(A))$ and replace (b) $t_A/\rhd_A/G_A/G_A^{-1}(1)$ by $t_{\vec{P}}/\rhd_{\vec{P}}/G_{f(\vec{P})}/G_{f(\vec{P})}^{-1}(1)$ and (c) f-IND/f-SMON by f-eIND/f-eSMON. Note that when ${G_{f(\vec{P})}}^{-1}(1) \neq \emptyset$, ${G_{f(\vec{P})}}^{-1}(1)$ always has the infimum and thereby we can set $\rhd_{f(\vec{P})} :=\geq$, because given $\vec{P}$, $G_{f(\vec{P})}^{-1}(1)(\subseteq \{f(\vec{P})(A)|A \in \mathcal{P}(W)\})$ is finite since $\mathcal{P}(W)$ is finite.

(iv) Likewise, (a) let $F(\vec{P})(\overline{w}) = G_{f(\vec{P})}(f(\vec{P})(w))$ and replace (b) $t_A/\rhd_A/G_A/G_A^{-1}(1)$ by $s_{\vec{P}}/\rhd_{\vec{P}}/G_{f(\vec{P})}/G_{f(\vec{P})}^{-1}(0)$ and (c) f-IND/f-SMON by f-wIND/f-wSMON. when ${G_{f(\vec{P})}}^{-1}(0) \neq \emptyset$, we can set $\rhd_{f(\vec{P})} :=\geq$, because W is finite.

□

These two characterization theorems provide a useful framework to analyze and compare the eight classes of threshold-based rules. We first turn to part (i) of Theorem 3.1. It shows that the *direct rules with global event-thresholds* satisfy *IND*, which indicates that they are vulnerable to *the oligarchy result* described in the last chapter: if taken together with CDC (collective deductive closure), IND leads to the oligarchy result under the condition of CP (certainty preservation) and ZP (zero preservation) when the agenda is sufficiently complex. To circumvent this problem, we need to consider other approaches, e.g., procedures reducing the complexity of the agenda such as premise-based rules, methods with some inconsistency management like minimal change, or other kinds of threshold rules.

Global event-ths. might cause the same problem in *pooling(f)+threshold-based binarization* as well. As seen in part (i) of Theorem 3.2, it satisfies *f-IND*. If the OP f satisfies independence as well — e.g., linear pooling — in the sense that

$$\text{for every } A \in \mathcal{A}, \text{ if } \vec{P}(A) = \vec{P'}(A), \text{ then } f(\vec{P})(A) = f(\vec{P'})(A) \text{ for all } \vec{P}, \vec{P'},$$

the whole procedure satisfies *IND*, which leads to the oligarchy result as above. It is also worth pointing out that binarization with global event-ths. does not ensure collective deductive closure (CDC) as the lottery paradox shows.

Next let us move to *direct threshold rules*. In this research we restrict our focus to the simplest forms of direct threshold rules in which every individual's probability should *unanimously* exceed his/her own threshold. This requirement corresponds to *Conj*. This could be justified in some situations in which every individual's opinion should be respected, but in many other situations it would appear not so reasonable.

Now we turn to *pooling + threshold based binarization with world-ths. or local event-ths.* in part (ii)-(iv) of Theorem 3.2. As highlighted before, binarization with local thresholds involves the rules satisfying *Humean thesis (HT^r)* in Leitgeb (2014, 2017a) and the *Camera Shutter (CS^s)* rules in Lin & Kelly (2012b). In contrast to global event-ths., not only rules with world-ths. but also certain rules with local event-ths. elude the above problem of rationality. Firstly, the rules with world-ths. like the CS^s rules guarantee CDC, as shown in part (ii) and (iv). Secondly, not every local event-ths.-based binarization ensures rationality. However, notice that HT^r can generate a special kind of local event-ths.-based rules. If we combine our results with the stability theory of belief, we can see that the following holds:

$$\begin{gathered}\text{a HA F satisfies f-eMON (f-eNEU plus f-eSMON) and CDC}\\ \text{iff}\\ (f(\vec{P}), F(\vec{P})) \text{ satisfies } \mathrm{HT}^{\frac{1}{2}} \text{ for all } \vec{P} \text{ in the domain}\end{gathered}$$

where $\mathrm{HT}^{\frac{1}{2}}$ is the Humean thesis with $r = \frac{1}{2}$, which we will explain in full detail in the next section. It also deserves special mention that according to the stability theory of belief, the rules generated by HT^r can be seen as local world-threshold rules as well, which implies that the rules satisfy f-wMON(f-wNEU plus f-wSMON). Accordingly, the following also holds:

$$\begin{gathered}\text{a HA } F \text{ satisfies f-eMON, f-wMON and CDC}\\ \text{iff}\\ (f(\vec{P}), F(\vec{P})) \text{ satisfies } \mathrm{HT}^{\frac{1}{2}} \text{ for all } \vec{P} \text{ in the domain}\end{gathered}$$

Threshold-based binarization with local event/world ths. — that corresponds to Lockean/coherent binarization in the next section, if we assume belief sets/belief cores to be non-empty(NSK/CCS)— will be addressed in full detail in the next section.

3.3 Lockean, Coherent, Stable or Rational-likely Belief Binarization

$$\begin{tikzcd}
\begin{array}{c}\text{Individuals'}\\ \text{probabilistic}\\ \text{Beliefs } \vec{P}\end{array} \arrow[r, "\text{Opinion}", "\text{Pooling } f"'] \arrow[dr, "HA"'] & \begin{array}{c}\text{Group's}\\ \text{probabilistic}\\ \text{Belief } f(\vec{P})(=: P)\end{array} \arrow[d, "\text{Binarization } G"] \\
& \begin{array}{c}\text{Group's}\\ \text{binary}\\ \text{Belief } F(\vec{P})\end{array}
\end{tikzcd}$$

We have studied both direct threshold rules and pooling + threshold-based binarization so far. In the remainder of this chapter, we will focus only on belief binarization, which can be employed for heterogeneous belief aggregation if combined with opinion pooling, and we will look more closely at threshold-based binarization. Recall that $G : P \mapsto G(P)$ is a *binarization rule (BR)* on a finite space $(W, \mathcal{P}(W))$, where P is a probability function on $(W, \mathcal{P}(W))$ and $G(P) : \mathcal{P}(W) \to \{0, 1\}$ is a binary belief.

$$\begin{tikzcd}
\begin{array}{c}\text{Group's}\\ \text{probabilistic Belief}\\ P\end{array} \arrow[d, "\text{Binarization } G"] \\
\begin{array}{c}\text{Group's}\\ \text{binary Belief}\\ G(P)\end{array}
\end{tikzcd}$$

In this section, we will collect some *properties* of belief binarization that concern *binarization with local event-ths. and local world-ths.* The first aim of this section is to investigate the interrelation between the properties and suggest *geometrical characterizations* of some of them. Then, we will introduce certain threshold-based binarization *rules* in the literature on belief binarization and moreover, we will propose some other binarization rules. The second aim is to examine whether each of the binarization rules satisfy each property.

Properties of Belief Binarization The first two properties are *Lockean binarization* and *coherent binarization* that pertain to binarization with local event thresholds and local world thresholds, respectively. Let $Bel : \mathcal{P}(W) \to \{0, 1\}$ be a binary belief and P be a probability function on a finite space $(W, \mathcal{P}(W))$. We call an ordered pair

(P, Bel) a *belief state* on $(W, \mathcal{P}(W))$. We first define what a Lockean/coherent belief state means and using this, we will define Lockean/coherent binarization as follows.

Definition 3.11 (Lockean Binarization and Coherent Binarization)**.** *Let W be a finite non-empty set. Let (P, Bel) be a belief state on $(W, \mathcal{P}(W))$ and G be a BR.*

(1) (P, Bel) is called Lockean if

(i) Bel has a non-empty belief set ($Bel^{-1}(1) \neq \emptyset$), and

(ii) for all $A, C \in \mathcal{P}(W)$, if $P(A) \leq P(C)$ and if $Bel(A) = 1$, then $Bel(C) = 1$.

(Equivalently, (P, Bel) is called Lockean if Bel has a non-empty belief set and there exists a threshold $t_P \in [0, 1]$ such that $Bel(A) = 1$ iff $P(A) \geq t_P$ for all $A \in \mathcal{P}(W)$.)

(2) (P, Bel) is called coherent if

(i) Bel has a non-empty belief core B, and

(ii) for all $v, w \in W$ if $P(v) \leq P(w)$ and if $v \in B$, then $w \in B$.

(Equivalently, (P, Bel) is called coherent if Bel has a non-empty belief core B and there exists a threshold $s_P \in [0, 1]$ such that $w \in B$ iff $P(w) \geq s_P$ for all $w \in W$.)

(3) G is Lockean/coherent iff $(P, G(P))$ is Lockean/coherent for all P in the domain of G.

In a Lockean belief state, high probabilities of events matter, while in a coherent belief state, high probabilities of worlds matter. To be more precise, a Lockean/coherent belief state means that the belief set/belief core is (i) non-empty and (ii) contains all and only the events/worlds with probability above some thresholds. The thresholds depend on the probability and so we call them local thresholds. So, for a belief state to be Lockean/coherent, two conditions are required: (i) the belief set/the belief core being non-empty and (ii) a kind of monotonicity of probabilities of events/worlds. These two conditions can be rephrased in terms of local thresholds. Since it is guaranteed that the belief set/belief core is non-empty and $(W, \mathcal{P}(W))$ is a finite space, we can set $\rhd = \geq$ (and let $t_P = minG_P^{-1}(1)/s_P = minG_P^{-1}(0)$ where $G(P)(A) = G_P(P(A))/G(P)(\overline{w}) = G_P(P(w))$).

A BR G being Lockean/coherent are closely related to pooling + threshold-binarization with local event/world ths. in the last section. Let f be an OP and G be a BR where the image of $f(:= \{f(\vec{P})|\vec{P} \in \mathbb{P}(W)^n\})$ is included in the domain of G (e.g., G has the universal domain $\mathcal{P}(W)$), which we need to define $f + G$. Then the following two statements hold:

G is *Lockean* iff (i) F satisfies *NSK* and (ii) $F = f + G$ is a rule with *local event-ths.*.

G is coherent iff (i) F satisfies *CCS* and (ii) $F = f + G$ is a rule with *local world-ths.*.

Even though we will focus on Lockean and coherent binarization in this section, the above two statements say that this amounts to the study of pooling + local event/world threshold based binarization if we take NSK/CCS for granted.

One natural question would be when a binarization rule is Lockean and coherent simultaneously. This question leads us to introduce the *stability theory of belief* in Leitgeb (2013), Leitgeb (2014a) and Leitgeb (2017a). The theory will provide us with the notion of being stable that will turn out to be equivalent to being Lockean and coherent simultaneously. To define this notion, we will need the following definitions of HT^r and P-stable^r. *HT*^{r} is a joint constraint on the binary belief and the probability function of a belief state, concerning how they should rationally relate each other. It says that an agent should believe events of *stably high probability*. Here, an event having stably high probability means that its probability is above a given threshold $r(\in(\frac{1}{2},1])$, and it remains so high even with *conditionalization* on every event (with positive probability) not excluded in light of the agent's binary belief — i.e., every event whose complement is not believed. It turns out that binary beliefs that together with a probability function P satisfy HT^r are deductively closed, which means that each of them has its *belief core*.[3] What is more, a belief core must be *P-stable^r* to satisfy HT^r, in the sense that its conditional probability on every event (with positive probability) that is consistent with it should exceed the threshold r. Indeed, any P-stable^r sets with no zero-world (a world with probability zero) can be a belief core that induces a binary belief, together with P, satisfying HT^r. The following definition and lemma will make this more explicit and explain the reason for this.

Definition 3.12 (HT^r and P-stable^r (Leitgeb (2017a))). *Let W be a finite non-empty set and $r\in[\frac{1}{2},1)$. Let (Bel,P) denote a belief state, where P is a probability function on $(W,\mathcal{P}(W))$ and $Bel:\mathcal{P}(W)\to\{0,1\}$ is a binary belief. Let B be a subset of W.*

(1) (P,Bel) satisfies HT^r if the following holds:

$$\textit{for all } A\in\mathcal{P}(W),\ Bel(A)=1 \textit{ iff for all } Y\in\mathcal{P}(W) \textit{ with } Bel(\overline{Y})=0 \textit{ and } P(Y)>0\textit{: } P(A|Y)>r.$$

(2) B is P-stable^r if the following holds:

$$\textit{for all } Y\in\mathcal{P}(W) \textit{ with } Y\cap B\neq\emptyset \textit{ and } P(Y)>0\textit{: } P(B|Y)>r.$$

So, HT^r in (1) expresses that any event is believed iff it has a stably high probability and (2) says that a P-stable^r set B has a stably high probability when we

[3]See Definition 3.12 (1) below and the proof of Theorem 5 in Chapter 2 of Leitgeb (2017a). For the sake of self-containment, we repeat it here. (i) According to HT^r, $Bel(W)=1$. (ii) Let $A\subseteq B$. If $P(A|Y)>r$ then $P(B|Y)\geq P(A|Y)>r$. Thus, Bel is closed under superset. (iii) For a contradiction with closure under conjunction, suppose $Bel(A)=Bel(B)=1$ but $Bel(A\cap B)\neq 1$. Then $\overline{A\cap B}$ can be Y such that $Bel(\overline{Y})=0$ and $P(Y)>0$ (for otherwise $Bel(A\cap B)=1$). However, $P(A\cup B|\overline{A\cap B})=P(A|\overline{A\cap B})+P(B|\overline{A\cap B})>2r\geq 1$, a contradiction.

regard B as a belief core. It is because if B is a belief core, then the binary belief Bel induced by it satisfies that $Bel(\overline{Y}) = 0$ iff $Y \cap B \neq \emptyset$.

Notice that the empty set $\emptyset$ is always P-stable^r and so is every set of probability 1, which we call a *trivial P-stable^r* set. This distinction between trivial and *non-trivial P-stable^r* sets will be needed to explicate the relation between HT^r and being P-stable^r in the following lemma. And thus, it is worthwhile to discuss the following two points here. First, it is clear that except for the smallest set of probability 1, the other sets of probability 1 — the sets with a zero-world — cannot be the belief core of a binary belief satisfying HT^r together with any P. The reason is that the smallest set of probability 1 is always believed by any Bel satisfying HT^r, and consequently, the greater sets cannot be the belief core, as a belief core is the smallest believed set. Second, it is easily seen that *non-trivial P-stable^r sets and the smallest trivial P-stable^r set* are those sets B satisfying the following[4]:

$$\text{for all } w \in B,\ P(w) > \tfrac{r}{1-r} P(\overline{B})$$

This formula shows that there is such a big gap between the probabilities of the worlds in B and the probabilities of the worlds in $\overline{B}$ that each of the former probabilities is bigger than even the sum of all the latter probabilities when $r = \frac{1}{2}$. The greater r is, the bigger the gap becomes.

Now we formulate the relation between a belief state satisfying HT^r and a set being P-stable^r. We need the following lemma to define *binarization*, and moreover, we will use it to define a binarization rule that we will call HT^r(S) later and to prove some theorems in the next section as well.

Lemma 3.3 (HT^r and P-stable^r (Leitgeb (2017a)))**.** *Let W be a finite non-empty set and (P, Bel) be a belief state on $(W, \mathcal{P}(W))$. Let $r \in [\frac{1}{2}, 1)$. The following statements are equivalent:*

(i) *(P, Bel) satisfies HT^r and $Bel(\emptyset) = 0$*

(ii) *Bel has a non-empty belief core B such that B is a non-trivial P-stable^r set or the smallest set with probability* 1

(iii) *Bel has a non-empty belief core B such that for all $w \in B$, $P(w) > \frac{r}{1-r} P(\overline{B})$.*

Proof. Theorem 5 in Chapter 2 in Leitgeb (2017a), p. 108. □

Notice that $Bel(\emptyset) = 0$ amounts to Bel being *consistent*, if Bel is deductively closed. Even though HT^r ensures deductively closed binary beliefs, it does not guarantee consistency: it allows for a belief core being $\emptyset$, i.e., when $Bel^{-1}(1) = \mathcal{P}(W)$, it satisfies HT^r with any P. Since we take consistency for granted and we are interested only in consistent binary beliefs, we require consistency in our definition of being

[4]The proof is the following: Let v be a world with the minimum probability in B (if B is empty, then the claim is vacuously true.) Then we have $P(B|Y) \geq \frac{P(v)}{P(v)+P(\overline{B})}$ for all $Y \in \mathcal{P}(W)$ with $Y \cap B \neq \emptyset$ and $P(Y) > 0$. Since $\frac{P(v)}{P(v)+P(\overline{B})} > r$ iff $P(v) > \frac{r}{1-r} P(\overline{B})$, the claim holds.

stable, which corresponds to the condition of a belief core being non-empty, when a belief core is presupposed to exist.

This lemma provides a way to find a consistent binary belief satisfying HT^r given a probability function P: (i) pick one of any non-empty and non-trivial P-stable^r sets or the smallest set with probability 1 and (ii) believe all its supersets.

We not only focus on consistent beliefs but also want to allow as many binary beliefs satisfying the stability theory of belief together with a given probability function as possible, and thereby we want to attain the weakest, i.e., the least demanding, form of stability. Accordingly, we plug the smallest value $\frac{1}{2}$ into r in the above lemma and this leads to the following definition.

Definition 3.13 (Stable Binarization)**.** *Let W be a finite non-empty set and (P, Bel) be a belief state on $(W, \mathcal{P}(W))$. Let G be a BR.*

(1) (P, Bel) is stable iff one of the following equivalent statements holds:

- *(i) (P, Bel) satisfies $HT^{\frac{1}{2}}$ and $Bel(\emptyset) = 0$*
- *(ii) Bel has a nonempty belief core B such that B is a non-trivial P-stable set or the smallest set with probability* 1
- *(iii) Bel has a nonempty belief core B such that for all $w \in B$, $P(w) > P(\overline{B})$.*

(2) G is stable iff $(P, G(P))$ is stable for all P in the domain of G.

Simply stated, part (i) expresses that being stable is defined to be consistent and stable under conditionalization. This is equivalent to (ii), which turns out to be equivalent to (iii) the existence of non-empty belief core that is so stable that even a world with the minimum probability in it has a greater probability than its complement. We will use only (iii) in this chapter.

It is important to recognize the distinction between the property of a binarization being *stable*, the relation *HT^r* between the binary belief and the probability function of a belief state and the binarization rule *HT^r(S)*, to be defined later, which picks the smallest non-empty P-stable^r set as the belief core. The first one will serve as a norm that binarization rules should satisfy. The second one is not only used to define the first and third ones, but also gives a binarization relation and binarization correspondence in the next section, while the third one gives a binarization rule, which is a reduction method to specify which to believe given any probability function. They are different notions that have a different role, but closely related by definition: each input and its resulting output of the rule HT^r(S) satisfy HT^r and consistency and thus, they are stable belief states, which implies that the rule is stable.

Next let us turn to another notion, *rational likely binarization*, which is less demanding than being stable.

Definition 3.14 (r-likely Binarization)**.** *Let W be a finite non-empty set and (P, Bel) be a belief state on $(W, \mathcal{P}(W))$. Let G be a BR.*

(1) (P, Bel) is called rational-likely(r-likely) if

(i) *Bel has a nonempty belief core B, and*

(ii) *for all $A \in \mathcal{P}(W)$, if $Bel(A) = 1$, then $P(A) > P(\overline{A})$, i.e., $P(A) > \frac{1}{2}$.*

(Equivalently, (P, Bel) is called r-likely if (i) Bel has a nonempty belief core B, and (ii) $P(B) > \frac{1}{2}$.)

(2) G is r-likely iff $(P, G(P))$ is r-likely for all P in the domain of G.

This definition of a rational-likely belief state consists of two parts. Part (i) expresses that the binary belief of a belief state should be rational, i.e., consistent, and deductively-closed. The reason why we add this part is that we are basically interested in binarization that generates rational binary beliefs. Moreover, the requirement makes it simpler to compare the property of being r-likely with being coherent and being stable, since they imply being rational. And every rule we will consider in this section generates rational binary beliefs. Thus, the restriction of being rational will not affect our discussion in this section.

Part (ii) is the main part of this definition. It is required to compare each event and its complement, and then not to believe one of them that is less likely. It is immediate that under the assumption of the existence of a non-empty belief core, this is equivalent to the requirement that the probability of the belief core should be greater than one-half. This second part is weaker than being Lockean with a threshold greater than $\frac{1}{2}$, because it is just one direction of the requirement of being Lockean that $Bel(A)$ iff $P(A) > \frac{1}{2}$ — having a probability above a half is just a necessary condition for belief in the definition of being r-likely. Thus, it might elude the problem of rationality of Lockean binary belief. Even though we cannot respect the whole structure of the probabilities of events due to this problem, we could require to respect at least the probability pair of each event and its complement.

The last property we will introduce is rationality of binarization, which encompasses being coherent, stable, and r-likely. One might notice that in all of the three definitions of a belief state being coherent, stable, and r-likely, the existence of a non-empty belief core is required, which is equivalent to rationality of the binary belief of the belief state. A rational binarization is defined to be a binarization that generates rational binary beliefs, as follows.

Definition 3.15 (Rational Binarization)**.** *Let G be a BR. G is rational iff $G(P)$ is rational for all P in the domain of G.*

The requirement of non-empty belief core will play a crucial role to define various rational binarization rules.

So far, we have introduced all of properties of a belief state and binarization we want to explore, and we are now ready to look at the relation between them. Figure 3.1 depicts the relation between the properties. Lockean and rational binarization is exactly the same as stable binarization, which is coherent and r-likely, hence rational.

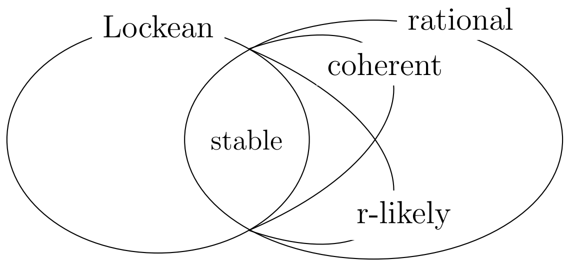


Figure 3.1: Properties of Binarization Rules

This relation can be formulated as follows. The theorem and proof can be found in Letigeb (2013), Leitgeb (2014a) or Leitgeb (2017a), but to make our exposition self-contained, we offer the proof here.

Theorem 3.4 (Stability and other properties (Letigeb (2013), Leitgeb (2014a), Leitgeb (2017a)))**.** *Let G be a BR. The following statements are equivalent.*

(i) G is stable.

(ii) G is Lockean and coherent.

(iii) G is Lockean and r-likely.

(iv) G is Lockean and rational.

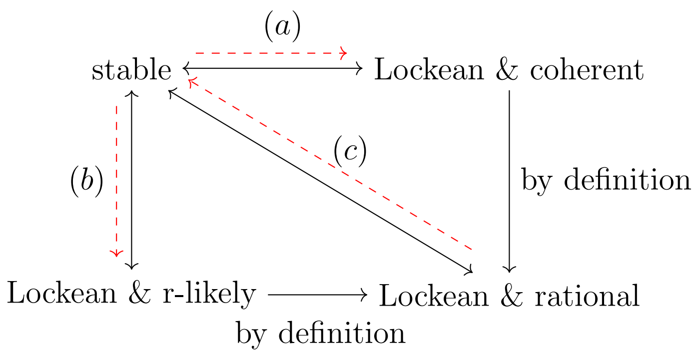


Figure 3.2: Equivalence among properties

Proof. It is sufficient to show the following(See the figure):
(a) if (P, Bel) is stable, then (P, Bel) is Lockean and coherent
(b) if (P, Bel) is stable, then (P, Bel) is Lockean and r-likely
(c) if (P, Bel) is Lockean and rational, then (P, Bel) is stable.

To (a) and (b): assume that (P, Bel) is stable. Thus Bel has a non-empty belief core B such that

$$P(v) > P(\overline{B}) \tag{3.1}$$

where v is a world with the minimum probability in B, and thereby if $P(v) > P(w)$, then $w \notin B$. If $P(u) \geq P(v)$, then $u \notin \overline{B}$, for otherwise $P(v) \leq P(\overline{B})$. Thus (P, Bel) is coherent.

Now suppose that $P(B) < \frac{1}{2}$ and thus $P(v) < \frac{1}{2}$ and $P(\overline{B}) \geq \frac{1}{2}$. This contradicts the inequality in (3.1). Thus, (P, Bel) is r-likely.

Moreover, if $Bel(A) = 1$, i.e., $A \supseteq B$, then $P(A) \geq P(B)$. Now consider the set that is not believed and has the maximal probability. This is $\overline{\{v\}}$, and $P(\overline{\{v\}}) = P(B) - P(\{v\}) + P(\overline{B}) < P(B)$ because of the inequality in (3.1). Therefore, if $Bel(A) = 0$, then $P(A) < P(B)$. Hence (P, Bel) is Lockean.

To (c): assume that (P, Bel) is rational, and thereby there exists a nonempty belief core B. Suppose that (P, Bel) is not stable. Then, $P(v) \leq P(\overline{B})$ where v is a world with the minimum probability in B. From this, it follows that $P(\overline{\{v\}}) = P(B) - P(\{v\}) + P(\overline{B}) \geq P(B)$ even though $Bel(\overline{\{v\}}) = 0$ and $Bel(B) = 1$, which shows that (P, Bel) is not Lockean.

□

Geometrical Characterization Now, we will examine how the properties presented so far can be geometrically characterized. These characterizations not only give us a geometric intuition for the properties, but also provide convenient criteria for testing which rules satisfy each property.

For geometrical characterizations, we first introduce a typical way to represent a probability function in Euclidean space. Let $W := \{w_1, ..., w_m\}$ be a finite nonempty set of possible worlds with $m \geq 1$. For any probability function $P \in \mathbb{P}(W)$ on $(W, \mathcal{P}(W))$, we will use the small letter p to denote its *representation point* $(p_1, ..., p_m)$ in $\mathbb{R}^m$ where $p_i = P(w_i)$ for all $i \leq m$. The *probability simplex* (with dimension $m-1$) $\triangle^m$ is defined as the set of the representation points of all probability functions, i.e.,

$$\triangle^m := \{p \in \mathbb{R}^m | P \in \mathbb{P}(W)\}$$

Hence it is evident that there is one-to-one correspondence between $\mathbb{P}(W)$ and $\triangle^m$. In this way we can represent the input of a BR in $\triangle^m (\subseteq \mathbb{R}^m)$.

Our next question would be how to represent outputs of a BR, binary beliefs. We propose a convenient way for a *rational BR*. Let G be a rational BR, then it enables us to regard a binary belief $G(P)$ as a *non-empty belief core*, a non-empty subset of W. Our main idea is to identify any non-empty belief core B with the *representation point of the uniform distribution* $U(B)$ on B. We denote the point corresponding to a belief core B by $b \in \triangle^m$. Let $U^m (\subseteq \triangle^m)$ be the set of such points, i.e.,

$$U^m := \{b \in \triangle^m | b \text{ is the representation point of } U(B) \text{ for some } B \in \mathcal{P}(W) \setminus \{\emptyset\}\}$$

Then there is one-to-one correspondence between U^m and $\mathcal{P}(W) \setminus \{\emptyset\}$. Accordingly, we can treat G as a function from $\triangle^m$ to U^m. For $b \in U^m$, we define

$$G^{-1}(b) := \{p \in \triangle^m | \; G(P) = B\}$$

(the set of the representation points of the probability functions in the preimage of B under G). We call $G^{-1}(b)$ the preimage-region of b under G.

With this in place, let us start to characterize *coherent binarization*. We will do it using the *Voronoi diagram*, which will also be useful throughout the rest of this chapter (Later, we will introduce several distance-based binarization methods that can be expressed by Voronoi diagrams). Here is its definition:

Definition 3.16 (Voronoi Diagram). *Let $R = \{r_1, r_2, ..., r_l\} \subseteq \mathbb{R}^m$ ($l \geq 1$) and $r_i \in R$, which is called a generator. We call the set given by*

$$V(r_i|R) = \{x \in \mathbb{R}^m | \ ||x - r_i|| \leq ||x - r_j|| \ \forall j\}$$

the Voronoi cell of r_i with respect to R where $||\cdot||$ is the Euclidean distance. We also use $V(r_i|R)^\circ$ to denote the set $\{x \in \mathbb{R}^m | \ ||x - r_i|| < ||x - r_j|| \ \forall j \neq i\}$. We call the set given by

$$\mathcal{V}(R) = \{V(r_1|R), V(r_2|R), ..., V(r_l|R)\}$$

the Voronoi diagram of R.

So, given generators in $\mathbb{R}^m$, the Voronoi diagram decomposes $\mathbb{R}^m$ into the regions close to each generator. More precisely, the Voronoi cell of a generator consists of all points that have the minimum distance from that generator rather than from all other generators, and the Voronoi diagram is the set of the Voronoi cells. Geometrically speaking, the Voronoi cell of a generator is the intersection of the half-spaces containing the generator that are bounded by the perpendicular bisectors of the line segment connecting the generator and one of the other generators. For example, assume that the points in Figure 3.3 are the generators in $\mathbb{R}^2$. The line segments connecting two generators are represented by blue dotted line segments and the boundaries of the Voronoi cells are drawn with black solid line segments.

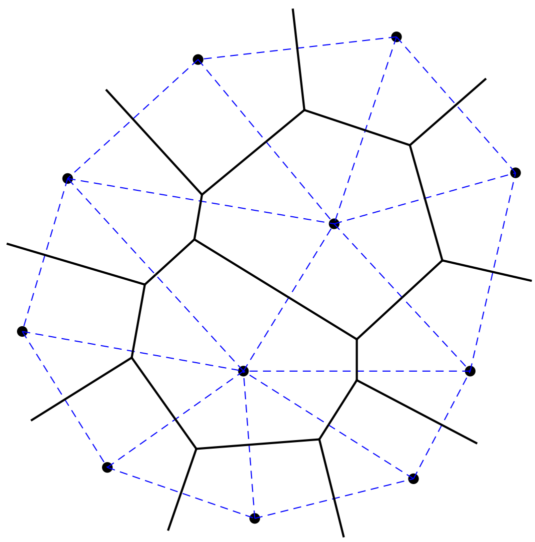


Figure 3.3: A Voronoi Diagram in $\mathbb{R}^2$

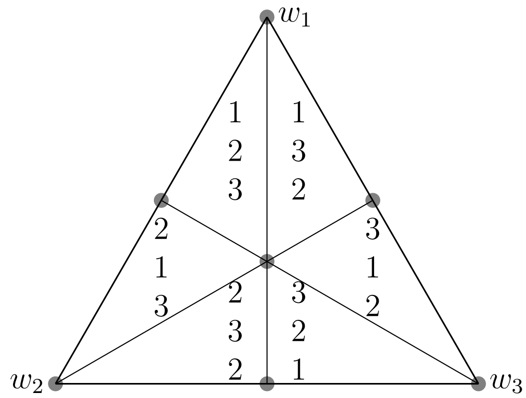


Figure 3.4: Probability Simplex $\triangle^3$ and Belief Cores

Using the notion of the Voronoi diagram, let us first characterize coherent binarization when $W = \{w_1, w_2, w_3\}$. Figure 3.4 represents the probability simplex $\triangle^3$ in $\mathbb{R}^3$. The seven grey points b correspond to the seven non-empty subsets $B \subseteq W$, and constitute U^3 — i.e., $U^3 = \{b_1, b_2, b_3, b_{12}, b_{13}, b_{23}, b_{123}\}$ where $b_1 = (1, 0, 0)$, $b_2 = (0, 1, 0)$, $b_3 = (0, 0, 1)$, $b_{12} = (\frac{1}{2}, \frac{1}{2}, 0)$, $b_{13} = (\frac{1}{2}, 0, \frac{1}{2})$, $b_{23} = (0, \frac{1}{2}, \frac{1}{2})$ and $b_{123} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The simplex is divided into the 6 smallest triangles, in which the indices of worlds are listed, from top to bottom, in decreasing order of the probability values of the worlds: for example, the numbers $\substack{1\\2\\3}$ in the upper left triangle indicates that $p_1 > p_2 > p_3$ for any p in that triangle that is neither on the vertical edge nor the lower left edge of the triangle. Figure 3.4 illustrates three Voronoi diagrams as well. Let us divide U^3 into three sets $U^3_1 = \{b_1, b_2, b_3\}$, $U^3_2 = \{b_{12}, b_{23}, b_{13}\}$ and $U^3_3 = \{b_{123}\}$. Then we can make the three Voronoi diagrams $\mathcal{V}(U^3_1)$, $\mathcal{V}(U^3_2)$ and $\mathcal{V}(U^3_3)$ (they will be depicted in Figure 3.5). The line segments inside of the simplex correspond to the Voronoi cell boundaries of the three Voronoi diagrams.

Now, consider the region where $p_1 > p_2, p_3$ — the upper left and right triangles. It lies in $V(b_1|U^3_1)^\circ$. And the region where $p_1, p_2 > p_3$ — the triangles with $\substack{1\\2\\3}$ and $\substack{2\\1\\3}$ — lies in $V(b_{12}|U^3_2)^\circ$. If G is coherent and $G(P) = \{w_1\}/G(P) = \{w_1, w_2\}$, then p lies in the region where $p_1 > p_2, p_3/p_1, p_2 > p_3$, respectively. Accordingly, for a BR G to be coherent, it must hold that

$$G^{-1}(b_1) \subseteq V(b_1|U^3_1)^\circ$$

and

$$G^{-1}(b_{12}) \subseteq V(b_{12}|U^3_2)^\circ$$

Figure 3.5 shows these subset relations. The leftmost/centered/rightmost figure corresponds to the Voronoi diagram $\mathcal{V}(U^3_1)/\mathcal{V}(U^3_2)/\mathcal{V}(U^3_3)$, respectively, restricted to $\triangle^3$. The dotted lines show the way each Voronoi diagram divides each simplex. Each region containing b and bounded by red line represents $G^{-1}(b)$ for some G. For a rational BR to be coherent, every preimage-region should be strictly included in the corresponding Voronoi cell and thus, the red lines should not touch the dotted lines.

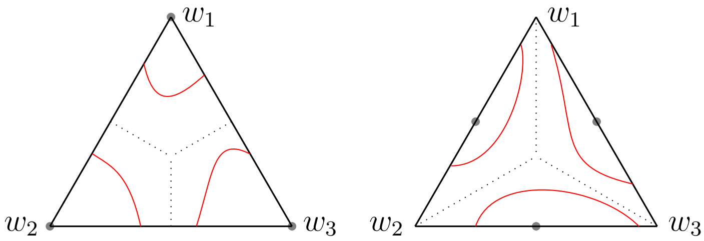


Figure 3.5: Coherence and the Voronoi Diagrams $\mathcal{V}(U_1^3)$(leftmost)/$\mathcal{V}(U_2^3)$(centered)/$\mathcal{V}(U_3^3)$(rightmost), restricted in $\triangle^3$.

Now, let us generalize this reasoning to the case where $|W| = m$. For a representation point $p \in \triangle^m$, we define the support $Supp(p)$ of p to be the same as the support of the corresponding probability function P, i.e.,

$$Supp(p) = Supp(P) = \{w \in W | P(w) \neq 0\}$$

Assume that a rational BR G is coherent. Let $G(P) = B$ (i.e., $p \in G^{-1}(b)$). Since G is coherent, B consists of the worlds with higher probability values of P than the worlds in $\overline{B}$. This implies that p lies in the region where for all $i \in Supp(b)$ (i.e., $w_i \in B$) and for all $j \notin Supp(b)$ (i.e., $w_j \in \overline{B}$)

$$p_i > p_j$$

which can be proven to be equivalent to

$$p \in V(b|U_k^m)^\circ$$

where $k = |Supp(b)|$ and $U_k^m = \{b' \in U^m ||Supp(b')| = k\}$. Therefore

$$G^{-1}(b) \subseteq V(b|U_k^m)^\circ \tag{3.2}$$

Conversely, if we assume (3.2) for all $b \in U^m$, we can check that G is coherent. The following theorem and its proof shows this intuition is true.

Theorem 3.5 (Geometric Characterization of Coherence)**.** *Let $|W| = m$. A rational binarization rule G is coherent iff for all $k \leq m$ and for all $b \in U_k^m$,*

$$G^{-1}(b) \subseteq V(b|U_k^m)^\circ$$

where $U_k^m = \{b' \in U^m|\ |Supp(b')| = k\}$

Proof. ($\rightarrow$) Suppose that a binarization rule G is coherent. Suppose that there are $k \leq m$, $b, b' \in U_k^m$, and $p \in \triangle^m$ such that $p \in G^{-1}(b)$, $b \neq b'$, and $p \in V(b'|U_k^m)$. Since B and B' are different sets with the same cardinality, there are w and w' such that $w \in B \backslash B'$ and $w' \in B' \backslash B$. Let us compare p_w and p'_w. Define $B'' = (B' \backslash \{w'\}) \cup \{w\}$. Note that $b'' \in U_k^m$. Since $p \in V(b'|U_k^m)$, we have

$$||p - b'||^2 \leq ||p - b''||^2$$

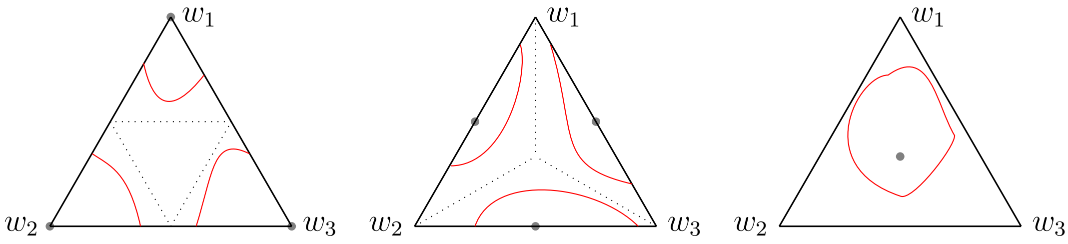


Figure 3.6: stable for $|W| = 3$ for $|W| = 3$

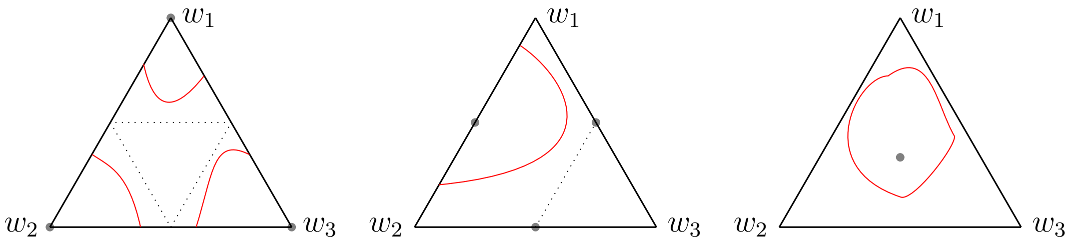


Figure 3.7: r-likely for $|W| = 3$

It follows that

$$(p_{w'} - \frac{1}{k})^2 + p_w^2 \leq (p_w - \frac{1}{k})^2 + p_{w'}^2$$

which implies that $p_{w'} \geq p_w$. Since G is coherent, $G(P) = B$, and $w \in B$, we have $w' \in B$, which contradicts our assumption that $w' \in B' \setminus B$.
($\leftarrow$) Assume the RHS. And suppose that there are $w, w' \in W$, $p \in \triangle^m$, and $b \in U_k^m$ such that $p \in G^{-1}(b)$, $w \in B$, $P(w') \geq P(w)$, and $w' \notin B$. Then there is a $b' \in U_k^m$ such that $B' = (B \setminus \{w\}) \cup \{w'\}$. Since $G^{-1}(b) \subseteq V(b|U_k^m)^\circ$, we have

$$||p - b||^2 < ||p - b'||^2$$

which implies that $p_{w'} < p_w$, which contradicts our assumption that $P(w') \geq P(w)$. $\square$

The above theorem says that a coherent BR can be fully characterized by the geometrical property saying that each preimage-region of the BR is strictly contained in the respective Voronoi cell.

Next we turn to the property of being *stable* and being *r-likely*. In this research, we provide geometrical characterization only when $W = \{w_1, w_2, w_3\}$ and leave a generalization for the future work. It will enable one to see whether various kinds of rational BRs are stable or r-likely when $|W| = 3$. Figure 3.6 and Figure 3.7 illustrates the conditions for a rational BR to be stable and to be r-likely. In both figures, the preimage-regions are represented by red curved lines. Each leftmost/centred/rightmost figure illustrates the preimage-regions of b_1, b_2 and b_3/b_{12}, b_{21} and b_{13}/b_{123}, respectively. For a rational BR to be stable$^{\frac{1}{2}}$/r-likely, each preimage-region should not cross the dotted lines. The reason is the following. Firstly, for a singleton represented by b_i to be a belief core by a BR that is stable$^{\frac{1}{2}}$, it must hold that $p_i > p_j + p_k$ where

$j \neq k \in \{1, 2, 3\}$, i.e.,

$$p_i > \frac{1}{2}$$

which is the same condition for being r-likely. The dotted lines in both leftmost figures represent the region where $p_i = \frac{1}{2}$ for $i \in \{1, 2, 3\}$. Secondly, as for b_{ij}, the centred figures give the criteria. To be stable, it must hold that $p_i, p_j > p_k$, which is equivalent to the condition to be coherent. Thus the centred one in Figure 3.6 is the same as the one in Figure 3.5. By contrast, the condition to be r-likely is that

$$p_i + p_j > \frac{1}{2}$$

which is shown in the centred one of Figure 3.7 (We described only the case for $b_{ij} = b_{12}$. The other cases are obvious from symmetry).

3.4 Threshold-based Binarization Rules

We now introduce several belief binarization rules, which we will examine regarding whether the properties in this section are satisfied or not. A lot of research has been done on the first three rules while the forth rule has been rarely studied and the fifth rule is proposed for the first time in this thesis.

Definition 3.17 (Threshold-based Binarization Rules). *Let G be a rational BR with the domain $\mathbb{P}(W)$, with $G(P)$ being regarded as a non-empty belief core (a non-empty subset of W) for all $P \in \mathbb{P}(W)$.*

(i) Let $r \in [\frac{1}{2}, 1)$. G is the Humean Thesis with the Smallest-stable-selection rule (HT^r(S)) if $G(P)$ is the smallest non-empty P-stable^r set $B \subseteq W$ for all $P \in \mathbb{P}(W)$.[5]

(ii) Let $s > 1$. G is the Camera Shutter rule (CS^s) if $G(P) = \{v \in W | P(v) \geq \frac{max_{w \in W} P(w)}{s}\}$ for all $P \in \mathbb{P}(W)$.[6]

(iii) Let $s > 1$, $\sum_{w \in W} m_w = 1$, and $m_w > 0$ for all $w \in W$. G is the generalized Camera Shutter rule (gCS^s) if $G(P) = \{v \in W | m_v P(v) \geq \frac{max_{w \in W} m_w P(w)}{s}\}$ for all $P \in \mathbb{P}(W)$.[7]

(iv) Let $g \in (0, 1]$. G is the Coherent Core-Threshold rule (CCT^g) if $G(P)$ is the smallest coherent set satisfying $P(G(P)) \geq g$ for all $P \in \mathbb{P}(W)$.[8]

(v) G is the Distance Minimization rule with Squared Euclidean distance (DM(SE)$^+$) if $G(P) \in \text{argmin}_{B \subseteq W} D(P, B) := \{B \subseteq W| \; ||b - p||^2 \leq ||b' - p||^2 \; \forall b' \in U^m\}$ for all $P \in \mathbb{P}(W)$ (combined with some tie-breaking rule).

It is easy to check that every rule is well-defined and rational, i.e., for every rule G in the above definition, for each $P \in \mathbb{P}(W)$ there exists a unique $G(P)$ that is non-empty. We now go through the rules in the above definition one by one.

Let us begin with part (i) of the definition. HT^r allows, given a probability function P, multiple candidates for the non-empty belief core: non-trivial non-empty P-stable sets and the smallest set with probability 1 (See Lemma 3.3 (ii)) From Lemma 3.3 (iii), it is easy to check that they are in a subset relationship each other, and thus form a system of spheres or a ranked system of sets. Then we can devise an easily definable and justifiable selection rule HT^r(S) according to which the innermost sphere is always chosen.[9] This selection function can be useful when the binarization context demands the largest belief set. Figure 3.8 depicts

[5] See Leitgeb (2013). For the discussion about HT^r(S) see Cariani (2016), Thorn (2018), and Wright (2018).

[6] See Lin & Kelly (2012a, 2012b).

[7] See Lin & Kelly (2012b).

[8] See Cantwell & Rott (2019).

[9] Of course there can be other selection rules. For example, we can devise a new rule by measuring distances from the point representing the given probability function to the points representing the candidates for the non-empty belief core that are permitted by HT^r. Formally, HT^r equipped with

which belief core $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$(left)/$\mathrm{HT}^{\frac{3}{4}}(\mathrm{S})$(right) assigns to each probability function on $W = \{w_1, w_2, w_3\}$. In the left figure, the preimage-regions of b_1,b_2 and b_3 are respectively the upmost, leftmost, and rightmost small triangles excluding their edge inside of the simplex. The preimage-regions of b_{12}, b_{13}, and b_{23} are respectively the left, right, and lower quadrangles in the middle excluding the bold lines, which constitute the preimage-region of b_{123}. In the right figure, the simplex is divided into seven regions. Each of them is the preimage-region of the point in it.

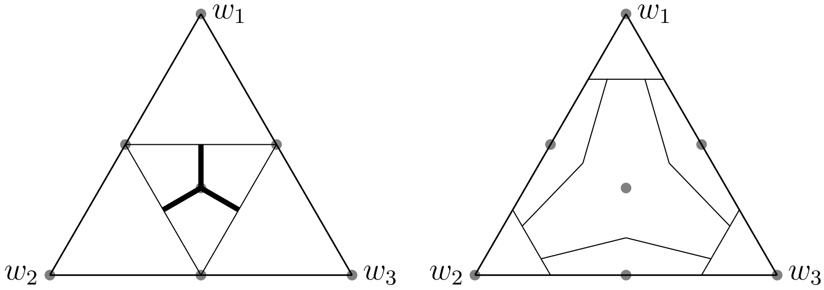


Figure 3.8: $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$(left) and $\mathrm{HT}^{\frac{3}{4}}(\mathrm{S})$(right)

We next turn to CS^s and gCS^s. According to CS^s, what matters is probabilities of worlds. Contrary to $\mathrm{HT}^r(\mathrm{S})$, probabilities of events play no role. Given $s > 1$ and a probability function P, the way to choose the worlds with higher probabilities to constitute a belief core $G(P)$ is using a world-ordering $\prec_P$ such that

$$w \prec_P v \text{ iff } sP(w) < P(v)$$

for all $v, w \in W$. *The maximal elements* with respect to the world-ordering $\prec_P$ constitute the belief core of $G(P)$. Accordingly, we have

$$v \in G(P) \text{ iff } P(v) \geq \frac{max_{w \in W} P(w)}{s}$$

That is, for a world to be in the belief core, *the ratio of its probability to the maximum probability* must be above the threshold $\frac{1}{s}$. So this rule utilizes only ratios between worlds' probabilities, regardless of any sums of worlds' probabilities — i.e., probabilities of events. Turning to gCS^s, it generalize CS^g by introducing a weighting vector $(m_w)_{w \in W}$ to worlds and the ordering such that $w \prec'_P v$ iff $sm_wP(w) < m_vP(v)$. In Figure 3.9, the left one depicts CS^s when $s = 3$, and the right one is for gCS^3, in which the weighting vector has the effect of breaking the rotational symmetry of order-3.[10] The symmetry is so broken that even b_{123} and b_{23} are not in their preimage-region.

Now let us move to CCT^g. Just as CS^s, a belief core of $G(P)$ consists of worlds with higher probabilities. That is, all coherent sets are candidates for a belief core, which are related each other in a subset relationship, and thus form a system of

distance minimization, denoted by $\mathrm{HT}^r(\mathrm{D})$ is defined as following: G is $\mathrm{HT}^r(\mathrm{D})$ if $G(P) = B$ where B is the non-empty P-stable^r set $B \subseteq W$ whose corresponding point $b(:= (u(B)(w))_w \in U^m)$ is the closest from $p(:= (P(w))_w \in \triangle^m)$ with respect to Euclidean distance for all $P \in \mathbb{P}(W)$.

[10]The rotational symmetry of order-n means that it looks the same after rotation by $\frac{2\pi}{n}$ radians.

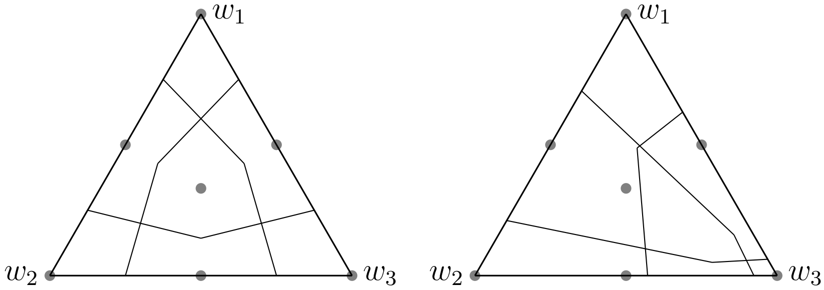


Figure 3.9: CS^3 and gCS^3 with the weighting vector $(0.5, 0.4, 0.1)$.

spheres like the case of HT^r. Confronting multiple options, CCT^g chooses the innermost sphere B whose probability $P(B)$ exceeds the threshold g, while HT^r chooses the innermost non-empty P-stable^r sphere and CS^s chooses the set of the maximal elements with respect to $\prec_P$. That is, the belief core should be the smallest coherent one with probability above g, called a *belief core threshold*. It would be one of the simplest way to determine a belief core among coherent sets. It is worth noting that while CS^s determines belief cores solely based on the ratio between probabilities of worlds, CCT^g obtains belief cores using a threshold applied to the probability of a belief core, i.e., the sum of probabilities of worlds constituting a belief core. The left simplex in Fig 3.10 illustrates $\mathrm{CCT}^{0.75}$. Note that the preimage-region of the centre consists of the gray triangle region and three gray line segments.

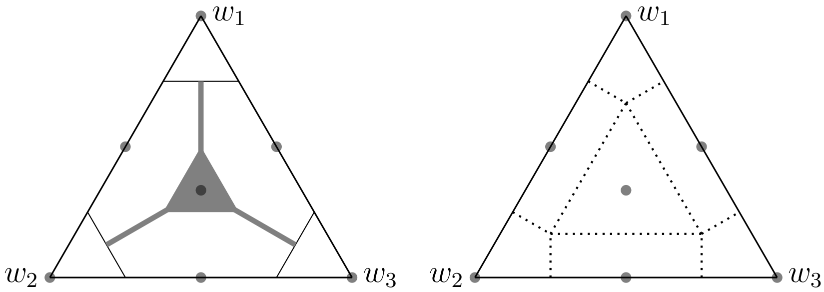


Figure 3.10: $\mathrm{CCT}^{0.75}$ and $\mathrm{DM(SE)}^+$

Finally, let us get into $\mathrm{DM(SE)}^+$, which is a new rule we are proposing in this research. As a belief binarization procedure, we might measure some kind of distance between inputs and outputs, and choose the closest one to a given input. The input is a probability function and the output is a binary belief. Thus, we need to find a way to measure distances between them. Usually, a probability function is represented as a point in a probability simplex, and we proposed, in this section, a new way to represent a rational binary belief as a point (the representation point of the uniform probability function on the belief core of a rational binary belief) in a probability simplex. Thus, we can use any distance defined in $\triangle^m$. We call this binarization procedure distance minimization procedure, which will be the main subject of the next chapter. In this chapter, we address the most basic distance minimization procedure, which is the one with the squared Euclidean distance.[11] Since this distance comparison might allow multiple options, we may have two different ways: (1) allowing multiple options and (2) adding a tie-breaking rule. If qualitative beliefs can be seen as intrinsically vague concepts related to credal states, (1) would not be so indecisive. In the next section, we will address this possibility. In this section we focus only on a binarization rule, by which we mean a function mapping each input into one binary belief. So by adding some tie-breaking rule, we can make DM(SE)$^+$. In Fig 3.10, the right one illustrates DM(SE)$^+$. The dotted lines are drawn to emphasize the place where any arbitrary tie-breaking rule is needed. As previously mentioned, DM(SE)$^+$ and the Voronoi diagram are closely related. Let $\mathcal{V}(U^m)$ be the Voronoi diagram of U^m on $\mathbb{R}^m$ and G be a DM(SE)$^+$. Then for any $b \in U^m$ we have the following subset relation:

$$(V(b|U^m)^\circ \cap \triangle^m) \subseteq G^{-1}(b) \subseteq (V(b|U^m) \cap \triangle^m)$$

Exactly where $G^{-1}(b)$ lies between $(V(b|U^m)^\circ \cap \triangle^m)$ and $(V(b|U^m) \cap \triangle^m)$ depends on which tie-breaking rule we adopt. If we ignore the boundaries of $G^{-1}(b)$ and $V(b|U^m) \cap \triangle^m$, they are the same.[12]

Properties of Rational Binarization Rules Now we will check whether the introduced BRs satisfy the properties introduced in the beginning of this section. In what follows, to say a rule satisfies a property means that the rule satisfies the property for all contexts embracing all numbers of possible worlds and all values of parameters. Thus, to say a rule does not satisfy a property means that there is a context — i.e., a probability space and a value of parameter — in which the rule does not satisfy the property.

Let us begin with simple ones. Some positive results can be easily proved. Since HT^r(S) is, by definition, stable, by Theorem 3.4, it holds that HT^r(S) is Lockean, coherent, and r-likely. By definition, CS^s and CCT^g are coherent, and CCT^g is r-likely given $g > \frac{1}{2}$. By geometrical characterization of being coherent (Theorem 3.5) and being r-likely, we see that gCS^s is neither coherent nor r-likely, and thus neither stable nor Lockean.

To answer the remaining questions, let us use the geometrical characterizations for $|W| = 3$ to get some intuition. We put off discussing whether DM(SE)$^+$ is coherent, and focus on the properties of being stable and being r-likely. Comparing the figures for CS^3, CCT$^{0.75}$ and DM(SE)$^+$, and the figures for being stable and being r-likely for $|W| = 3$ (Figure 3.6 and 3.7) leads one to conclude that CS^3, CCT$^{0.75}$, and DM(SE)$^+$ are stable and r-likely. Are these statements still be the case in other contexts where there are more than three possible worlds or other parameter values are used? We will

[11] DM(SE)$^+$ uses the squared Euclidean distance as the distance measure, which produces the same result as when we adopt the Euclidean distance. The reason why we utilize the squared Euclidean distance is that this metric belongs to the Bregmann divergence contrary to the Euclidean metric. The Bregman divergence will play a central role in the next chapter. In this chapter, the difference between the squared Euclidean distance and the Euclidean distance has no effect.

[12] Later, we will introduce the notion of relative interior. Then this informal observation can be formalized as the following: $ri(G^{-1}(b)) = ri(V(b|U^m) \cap \triangle^m)$ where $ri(A)$ means the relative interior of some subset $A \subseteq \mathbb{R}^m$.

	$\text{HT}^r(\text{S})$	CS^s	gCS^s	CCT^g	DM(SE)^+
Lockean	O	X	X	X	X
coherent	O	O	X	O	interior
$\text{stable}^{\frac{1}{2}}$	O	X	X	X	X
r-likely	O	X	X	$g > \frac{1}{2}$	O

Table 3.3: Rational BRs and Properties. In the table 'O'/'X' means that the rule satisfies/does not always satisfy the property. 'interior' means that the rule is interior coherent and '$g > \frac{1}{2}$' means that the rule satisfy the property when $g > \frac{1}{2}$.

answer all of these questions. The results are collected in Table 3.3 and illustrated in Figure 3.11. In the remainder of this section, we will prove every claim.

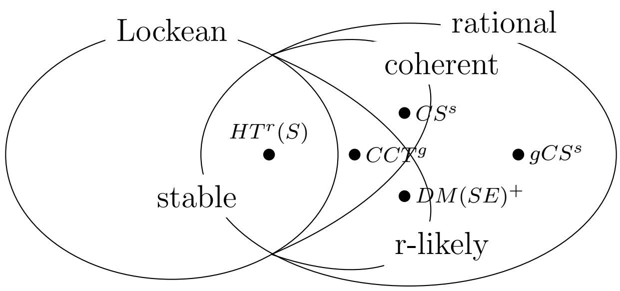


Figure 3.11: Properties of Rational Binarization Rules. In this figure, we assume $g > \frac{1}{2}$. Note that DM(SE)^+ is interior coherent.

In the following theorem, some negative results are proved. Every proof includes a counterexample.

Theorem 3.6. *(1) CS^s is neither $\text{stable}^{\frac{1}{2}}$ nor r-likely.*

(2) CCT^g is not $\text{stable}^{\frac{1}{2}}$.

(3) $DM(SE)^+$ is not $\text{stable}^{\frac{1}{2}}$.

Proof. (1) Let G be $\text{CS}^{1.2}$. Consider a probability distribution P on $W = \{w_1, w_2, w_3\}$.

$$\begin{array}{c||cc} w_1 & w_2 & w_3 \\ 0.4 & 0.3 & 0.3 \end{array}$$

Since $G(P) = \{w_1\}$, and $P(w_1) = 0.4$, it follows that G is neither $\text{stable}^{\frac{1}{2}}$ nor r-likely.

(2) Let G be $\text{CT}^{0.55}$. Consider the following probability distribution P on $W = \{w_1, w_2, w_3, w_4, w_5\}$:

$$\begin{array}{cc||ccc} w_1 & w_2 & w_3 & w_4 & w_5 \\ 0.3 & 0.25 & 0.15 & 0.15 & 0.15 \end{array}$$

Since $G(P) = \{w_1, w_2\}$, and $P(w_2) < P(W \setminus \{w_1, w_2\})$, it follows that G is not stable$^{\frac{1}{2}}$.

(3) Let G be a DM(SE)$^+$. Consider a probability distribution P on $W = \{w_1, w_2, w_3, w_4, w_5\}$.

	w_1	w_2	w_3	w_4	w_5
P	0.5	0.16	0.16	0.09	0.09
b	0.2	0.2	0.2	0.2	0.2
b'	0.25	0.25	0.25	0.25	0
b''	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0

By computation, we have $||p - b||^2 = 0.1174, ||p - b'||^2 = 0.1124, ||p - b''||^2 = 0.10406$, which implies $G(P) = \{w_1, w_2, w_3\}$. Since $P(w_3) < P(W \setminus G(P))$, G is not stable$^{\frac{1}{2}}$. □

Since CS^s, CCT^g, and DM(SE)$^+$ are not stable, and all of them are rational, by Lemma 3.4, it follows that they are not Lockean.

There remain two questions. The first one is whether DM(SE)$^+$ is r-likely; the second one is whether DM(SE)$^+$ is coherent. Let us deal with the first one. The next theorem shows a positive result.

Theorem 3.7. *DM(SE)$^+$ is r-likely.*

Proof. Suppose that $|W| = m, G(P) = B = \{w_1, w_2, ..., w_k\}$. If $k = m$, then $P(B) = 1$. Suppose $k < m$. Geometrically, we set $p = (p_1, p_2, ..., p_m), u = (\frac{1}{m}, \frac{1}{m}, ..., \frac{1}{m})$, and $b = (\frac{1}{k}, \frac{1}{k}, ..., \frac{1}{k}, 0, ..., 0)$ where the first k-components are $\frac{1}{k}$. By the definition of the DM(SE)$^+$, it follows:

$$\begin{aligned} ||p - b||^2 &= \sum_1^k (p_i - \frac{1}{k})^2 + \sum_{j=k+1}^m p_j^2 \\ &\le \sum_1^m (p_i - \frac{1}{m})^2 = ||p - u||^2 \end{aligned}$$

By computation, we have

$$\sum_{i=1}^m p_i^2 - \frac{2}{k}\sum_{i=1}^k p_i + \frac{1}{k} \le \sum_{i=1}^m p_i^2 - \frac{1}{m}$$

It follows that

$$\begin{aligned} \sum_{i=1}^k p_i &\ge (\frac{1}{k} + \frac{1}{m})\frac{k}{2} \\ &= \frac{1}{2} + \frac{k}{2m} > \frac{1}{2} \end{aligned}$$

□

In the above proof, we compared the distance from a probability function to the point representing its belief core with the distance to the centre point. This simple comparison guarantees that a belief core selected by DM(SE)$^+$ is more likely than its complement.

Now we turn to the last question of whether DM(SE)$^+$ is coherent. First, we will, by means of a counterexample, show that it is not coherent. Then we will find a weaker but sufficiently plausible notion of coherence that DM(SE)$^+$ can satisfy. In the right one in Figure 3.10, consider the point $p = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$. Among the elements of U^3, all of b_1, b_{12}, b_{13}, and b_{123} have the minimum distance to the point p. Since DM(SE)$^+$ allows any tie breaking rule, we can consider a rule that assigns to P $\{w_1, w_2\}$. Then, although $P(w_2) = P(w_3)$, the world w_2 belongs to the belief core, but w_3 does not, which violates the definition of a coherent BR. We can also check this using geometrical characterization in Theorem 3.5: p belongs to the preimage-region of $(\frac{1}{2}, \frac{1}{2}, 0)$ under the above tie breaking rule. However, p does not belong to the interior of the Voronoi cell of the generator $(\frac{1}{2}, \frac{1}{2}, 0)$ with respect to U_2^3, but located on the Voronoi cell boundary, which violates geometrical characterization of a coherent BR.

Notice that only the three points — $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, $(\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$ and $(\frac{1}{6}, \frac{1}{6}, \frac{2}{3})$ — violate the condition for DM(SE)$^+$ to be coherent. This makes us expect that any point inside the preimage-regions of DM(SE)$^+$ satisfies the condition to be coherent. As Theorem 3.8 will show, this conjecture is true. To formulate the exact notion of " inside", we need some basic topological notions.

Given any $x \in \mathbb{R}^n$ and $\epsilon > 0$, we define the open ball of the centre x and radius ϵ: $B_\epsilon(x) = \{y \in \mathbb{R}^n | \ ||x - y|| < \epsilon\}$; for $A \subseteq \mathbb{R}^n$, we define the interior of A:

$$int(A) = \{x \in A | \ \exists \epsilon > 0 B_\epsilon(x) \subseteq A\};$$

the closure of A:

$$cl(A) = \{x \in \mathbb{R}^n | \ \forall \epsilon > 0 \ B_\epsilon(x) \cap A \neq \emptyset\};$$

the boundary of A: $bd(A) = cl(A) \setminus int(A)$. Since the dimension of the probability simplex $\triangle^m \subseteq \mathbb{R}^m$ is $m-1$, $int(\triangle^m)$ is empty. To deal with low-dimensional objects placed in higher dimensional spaces, we need several refiner notions. A subset A of $\mathbb{R}^n$ is said to be affine if $l[x, y] \subset A$ for all $x, y \in A$ where $l[x, y] = \{\lambda x + (1-\lambda)y | \ \lambda \in \mathbb{R}\}$; for $A \subseteq \mathbb{R}^n$, we define the affine hull of A: $aff(A) = \bigcap\{C \subseteq \mathbb{R}^m | \text{ C is affine and } A \subseteq C\}$; We define the relative interior of A:

$$ri(A) = \{x \in A | \ \exists \epsilon > 0 (B_\epsilon(x) \cap aff(A)) \subseteq A\};$$

the relative boundary of A:

$$rb(A) = cl(A) \setminus ri(A).$$

Using these notions, we can make the condition of being inside a preimage-region formally precise: it refers to the condition of being *in the relative interior* of a preimage-region. Accordingly, we define slightly weaker notion of coherence that is applied only in the relative interior of every preimage-region as follows.

Definition 3.18. *A rational BR G is interior-coherent (coherent$^\circ$) if for all P in the domain of G and $B \subset W$ such that $p \in ri(G^{-1}(b))$, if $P(v) \leq P(w)$, and if $v \in B$, then $w \in B$ for all $v, w \in W$.*

We can formulate the relative interior of the preimage-region $G^{-1}(b)$ of $b \in U^m$ when G is a DM(SE)$^+$:

$$ri(G^{-1}(b)) = \{p \in ri(\triangle^m)|\ ||p - b|| < ||p - b'||\ \forall b'(\neq b) \in U^m\}$$

from which follows that

$$ri(G^{-1}(b)) = V(b|U^m)^\circ \cap ri(\triangle^m)$$

as mentioned before. So, for a DM(SE)$^+$ to be coherent$^\circ$, not all point $p \in \triangle^m$, but only the points in the above set for some $b \in U^m$ have to be mapped to a belief core generating a coherent belief state. As one might expect, this is the case as shown in the following theorem.

Theorem 3.8. *DM(SE)$^+$ is coherent$^\circ$.*

Proof. Let $|W| = m$ and G be a DM(SE)$^+$. Consider the restriction G' of G to the union of the relative interiors of the preimage-regions of all points in U^m. Obviously G' is a BR as well (it does not have the universal domain though) and $G'^{-1}(b) = ri(G^{-1}(b))$. Since $ri(G^{-1}(b)) = V(b|U^m)^\circ \cap ri(\triangle^m) \subseteq V(b|U^m_k)^\circ$ by geometrical characterization of coherence in Theorem 3.5, G' is coherent, that is, G is coherent$^\circ$. $\square$

For DM(SE)$^+$, in my opinion, being interior-coherent is sufficiently strong because this rule allows any tie-breaking rule, which presupposes that the decision at the boundary case is not so relevant.

3.5 Belief Binarization and Convexity

This section is based on joint work with Chisu Kim. In this section, we will investigate how the convexity requirement can be applied in the belief binarization problem. The notion of *convexity*[13] has long been studied and discussed in different fields in formal philosophy: according to *convex Bayesianism*, an epistemic state is represented by a set of probability functions, and the set must be *convex*[14]; according to *the conceptual spaces program*, concepts can be geometrically represented as regions in conceptual space equipped with a similarity measure (usually metric), and Gärdenfors (2000) argued that *'natural'* concepts have the feature of being represented by a *convex* set of points, and defended it as "a principle of cognitive economy"(p.70)[15]; in the probabilistic opinion pooling problem, the advocates of *linear pooling* methods implicitly support the convexity property because linear pooling methods produce a convex combination of the input probability functions. However, the notion of convexity has not been studied and discussed in the belief binarization problem, although many aspects of belief binarization are, to some extent, related to the research topics mentioned above. In this section, we begin the first discussion of this issue.

First of all, let us think about why we should respect the convexity norm in the binarization problem. In the following, we will present three possible arguments for the convexity norm, which are closely related to the above-mentioned research.

(1) (convex Bayes) Consider an epistemic state that is represented by a binary belief. Since binary beliefs are a coarser representation model of epistemic states than probability functions in terms of information quantity, there might be multiple probability functions all of which also represent the epistemic state together. Which probability functions correspond to the epistemic state? A binarization method can be viewed as a method to determine this correspondence. Accordingly, the set of such multiple probability functions can be understood as the preimage of a binary belief under a binarization method. According to convex Bayesianism, the set of such multiple probability functions should be convex because they represent an epistemic state.

(2) (Unanimity) Collective belief binarization might be a procedure for heterogeneous belief aggregation where multiple individual probability functions are aggregated into a collective probability function, and then reduced to a collective binary belief. In this context of heterogeneous belief aggregation, we suggest

[13] Let us remind the definition of convex set: a set $R \subseteq \mathbb{R}^m$ is called convex if for all $r, r' \in R$ it holds that $\alpha r + (1-\alpha)r' \in R$ for all $\alpha \in [0,1]$. Geometrically speaking, a convex set includes every point lying on the line segment connecting any two elements of the set. We call a linear average of two points a convex combination of two points. A rephrasing of the definition is that a set is convex if the set is closed under any convex combination.

[14] For convex Bayesianism see Levi (1980), Sterling and Morrell (1991). For critical discussion see Kyburg and Pittarelli (1996).

[15] For the conceptual space program see Gärdenfors (2000), Douven and Gärdenfors (2018). For a critical discussion about the convexity of natural concepts see Mormann (1993) and Hernandez-Code (2017).

the unanimity norm concerning collective belief binarization: if every individual probability function is rationally compatible with a binary belief according to the binarization method, then the pooled collective probability is as well. It seems particularly plausible, if the disagreement between individual probabilities does not stem from informational asymmetry between individuals. Now, we add the claim that linear pooling methods are permissible based on the normative arguments for linear pooling. Then it follows that if individual probability functions are compatible with a binary belief, then any convex combination of them is as well.

We can also draw the same conclusion based on a slightly different unanimity norm: assume that probability functions P and P' are compatible with a binary belief Bel according to the binarization method. If P approaches P', i.e., moves to a probability function P'' between P and P', we can say that the disagreement between P and P' is decreasing. Therefore, P'' should also be compatible with Bel. And we can take P'', a point between P and P', as any convex combination of P and P'.

(3) (Belief as a Natural Concept) The probability simplex can be viewed as the conceptual space with some distance measure, and rational binary beliefs can be thought of as natural concepts partitioning the probability simplex. Since every natural concept should be convex, each rational binary belief should have a convex region on the probability simplex.[16]

Although each of these arguments has some unanswered issues that deserve a lengthy discussion, we will leave them for further research.[17] For the rest of this section, our focus will be on more concrete and practical questions: how we can formulate convexity norms in the belief binarization context, and which binarization methods satisfy which convexity norms.

Regarding this problem, there are some problems we should be concerned about. First, many binarization rules do not satisfy the most straightforward convexity norm that each preimage-region of a BR should be convex. In particular, all the binarization rules introduced in the previous section do not satisfy it. Thus, we will present some weaker notions of convexity. Second, many binarization methods are not functions. Some binarization methods such as HT^r are relations between credences and

[16]If we adopt a typical method to generate natural concepts used in conceptual spaces program, the probability simplex can be partitioned in the following way: we set the distance measure to be the Euclidean metric, and for any rational binary belief, we set the uniform distribution on the belief core of the binary belief to be the prototype of the binary belief; from the set of prototypes we generate the Voronoi diagram consisting of the Voronoi cells that correspond to binary beliefs. Notice that this is exactly the same way as in the case of $\mathrm{DM(SE)}^+$.

[17]To defend (convex Bayes), we should ultimately justify the convex Bayesianism. To advocate (Unanimity), we should address the question whether there are differences between collective belief binarization and individual belief binarization. Moreover, we need to justify linear pooling. To vindicate (Belief as a Natural Concept), we should address how to interpret it as a normative requirement. And we need to explain why rational binary beliefs can be interpreted as natural concepts.

beliefs, and at the same time they can be viewed as correspondences since we can build a correspondence from any relation and vice versa[18]. Therefore, we will formulate convexity requirements imposed on relations and correspondences, respectively. Third, many binarization methods involve some intermediate stage between credences and binary beliefs. For example, $\mathrm{HT}^r(\mathrm{S})$ and CS^s employ a kind of ordinalization of worlds to find belief cores. Hence, it would also be important to study the convexity requirement on ordinalization in addition to binarization. Accordingly, this section consists of three topics: (1) convexity of belief binarization functions, (2) convexity of generalized belief binarization methods, and (3) convexity of ordinalization.

Before starting, let us outline the main results of this section. First, we will show that we can make convexity norms sufficiently refined to classify different binarization methods. We will formulate seven different types of convexity norms, and examine which binarization methods satisfy them. Second, we will see that distance-based binarization methods are generally better in terms of the convexity requirements.

Convexity of belief binarization functions We now present the most straightforward notion of convexity for a BR — preimage-convexity — and two weaker notions: interior-preimage-convexity and holistic monotonicity. Interior-preimage-convexity requires only the relative interior of each preimage-region to be convex.[19] This might be sufficiently strong if the BR allows arbitrary choices for the boundary cases like $\mathrm{DM(SE)}^+$. Holistic monotonicity requires that if a BR assigns to a probability function P a belief core B, then the BR should assign the same belief core B to any convex combination of P and $U(B)$. The reason why we call it 'holistic' monotonicity is that it takes into account the full information of the probability function in contrast to event-wise or world-wise monotonicity. This weaker notion of convexity seems even more plausible than preimage-convexity, because if P is changed to a convex combination with $U(B)$, then it means that P moves to B and we can say that the information of the resulting binary belief is more reflected. Let us give the formal definition of them.

Definition 3.19. *Let W be a finite non-empty set. Let G be a rational BR with the domain $\mathbb{P}(W)$ (the set of all probability functions on $(W, \mathcal{P}(W))$).*

(1) G satisfies preimage-convexity if for all $P, P' \in \mathbb{P}(W)$, if $G(P) = G(P')$, then

$$G(\alpha P + (1-\alpha)P') = G(P)(= G(P'))$$

for all $\alpha \in [0,1]$.

(2) G satisfies interior-preimage-convexity (preimage°-convexity) if for all $P, P' \in \mathbb{P}(W)$, if $p, p' \in ri(G^{-1}(b))$ for some $b \in U^m$, then

$$G(\alpha P + (1-\alpha)P') = G(P)(= G(P'))$$

for all $\alpha \in [0,1]$.

[18]For example, Dietrich & List (2021) sees HT^r as a correspondence.

[19]For the formal definition of relative interior and preimage-region see the previous section.

(3) A BR G satisfies holistic-monotonicity (h-monotonicity) if for all $P \in \mathbb{P}(W)$ and $B(\neq \emptyset) \subseteq W$, if $G(P) = B$, then

$$G(\alpha P + (1-\alpha)U(B)) = B$$

for all $\alpha \in [0, 1]$.

Note that interior-preimage-convexity and holistic monotonicity do not imply each other: geometrically speaking, h-monotonicity implies the star-convexity of every preimage-region.[20] Since the relative interior of star-shaped regions may not be convex, h-monotonicity does not imply preimage°-convexity. For the other direction, the counterexample will be provided when we examine CCT^g later.

Now we examine whether the binarization rules in the last section satisfy each kind of convexity. As in the previous section, to say a rule satisfies a convexity norm means that the rule satisfies the norm for all contexts embracing all numbers of possible worlds and all values of parameters. Thus, to say a rule does not satisfy a norm means that there is a context — i.e., a probability space and a value of parameter — in which the rule does not satisfy the norm.

First of all, let us check some easily obtainable results and get some geometrical intuitions regarding preimage-convexity and preimage°-convexity through the figures in the previous section. For a start, let us consider $HT^r(S)$. Figure 3.8 shows that $HT^r(S)$ and $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$ do not satisfy preimage-convexity. In particular, the preimage-region of the centre is not convex in each simplex in the figure. In the left simplex, we can see that the region where preimage-convexity is violated consists of three line segments, and therefore whose relative interior is empty.[21] This observation shows that $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$ satisfies preimage°-convexity for $|W| = 3$. To the question of whether this fact generally holds for $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$, the counterexample in the subsequent theorem answers negatively.[22] We now turn to CS^s. The left simplex in Figure 3.9 shows that

[20] A subset $A \subseteq \mathbb{R}^m$ is star-convex iff there is a $p \in A$ such that for all $q \in A$ it holds that $\alpha p + (1-\alpha)q \in A$ for all $\alpha \in [0, 1]$.

[21] This shape of the preimage-region of the centre implies that $\mathrm{HT}^{\frac{1}{2}}(\mathrm{S})$ violates what we call the strong suspension principle, which is defined as following: a rational BR G satisfies the strong suspension principle if for all $b \in U^m$ there is $\epsilon > 0$ such that $(B_\epsilon(b) \cap \triangle^m) \subseteq G^{-1}(b)$. This norm says that we should not allow a drastic change near the belief-core-point. Although $\mathrm{HT}^r(\mathrm{S})$ satisfies the suspension principle ($G(U(B)) = B$ for all $B(\neq \emptyset) \subseteq W$), it does not satisfy the strong suspension principle.

[22] One may wonder if $\mathrm{HT}^{\frac{1}{2}}(\mathrm{D})$ (for the definition of $\mathrm{HT}^r(\mathrm{D})$ see footnote[9]) satisfies preimage-convexity. We present a counterexample here. Consider the following probability distributions on $W = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7\}$:

	w_1	w_2	w_3	w_4	w_5	w_6	w_7
P	0.5	0.25	0.05	0.05	0.05	0.05	0.05
Q	0.05	0.25	0.5	0.05	0.05	0.05	0.05
$0.5P + 0.5Q$	0.275	0.25	0.275	0.05	0.05	0.05	0.05

According to the $\mathrm{HT}^{\frac{1}{2}}(\mathrm{D})$, the belief core of P and Q should be W while the belief core of $0.5P + 0.5Q$ should be $\{w_1, w_2, w_3\}$.

CS^s do not satisfy preimage-convexity: the preimage-region of the midpoint of each side is not convex. It is also shown that CS^s does not satisfy even the preimage°-convexity.

Next, we move to CCT^g. In the left simplex in Figure 3.10, the preimage of the centre has three line segments sticking out. Since the relative interior of the preimage-region of the centre dose not include the three line segments, we can see that CCT$^{0.75}$ satisfies preimage°-convexity for $|W| = 3$.[23] This makes us ask whether it can be generalized; we will answer this question positively in the following theorem. As to DM(SE)$^+$, see the right simplex in Figure 3.10. Consider the case where p, p' lie on the line segment on which it holds that $\text{argmin}_{B\subseteq W} D(P, B) = \text{argmin}_{B\subseteq W} D(P', B) = \{\{w_1\}, \{w_1, w_2\}\}$. Since DM(SE)$^+$ allows any tie-breaking rules, we can consider a rule G such that $G(P) = \{w_1\}, G(P') = \{w_1\}$, and $G(0.5P + 0.5P') = \{w_1, w_2\}$. Then G does not satisfy the preimage-convexity. However, from the figure, we can easily expect that DM(SE)$^+$ satisfies the preimage°-convexity, which turns out to be true as the next theorem shows. Now let us prove all the above claims about preimage-convexity and preimage°-convexity.

Theorem 3.9. *(1) $HT^r(S)$ does not satisfy the preimage°-convexity.*

(2) CS^s does not satisfy the preimage°-convexity.

(3) CCT^g does not satisfy preimage-convexity.

(4) CCT^g satisfies preimage°-convexity.

(5) DM(SE)$^+$ satisfies the preimage°-convexity.

Proof. (1) Let G be HT$^{\frac{1}{2}}$(S). Consider the following probability distributions on $W = \{w_1, w_2, w_3, w_4\}$:

	w_1	w_2	w_3	w_4
P	0.44	0.25	0.2	0.11
Q	0.25	0.44	0.2	0.11
$0.5P + 0.5Q$	0.345	0.345	0.2	0.11

For any $p' \in B_{0.0001}(p) \cap \triangle^4$ and any $q' \in B_{0.0001}(q) \cap \triangle^4$ it holds that $G(P') = G(Q') = \{w_1, w_2, w_3\}$, which implies that $p, q \in ri(G^{-1}(\{w_1, w_2, w_3\}))$. However, $G(0.5P + 0.5Q) = \{w_1, w_2\}$, which violates the preimage°-convexity.

(2) Let G be CS$^{1.4}$. Consider the following probability distributions on $W = \{w_1, w_2, w_3\}$:

[23] From this difference between the shapes of the preimage-region of HT$^{\frac{1}{2}}$(S) and CCT$^{0.75}$ we can see that CCT$^{0.75}$ satisfies the strong suspension principle for the case of $|W| = 3$. However, we will show that it does not hold in the more general setting. See footnote[26].

	w_1	w_2	w_3
P	0.41	0.32	0.27
Q	0.32	0.41	0.27
$0.5P + 0.5Q$	0.365	0.365	0.27

Note that for any $p' \in B_{0.0001}(p) \cap \triangle^3$ and any $q' \in B_{0.0001}(q) \cap \triangle^3$ it holds that $G(P') = G(Q') = \{w_1, w_2, w_3\}$, which implies that $p, q \in ri(G^{-1}(\{w_1, w_2, w_3\}))$. However, $G(0.5P + 0.5Q) = \{w_1, w_2\}$, which violates the preimage$^\circ$-convexity.

(3) Let G be a CCT$^{0.6}$. Consider the following probability distributions on $W = \{w_1, w_2, w_3\}$:

	w_1	w_2	w_3
P	0.5	0.25	0.25
Q	0.25	0.5	0.25
$0.5P + 0.5Q$	0.375	0.375	0.25

$G(P) = G(Q) = \{w_1, w_2, w_3\}$ but $G(0.5P + 0.5Q) = \{w_1, w_2\}$, which violates the preimage-convexity.

(4) Let G be a CCT^g and $P, P' \in \mathbb{P}(W)$. We will show that if $G(P) = G(P') = B$ but $G(\alpha P + (1-\alpha)P') = B' \neq B$ for some $\alpha \in [0, 1]$, then at least one of p and p' lies on $rb(G^{-1}(b))(:= cl(G^{-1}(b)) \setminus ri(G^{-1}(b))\ )$.

Assume that $G(P) = G(P') = B$ but $G(P''_\alpha) = B' \neq B$ where we write P''_α for $\alpha P + (1-\alpha)P'$. Since $P(B), P'(B) \geq g$, $\min_{w \in B} P(w) > \max_{w \in \overline{B}} P(w)$, and $\min_{w \in B} P'(w) > \max_{w \in \overline{B}} P'(w)$, we have $P''_\alpha(B) \geq g$ and

$$\min_{w \in B} P''_\alpha(w) \geq \alpha \min_{w \in B} P(w) + (1-\alpha) \min_{w \in B} P'(w)$$
$$> \alpha \max_{w \in \overline{B}} P(w) + (1-\alpha) \max_{w \in \overline{B}} P'(w) \geq \max_{w \in \overline{B}} P''_\alpha(w)$$

This indicates that not only B' but also B is one of the candidates (the P''_α - coherent sets with probability at least g) of belief cores of $G(P''_\alpha)$. Since all candidates of belief cores of $G(P''_\alpha)$ are linearly ordered by subset relation, we have

$$B' \subsetneq B$$

(for if not, it would be the case that $G(P''_\alpha) \neq B'$).

Now observe that it is not the case that $P(B') < g$ and $P'(B') < g$ because $P''_\alpha(B') \geq g$. W.l.o.g, let $P(B') \geq g$. We will show that p is on $rb(G^{-1}(b))$, that is, for any $\epsilon > 0$ there exists $Q \in B_\epsilon(p) \cap \triangle^m$ satisfying $G(Q) \neq B$. To see this, take one world in $\text{argmin}_{w \in B} P(w)$, say $w_{B,P}$, and another world in B, say v, and let

$$Q(w) := \begin{cases} P(w) - \delta & \text{if} \quad w = w_{B,P} \\ P(w) + \delta & \text{if} \quad w = v \\ P(w) & o/w \end{cases}$$

where δ is a sufficiently small positive number so as to satisfy $Q \in B_\epsilon(p) \cap \triangle^m$. It is possible because $P(w_B) > 0$ (a belief core generated by CCT^g cannot contain a world with probability 0) and $P(v) < 1$ (since $\emptyset \neq B' \subsetneq B$, B has at least two elements with non-zero probability). We can see that $B \setminus \{w_{B,P}\}$ is Q-coherent. Moreover, since we have $B' \subsetneq B$ and $w_{B,P}$ has the minimum probability in B, it follows that

$$Q(B \setminus \{w_{B,P}\}) > P(B \setminus \{w_{B,P}\}) \geq P(B') \geq g$$

This shows that $G(Q) \neq B$ because $B \setminus \{w_{B,P}\}$ is a smaller set than B and it satisfies the conditions to become $G(Q)$ except "the smallest" condition in the definition of $G(Q)$ and thus, B can never become $G(Q)$.

(5) Let $|W| = k$, $b \in U^m$, and G be a DM(SE)$^+$. Then $ri(G^{-1}(b)) = \{p \in \triangle^m | \; ||p-b||^2 < ||p-b'||^2$ for all $b'(\neq b) \in U^m\}$. For any $b' \in U^m$ it holds that

$$\begin{aligned} ||p-b||^2 < ||p-b'||^2 \text{ iff } & ||p||^2 - 2\langle p, b\rangle + ||b||^2 < ||p||^2 - 2\langle p, b'\rangle + ||b'||^2 \\ \text{iff } & 2\langle p, b'-b\rangle < ||b'||^2 - ||b||^2. \end{aligned}$$

Then we have

$$ri(G^{-1}(b)) = \bigcap_{b' \in U^m} \{p \in \triangle^m | \; 2\langle p, b'-b\rangle < ||b'||^2 - ||b||^2\} \setminus rb(\triangle^m)$$

Since the inner product is linear in the first argument, $ri(G^{-1}(b))$ is closed under convex combination.

□

Regarding DM(SE)$^+$, we can also prove the above theorem using the properties of the Voronoi diagram. Both $V(b|U^m)$ and $\triangle^m$ are a closed convex set[24], so is their intersection. Since $ri(G^{-1}(b)) = ri(V(b|U^m) \cap \triangle^m)$, and the relative interior of a convex set is also convex, we conclude that $ri(G^{-1}(b))$ is convex.

We now move to a discussion of h-monotonicity. For a start, let us consider the figures in the previous section. In Figure 3.8, 3.9, and 3.10, we can easily check that all 4 binarization rules satisfy h-monotonicity for the case $|W| = 3$. This observation make us expect that h-monotonicity might be very weak requirement, and the results for the $|W| = 3$ will be true in the generalized setting. This conjecture turns out to be true with one exception as the next theorem shows: except for CCT^g, the rest of the rules satisfy h-monotonicity. It is impressed that CCT^g does not satisfy h-monotonicity, which shows that preimage°-convexity does not imply h-monotonicity. Now let us prove all the above claims about h-monotonicity.

[24]It is well-known fact that every Voronoi cell is convex. Let us prove it. Let $R = \{r_1, ..., r_l\} \in \mathbb{R}^m$. $V(r_i|R) = \{p \in \mathbb{R}^m | \; ||p-r_i|| \leq ||p-r_j||$ for all $r_j \in R\}$. For any $j \in R$ it holds that

$$||p-r_i|| \leq ||p-r_j|| \text{ iff } ||p-r_i||^2 \leq ||p-r_j||^2 \text{ iff } 2\langle p, r_j - r_i\rangle \leq ||r_j||^2 - ||r_i||^2.$$

Then we have $V(r_i|R) = \bigcap_{r \in R}\{p \in \mathbb{R}^m | \; 2\langle r - r_i, p\rangle \leq ||r||^2 + ||r_i||^2\}$. Since the inner product is linear in the first argument, $V(r_i|R)$ is closed under convex combination.

Theorem 3.10.

(1) CCT^g does not satisfy h-monotonicity.

(2) $HT^r(S)$ satisfies h-monotonicity.

(3) CS^s satisfies h-monotonicity.

(4) $DM(SE)^+$ satisfies h-monotonicity.

Proof. (1) Let G be a $\mathrm{CCT}^{0.7}$. Consider the following probability distributions on $W = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7\}$:

	w_1	w_2	w_3	w_4	w_5	w_6	w_7
P	0.17	0.17	0.17	0.17	0.14	0.09	0.09
$U(B)$	0.2	0.2	0.2	0.2	0.2	0	0
$0.5p + 0.5U(B)$	0.185	0.185	0.185	0.185	0.17	0.045	0.045

where $B = \{w_1, w_2, w_3, w_4, w_5\}$. $G(P) = B$, but $G(0.5P+0.5U(B)) = \{w_1, w_2, w_3, w_4\}$, which violates h-monotonicity.

(2) Suppose $p \in G^{-1}(b), \hat{w} \in \operatorname{argmin}_{w\in B} P(w), |B| = k(\leq m = |W|)$. Let $\alpha \in (0, 1)$. We will show that $(q :=)\alpha p + (1 - \alpha)b \in G^{-1}(b)$. Since $\hat{w} \in \operatorname{argmin}_{w\in B} Q(w)$, We need to prove that (i)

$$Q(\hat{w}) > \frac{r}{1-r} Q(\overline{B})$$

and (ii) for all $v(\neq \hat{w}) \in B$ it holds that

$$Q(v) \leq \frac{r}{1-r} Q(\{w \in W | Q(v) < Q(w)\})$$

To (i): Since $p \in G^{-1}(b)$, we have $P(\hat{w}) > \frac{r}{1-r} P(\overline{B})$ which implies the following inequality:

$$Q(\hat{w}) = \alpha P(\hat{w}) + \frac{1-\alpha}{k} > \frac{\alpha r}{1-r} P(\overline{B}) = \frac{r}{1-r} Q(\overline{B})$$

To (ii): Let $v \in B$ with $v \neq \hat{w}$. Define $B' = \{w \in B |\ Q(w) \geq Q(v)\}$ with $|B'| = l < k$. We need to show the following inequality:

$$\begin{aligned} Q(v) &= \alpha P(v) + \frac{1-\alpha}{k} \\ &\leq \frac{r}{1-r}(\alpha P(B \setminus B') + \frac{1-\alpha}{k}(k-l) + P(\overline{B})) = \frac{r}{1-r} Q(\{w \in W | Q(w) < Q(v)\}) \end{aligned}$$

Since B is the smallest P-stable^r set, $v(\neq \hat{w}) \in B$, and $\alpha \in (0, 1)$, we have the following inequalities:

$$\begin{aligned} \alpha P(v) &\leq \frac{\alpha r}{1-r}(P(B \setminus B') + P(\overline{B})) \\ &< \frac{r}{1-r}(\alpha P(B \setminus B') + P(\overline{B})) \end{aligned} \tag{3.3}$$

Since $r > 0.5$ and $l < k$, we have $\frac{r}{1-r}(k-l) > 1$, which implies

$$\frac{r}{1-r}\frac{1-\alpha}{k}(k-l) - \frac{1-\alpha}{k} \geq 0 \tag{3.4}$$

Adding the LHS of 3.4 to the RHS of 3.3, we have the following inequality:

$$\alpha P(v) \leq \frac{r}{1-r}(\alpha P(B \setminus B') + P(\overline{B})) + \frac{r}{1-r}\frac{1-\alpha}{k}(k-l) - \frac{1-\alpha}{k}$$

which implies the desired result.

(3) Suppose that $p \in G^{-1}(b), |B| = k, \hat{w} \in \operatorname{argmax}_{w \in B} P(w)$, and $\alpha \in (0,1)$. We will show that $(q :=)\alpha p + (1-\alpha)b \in G^{-1}(b)$. It suffices to show that (1) for all $w \in B$ $sQ(w) \geq Q(\hat{w})$, (2) for all $w \in \overline{B}$ $sQ(w) < Q(\hat{w})$. For $w \in B$ let us rewrite (1) as the following:

$$s(\alpha P(w) + (1-\alpha)\frac{1}{k}) \geq \alpha P(\hat{w}) + (1-\alpha)\frac{1}{k}$$

which is equivalent to the following:

$$sP(w) + (s-1)\frac{1-\alpha}{\alpha}\frac{1}{k} \geq P(\hat{w})$$

Since $s > 1$, it holds that $(s-1)\frac{1-\alpha}{\alpha}\frac{1}{k} \geq 0$. Then it suffices to show that $sP(w) \geq P(\hat{w})$ which is true because $w \in B$.
Let us turn to (2). For $v \in \overline{B}$ we need to show the following:

$$sP(v) < P(\hat{w}) + \frac{1-\alpha}{\alpha}\frac{1}{k}$$

which is true because $\frac{1-\alpha}{\alpha}\frac{1}{k} \geq 0$ and $v \notin B$.

(4) Suppose that $p \in G^{-1}(b)$, $\alpha \in [0,1)$. Define $q = \alpha p + (1-\alpha)b$. Note that $(V(b|U^m)^\circ \cap \triangle^m) \subset G^{-1}(b) \subset (V(b|U^m) \cap \triangle^m)$. Consider the case where $p \in V(b|U^m)^\circ \cap \triangle^m$. Since $V(b|U^m)^\circ \cap \triangle^m$ is convex set, we have $q \in V(b|U^m)^\circ \cap \triangle^m$. Consider the case where $p \in (V(b|U^m) \setminus V(b|U^m)^\circ) \cap \triangle^m$. Since $V(b|U^m)^\circ$ is convex, $b \in V(b|U^m)^\circ (= ri(V(b|U^m))$, and $p \in V(b|U^m) (= cl(V(b|U^m)^\circ))$, we can use a well-known geometrical property of convex set: the half-open line segment $(p, b] \subset V(b|U^m)^\circ$.[25] Since $q \in (p, b]$, we have $q \in (V(b|U^m)^\circ \cap \triangle^m) \subset G^{-1}(b)$.

□

As for CCT^g, one may wonder if the counterexample in the above proof is also applicable to the problem of preimage$^\circ$-convexity of CCT^g. Although p and b generate a non-convex case, we can notice that b lies on the relative boundary of $G^{-1}(b)$ since even a small change in the coordinates of b yields a different belief core. For example, consider the point $p' = b + (\epsilon, -\epsilon, 0, 0, 0, 0, 0)$ where ϵ is an arbitrary small positive

convexity	$HT^r(S)$	CS^s	CCT^g	DM(SE)$^+$
preimage-convexity	X	X	X	X
preimage°-convexity	X	X	O	O
h-monotonicity	O	O	X	O

Table 3.4: Convexity and Binarization Rules. In the table 'O'/'X' means that the rule satisfies/does not always satisfy the property.

real number. Then $G(P') = \{w_1, w_3, w_4, w_5\}$ which is different from B.[26] Therefore this is not a counterexample to the preimage°-convexity of CCT^g.

All results so far are collected in Table 3.4. It is noteworthy that all of four BRs do not satisfy the preimage-convexity, DM(SE)$^+$ satisfies both weaker notions of convexity, and the remaining three rules satisfy one of two norms. Preimage°-convexity is favorable to CCT^g while h-monotonicity is favorable to HT^r and CS^s.

Convexity of generalized binarization methods We now extend bianrization functions to more generalized binarization methods, namely relations and correspondences. Contrast to functions, correspondences allow multiple outputs. When our goal is to find a most comprehensive rational bridge principle between credences and beliefs, rather than to make a decision to choose only one option — e.g., given a probability function, to choose only one binary belief rationally compatible with the probability function —, belief binarization relations can be appropriate. HT^r is an example of a belief binarization relation.

To state it formally, let us first define a binarization correspondence as the following: a binarization correspondence (BC) on $(W, \mathcal{P}(W))$ is a function that assigns to any $P \in \mathbb{P}(W)$ a set of some binary beliefs (a function from $\mathbb{P}(W)$ to $\mathcal{P}(\{0,1\}^{\mathbb{P}(W)})$). In this section, we will study two BCs. The first one is just the induced one by HT^r: C is the BC induced by HT^r if $C(P) = \{Bel \in \{0,1\}^{\mathbb{P}(W)} | (P, Bel) \in HT^r\}$; the second one is a new distance-based BC, called DM(SE), defined as follows

Definition 3.20. *Let C be a BC on $(W, \mathcal{P}(W))$. C is DM(SE) if $C(P)$ is the set of all binary beliefs induced by some belief core $B \in \mathrm{argmin}_{B\subseteq W} D(P,B) := \{B \subseteq W| \ ||p-b||^2 \leq ||p-b'||^2 \ \forall b' \in U^m\}$ for all $P \in \mathbb{P}(W)$.*

DM(SE) is just DM(SE)$^+$ without tie-breaking rule. In our opinion, DM(SE) is the most natural distance-based binarization method since distance-based binarization methods always come up with boundary cases, and it is difficult to justify a specific tie-breaking rule unless another criteria than distance minimization is given in advance.

Now let us formulate a binarization relation. A binarization relation on $(W, \mathcal{P}(W))$ is a subset of $\mathbb{P}(W) \times \{0,1\}^{\mathbb{P}(W)}$. We read "$(P, Bel) \in R$" as (P, Bel) satisfies R. Besides HT^r, we can also consider the binarization relation induced by DM(SE): R is the binarization relation induced by DM(SE) if it holds that $(P, Bel) \in R$ iff $B \in$

[25]See Prop. 1.2 in Hug & Weil (2020).

[26]This behavior of b implies that CCT^g does not satisfy the strong suspension principle either although it satisfies the suspension principle ($G(b) = B$ for all $b \in U^m$).

$\mathrm{argmin}_{B\subseteq W}\, D(P,B)$ where B is the belief core of Bel. Now let us define the convexity of binarization correspondence and binarization relation.

Definition 3.21 (C-convexity and R-convexity)**.** *(1) A BC C satisfies the convexity of binarization correspondence (C-convexity) iff for all $P, P' \in \mathbb{P}(W)$, if $C(P) = C(P')$, then*

$$C(\alpha P + (1-\alpha)P') = C(P)(= C(P'))$$

for all $\alpha \in [0,1]$.

(2) A binarization relation R satisfies the convexity of binarization relation (R-convexity) iff for all $P, P' \in \mathbb{P}(W)$ and for all binary belief Bel on $(W, \mathcal{P}(W))$, if (P, Bel) and (P', Bel) satisfy R, then $(\alpha P + (1-\alpha)P', Bel)$ satisfies R for all $\alpha \in [0,1]$.

Note that C-convexity and R-convexity do not imply each other.[27] As for $\mathrm{DM(SE)}^+$, the only obstacle to satisfying preimage-convexity was the boundary cases, and thus we expect that DM(SE) and the induced relation by it satisfy C-convexity and R-convexity, respectively. As to HT^r, let us consider two simplexes in Figure 3.8. Since the preimage-region of the centre is not convex, and there is no belief core bigger than the whole set W, we can see that HT^r does not satisfy C-convexity. By contrast, the following theorem shows that HT^r satisfies R-convexity. Let us prove all the claims.

Theorem 3.11.

(1) HT^r satisfies R-convexity.

(2) The BC induced by HT^r does not satisfy C-convexity.

(3) $DM(SE)$ satisfies C-convexity.

(4) The relation induced by $DM(SE)$ satisfies R-convexity.

Proof. (1) We will use Lemma 3.3 in the previous section. Let P, $P' \in \mathbb{P}(W)$, $\alpha \in (0,1)$ and $P''_\alpha = \alpha P + (1-\alpha)P'$. Consider the case where $Bel(\emptyset) = 1$, which implies that $Bel^{-1}(1) = \mathcal{P}(W)$. Then (P''_α, Bel) vacuously satisfies HT^r. Now assume that (P, Bel) and (P', Bel) with $\emptyset \neq \cap Bel^{-1}(1) = B$ satisfy HT^r. Then we have

$$\alpha \min_{w\in B} P(w) > \alpha \frac{r}{1-r} P(\overline{B});\ (1-\alpha) \min_{w\in B} P'(w) > (1-\alpha)\frac{r}{1-r} P'(\overline{B})$$

Adding two inequalities we have

$$\alpha \min_{w\in B} P(w) + (1-\alpha) \min_{w\in B} P'(w) > \frac{r}{1-r} P''_\alpha(\overline{B})$$

[27]To see that C-convexity does not imply R-convexity, consider the following case. Let $W = \{w_1, w_2\}$, and P and Q be probability functions with $P(w_1) = 0$ and $Q(w_1) = 1$, respectively. Define a BC C such that $C(P) = \{\{w_2\}, W\}, C(Q) = \{\{w_1\}, W\}$, and $C(\alpha P + (1-\alpha)Q) = \{\{w_1\}, \{w_2\}\}$ for all $\alpha \in (0,1)$. We can easily check that C satisfies C-convexity. Let R be the induced binarization relation from C. We can notice that $(P, W), (Q, W) \in R$, but $(0.5P + 0.5Q, W) \notin R$, which violates R-convexity. Moreover, the following theorem shows that R-convexity does not imply C-convexity.

Since

$$\min_{w\in B} P''_\alpha(w) \geq \alpha \min_{w\in B} P(w) + (1-\alpha)\min_{w\in B} P'(w)$$

(P''_α, Bel) satisfies HT^r.

(2) Let C be a BC induced by $\mathrm{HT}^{0.75}$. Consider the following probability distributions on $W = \{w_1, w_2, w_3\}$:

	w_1	w_2	w_3
P	0.58	0.3	0.12
Q	0.3	0.58	0.12
$0.5P + 0.5Q$	0.44	0.44	0.12

$C(P)(= C(Q)) = \{W\}$, but $C(0.5P + 0.5Q) = \{W, \{w_1, w_2\}\}$, which violates C-convexity.

(3) Let C be $DM(SE)$, and $P, P' \in \mathbb{P}(W)$. Suppose that $C(P) = C(P') = \mathcal{B} \subseteq \mathcal{P}(W)$. Let $\mathcal{B} = \{B_1, ..., B_k\}$. From the proof of Theorem 3.9 (5), it follows that for any $b, b' \in U^m$ it holds that for $\rhd \in \{=, <\}$

$$||p-b||^2 \rhd ||p-b'||^2 \text{ iff } 2\langle p, b'-b\rangle \rhd ||b'||^2 - ||b||^2.$$

Then we have the following equations:

$$\begin{aligned}
C^{-1}(\{b_1, ..., b_k\}) &= \{p \in \triangle^m |\ \operatorname*{argmin}_{b\in U^m} ||p-b||^2 = \{b_1, ..., b_k\}\}\\
&= \{p \in \triangle^m |\ ||p-b_1||^2 = ... = ||p-b_k||^2 < ||p-b||^2\ \forall b \in U^m \setminus \{b_1, ..., b_k\}\}\\
&= \{p \in \triangle^m |\ 2\langle p, b_i - b_1\rangle = ||b_i||^2 - ||b_1||^2 \text{ for } i = 1, ..., k.\\
&\quad \text{and } 2\langle p, b - b_1\rangle < ||b||^2 - ||b_1||^2 \text{ for } b \in U^m \setminus \{b_1, ..., b_k\}\}
\end{aligned}$$

Since the inner product is linear in the first argument, $C^{-1}(\{b_1, ..., b_k\})$ is closed under convex combination.

(4) Let $|W| = m$, and $p \in \triangle^m$. A belief state (P, Bel) satisfies the relation induced by DM(SE) iff $b \in \operatorname{argmin}_{b'\in\triangle^m} ||p-b'||^2$ iff $p \in V(b|U^m) \cap \triangle^m$ where $b \in U^m$ with $B = \bigcap Bel^{-1}(1)$. Since $V(b|U^m)$ and $\triangle^m$ are convex, the intersection of them is also convex.

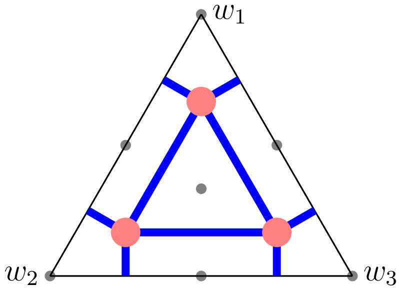
□

The above theorem shows that HT^r satisfies R-convexity while the BC induced by HT^r does not satisfy C-convexity. This may give the following lesson: when studying HT^r, how it is formalized[28] may be more important than it seems. Furthermore, it would be better to leave HT^r unaltered, at least with respect to the convexity norm. Now consider DM(SE). Geometrically speaking, the preimage-regions of DM(SE) partitions the probability simplex. Let us illustrate it by means of an example for the case of $W = \{w_1, w_2, w_3\}$ as follows

[28] Dietrich & List (2021) sees HT^r as a BC.

convexity	HT^r	DM(SE)
C-convexity	X	O
R-convexity	O	O

Table 3.5: convexity and generalized Binarization methods.



In the above simplex, the preimage-regions of DM(SE) partition the triangle into 19 pieces that consist of 7 white regions, 9 blue line segments, and 3 red points. The seven white regions are the preimage-regions of the singleton sets, the 9 blue line segments are the preimage-regions of the sets of 2-elements, and the 3 red points are the preimage-regions of the sets of 4-elements. Each preimage-region can be described using the Voronoi cells as follows

$$C^{-1}(\{b_1, ..., b_k\}) = [\bigcap_{i=1}^{k}(V(b_i|U^m) \cap \triangle^m)] \setminus [\bigcup_{b \in U^m \setminus \{b_1,...,b_k\}} V(b|U^m) \cap \triangle^m]$$

Since any convex polytope without some low dimensional face is also convex,
$C^{-1}(\{b_1, ..., b_k\})$ is convex. We will also prove a more general version of the above theorem in the next chapter, which provides an extension of the squared Euclidean distance in the above theorem to any Bregman divergence.
The induced relation by DM(SE) gives us a different picture: as shown in the above proof, for a given $b \in U^m$ the set of points in $\triangle^m$ that satisfy the relation induced by DM(SE) is exactly the same as the Voronoi cell of b restricted to $\triangle^m$. That is, the convexity defined by the relation induced by DM(SE) is no different than the convexity of Voronoi cells.

Every results about the convexity of HT^r and DM(SE) are collected in Table 3.5. It is noteworthy that DM(SE) satisfies both norms while HT^r satisfies one of them. In contrast to the BC induced by HT^r, DM(SE) satisfies C-convexity. Strictly speaking, we can extend the above table to include all belief binarization funtions. Since functions are also a correspondence/relation, BRs can be seen as a correspondence/relation. Since HT^r(S), CS^s, CCT^{g}and DM(SE)$^+$ do not satisfy preimage-convexity, they also do not satisfy C-convexity and R-convexity.

Convexity of ordinalization From many binarization methods, we can induce an ordering of possible worlds. In the following, we recall or introduce three world orderings formed by HT^r, CS^s, and CCT^g, respectively. Given $P \in \mathbb{P}(W)$,

(1) HT^r generates a unique total order $\precsim_P$ on W such that for all $v, w \in W$

$$\begin{aligned} v \prec_P w \text{ iff } & \text{ there is a } B \subseteq W \text{ such that } w \in B, v \notin B, \text{ and} \\ & P(w') > \frac{r}{1-r} P(\overline{B}) \text{ for all } w' \in B \\ v \sim_P w \text{ iff } & v \not\succ_P w \text{ and } v \not\prec_P w \\ v \precsim_P w \text{ iff } & v \prec_P w \text{ or } v \sim_P w; \end{aligned}$$

(2) CS^s generates a unique partial order $\prec_P$ on W such that for all $v, w \in W$

$$v \prec_P w \text{ iff } sP(v) < P(w);$$

(3) CCT^g generates a unique total order $\precsim_P$ on W such that for all $v, w \in W$

$$v \precsim_P w \text{ iff } P(v) \leq P(w)$$

Both CS^s and HT^r(S) generate its own non-trivial world ordering while the world ordering generated by CCT^g copies the total order of probability values, which makes CCT^g seem natural on the one hand, and vulnerable to even small changes in probabilities on the other hand. In our opinion, this 'natural' ordering is what caused CCT^g to fail to satisfy h-monotonicity[29]

Let us now introduce two different notions of the convexity of ordinalization. The first one, called (B,wo)-convexity, aims to weaken the preimage-convexity so that CS^s and CCT^g can satisfy it: (B,wo)-convexity is weaker than the preimage-convexity norm in the sense that the condition that a BR G generates the same world-ordering is added to the antecedent of the preimage-convexity norm. The second one is the straightforward formulation of the convexity of ordinalization, and is called wo-convexity. Note that (B,wo)-convexity and wo-convexity do not imply each other.

Definition 3.22. *(1) A belief binarization rule G satisfies the convexity of belief and world ordering ((B, wo)-convexity) if for all $P, P' \in \mathbb{P}(W)$ if P and P' induce the same world-ordering by G, and $G(P)(= G(P')) = B$, then $G(\alpha P + (1-\alpha)P') = B$ for all $\alpha \in [0,1]$.*

(2) A belief binarization method generating a world ordering satisfies the convexity of world ordering (wo-convexity) if for all $P, P' \in \mathbb{P}(W)$ if P and P' induce the same world-ordering , then $\alpha P + (1-\alpha)P'$ induces the same world-ordering as well for all $\alpha \in [0,1]$.

As mentioned above, CS^s and CCT^g satisfy (B,wo)-convexity. Then one may wonder if (B,wo)-convexity is so weak that it is generally satisfied. However, the next theorem shows that HT^r(S) does not satisfy it. It is also shown that the results regarding wo-convexity are similar.

Theorem 3.12.

[29]It makes CCT^g fail to satisfy the strong suspension principle too.

(1) $HT^r(S)$ satisfies neither (B,wo)-convexity nor wo-convexity, and HT^r does not satisfy wo-convexity.

(2) CS^s satisfies (B,wo)-convexity and wo-convexity.

(3) CCT^g satisfies (B,wo)-convexity and wo-convexity.

Proof. (1) Let G be a $\mathrm{HT}^{0.75}$(S). Consider the following probability distributions on $W = \{w_1, w_2, w_3\}$:

	w_1	w_2	w_3	$\precsim$	$(\cdot)$-stable$^{0.75}$ sets
P	0.58	0.3	0.12	$w_1 \sim w_2 \sim w_3$	$\{w_1, w_2, w_3\}$
Q	0.3	0.58	0.12	$w_1 \sim w_2 \sim w_3$	$\{w_1, w_2, w_3\}$
$0.5P + 0.5Q$	0.44	0.44	0.12	$w_1 \sim w_2 \succ w_3$	$\{w_1, w_2\}, \{w_1, w_2, w_3\}$

Since $G(P)(= G(Q)) = \{w_1, w_2, w_3\}$ with $\precsim_P = \precsim_Q$ but $G(0.5P+0.5Q) = \{w_1, w_2\}$, which violates (B,wo)-convexity. From $\prec$ column, we see that HT^r(S) does not satisfy wo-convexity.
If we apply $\mathrm{HT}^{0.75}$ to the above probability distributions, we also have the $\prec$ column, which implies that HT^r does not satisfy wo-convexity.

(2) Let G be a CS^s, and $P, P' \in \mathbb{P}(W)$. Suppose that $G(P) = G(P') = B$, and $\prec_P = \prec_{P'}$. Let $\alpha \in (0, 1)$, and $P''_\alpha = \alpha P + (1-\alpha)P'$. First, we will show wo-convexity, i.e., for all $v, w \in W$,

$$v \prec_P w \text{ iff } v \prec_{P''_\alpha} w$$

From our assumptions, we have

$$\begin{aligned} &v \prec_P w \text{ iff } \alpha s P(v) < \alpha P(w) \\ &v \prec'_P w \text{ iff } (1-\alpha)sP'(v) < (1-\alpha)P'(w) \end{aligned}$$

Suppose that $v \prec_P w$. Adding thw above two inequalities, we have $sP''_\alpha(v) < P''_\alpha(w)$. Now suppose that $v \not\prec_P w$. Similarly, we have $sP''_\alpha(v) \geq P''_\alpha(w)$, as desired. Since the same world ordering gives rise to the same maximal elements, we have also $G(P''_\alpha) = B$.

(3) Let G be a CCT^s, and $P, P' \in \mathbb{P}(W)$. Suppose that $G(P) = G(P') = B$, and $\precsim_P = \precsim_{P'}$. Let $\alpha \in (0, 1)$, and $P''_\alpha = \alpha P + (1-\alpha)P'$. With similar proof to the case of CS^s, it can be easily proved that P''_α preserves the given ordering on W. It is also easily checked that if $P(B), P'(B) \geq g$, then $P''_\alpha(B) \geq g$. It remains to show that $P''_\alpha(B) - \min_{w \in B} P''_\alpha(w) < g$. From our assumptions, we have

$$\alpha P(B) - \alpha \min_{w \in B} P(w) < \alpha g$$
$$(1-\alpha)P'(B) - (1-\alpha) \min_{w \in B} P'(w) < (1-\alpha)g$$

Adding two inequalities, we have

$$P''_\alpha(B) - (\alpha \min_{w \in B} P(w) + (1-\alpha) \min_{w \in B} P'(w)) < g$$

convexity	HT^r	$HT^r(S)$	CS^s	CCT^g
(B,wo)-convexity	—	X	O	O
wo-convexity	X	X	O	O

Table 3.6: Convexity of Ordinalization

Since $\precsim_P = \precsim_{P'} = \precsim_{P''_\alpha}$, it also holds that

$$\operatorname*{argmin}_{w \in B} P(w) = \operatorname*{argmin}_{w \in B} P'(w) = \operatorname*{argmin}_{w \in B} P''_\alpha(w)$$

, which implies our claim.

□

In order to formulate a kind of convexity that CS^s satisfies, why do we need the additional antecedent saying that two probability functions induce the same world ordering? Let us explain it by means of an example. Let G be CS^4. Consider the probability distributions on $W = \{1, 2, 3\}$ used in the proof of Theorem 3.12 (1). From those distributions, we can build the following partial orders (among two worlds with an edge, the lower world is less than ($\prec$) the upper world; otherwise no order):

$$\begin{array}{cc} w_1 & w_2 \\ | & \\ w_3 & \\ & \\ \multicolumn{2}{c}{\prec_P} \end{array} \quad \Bigg| \quad \begin{array}{cc} w_1 & w_2 \\ & | \\ & w_3 \\ & \\ \multicolumn{2}{c}{\prec_{P'}} \end{array} \quad \Bigg| \quad \begin{array}{ccc} w_1 & w_2 & w_3 \\ & & \\ & & \\ & & \\ \multicolumn{3}{c}{\prec_{(0.5P+0.5P')}} \end{array}$$

Note that G assigns to P and P' the same belief core $\{w_1, w_2\}$ although $\prec_P$ and $\prec_{P'}$ are different world orderings. So this case is ruled out by the antecedent and does not work as a counterexample.

Every result about convexity of ordinalization is indicated in Table 3.6. It is noteworthy that regarding the convexity of ordinalization, the performance of CS^s and CCT^g is superior to HT^r(S). How do we interpret this result? The results for CS^s and CCT^g might show that the binarization is only a part of ordinalization, and without respecting world-ordering, they are vulnerable to some mixing of probability functions. In contrast, for HT^r(S), even respecting world-ordering does not make it satisfy the convexity norm.

Now let us collect all the results proved in this section. It is worth pointing out that every binarization method has some convexity requirements favorable to it, which make it possible to distinguish every binarization method from the other methods.

convexity	HT^r	$HT^r(S)$	CS^s	CCT^g	DM(SE)$^+$	DM(SE)
preimage-convexity	—	X	X	X	X	—
preimage°-convexity	—	X	X	O	O	—
h-monotonicity	—	O	O	X	O	—
C-convexity	X	X	X	X	X	O
R-convexity	O	X	X	X	X	O
(B,wo)-convexity	—	X	O	O	—	—
wo-convexity	X	X	O	O	—	—

Table 3.7: Convexity and Binarization methods

3.6 Conclusion

In this chapter, we introduced, in a systematic way, classes of threshold-based heterogeneous belief aggregation rules. To characterize them, we defined various forms of monotonicity, and to uniquely characterize the direct rules, we devised the notion of conjunctiveness. We characterized all the classes we introduced and using the characterization we analyzed threshold-based rules. Moving on to collective belief binarization, we focused on local threshold rules that yield the property of being Lockean and coherent. We added their conjunction, being stable, and a weaker property of being r-likely. We reviewed the inclusion relation between the properties and provided their geometrical characterizations. With help of this, we identified which rational rules satisfy the properties:

- CS^s, CCT^g and DM(SE)$^+$ are not always stable.
- CS^s is not always r-likely but DM(SE)$^+$ is r-likely.
- DM(SE)$^+$ is interior-coherent.

Moreover, we proposed various kinds of convexity norms and examined which of them are satisfied by which binarization methods including not only functions but also correspondences, relations and ordinalizations:

- HT$^{\frac{1}{2}}$(S), HT^r(S), CS^s, CCT^g and DM(SE)$^+$ do not always satisfy preimage-convexity.
- CCT^g and DM(SE)$^+$ satisfy preimage°-convexity.
- CCT^g does not always satisfy h-monotonicity.
- DM(SE) satisfies C-convexity.
- HT^r and DM(SE) satisfy R-convexity.
- CS^s and CCT^g satisfy (B,wo)- and wo-convexity.

Notice that we investigated whether each rule *always* satisfies each property or not, and we suggested counterexamples when it fails. It will have to be left for future work to study the exact conditions under which a rule fails to satisfy the properties. Moreover, most of the analysis of collective belief binarization is not specialized for group decision context. We need to examine how we can reflect on the difference between individual and collective beliefs.

Chapter 4

Distance- and Utility-based Heterogeneous Belief Aggregation

In Chapter 1, we proposed a new research topic of aggregating multiple credences to a binary belief: heterogeneous belief aggregation. In Chapter 2, we collected some seemingly desirable conditions on heterogeneous belief aggregation, and proved that only trivial rules obey those conditions. To avoid the triviality results, we relax the independence condition and preserve rationality in Chapters 3 and 4. In Chapter 3, we studied threshold-based binarization rules and utilized them for heterogeneous belief aggregation. In this chapter, we shall propose and investigate distance- and utility-based rules: (1) belief binarization rules minimizing distance or maximizing expected utility, which will be combined with opinion pooling and (2) direct rules based on distance or utility.

$$\begin{tikzcd} \text{Individuals' probabilistic Beliefs} \arrow[r, "\text{Opinion}", "\text{Pooling}"'] \arrow[dr, "\text{(2) Direct rules (Dist.Min/Util.Max)}"'] & \text{Group's probabilistic Belief} \arrow[d, "\text{(1) Binarization (Dist.Min/Util.Max)}"] \\ & \text{Group's binary Belief} \end{tikzcd}$$

Figure 4.1

4.1 Introduction

4.1.1 Setting Out

This chapter proposes two novel belief binarization methods based on distance and utility and applies them to the heterogeneous belief aggregation problem. Most of the belief binarization literature has focused on threshold-based rules, as discussed in Chapter 3. The threshold-based rules associate beliefs/disbeliefs with high/low probabilities, respectively. We saw that global event-threshold rules — believing all and only the events with a probability above a given threshold — lead to the lottery paradox, which demonstrates the conflict between the independence norm and rationality. To preserve rationality, we may choose thresholds in a way that avoids the paradox (in local event-threshold rules) or set them for worlds' probabilities (in world-threshold rules). However, they are still concerned with each event's probability or world's probability.

Our arguments for new binarization rules begin by criticizing these event-wise or world-wise rules. There is no reason to determine a binary belief event-wise or world-wise, given a probability function on logically interconnected events. Rather, event-wise or world-wise procedures should be circumvented when the logical relation of the events in consideration is complex. We want belief binarization methods to take those logical relations into account. Therefore, we will not collect events/worlds with high probabilities to obtain a belief set/belief core, respectively, but directly choose a belief core — whose supersets constitute a belief set — among non-empty subsets of the set of all possible worlds. Now our problem is how a belief core can be determined given a probability function. To solve this problem, we return to the fundamental reasons why high probabilities are associated with beliefs and find a more direct and holistic way to obtain rational beliefs.

Probabilistic and binary beliefs aim at the truth. Thus, our purpose is to provide belief binarization rules that track the truth well — the closer to the truth, the better the rule. However, our perfect rational agents do not know what the truth is and just have a subjective probability, which aims at the truth as well. So to get close to the truth, the best ways for the agent with a probability to determine what to believe would be (i) the ways to *get close to the probability* or (ii) the ways to expectedly get close to the truth. Since probabilistic beliefs aim at the truth and are more fine-grained than binary beliefs, rational belief binarization methods seek to find a binary belief as close to the probabilistic belief as possible. In this sense, the former method tracks the truth by tracking the probabilistic belief and therefore is a somewhat implicit method to follow the aiming-at-truth norm. In contrast, we can also devise belief binarization methods that explicitly consider the truth-tracking norm of beliefs. Since our agent does not have access to the truth but only to a subjective probability which encodes all internally accumulated evidence of the truth, the agent, at best, can expect beliefs to be correct from the point of view of her own credal state. In this sense, the latter method can be seen as an explicit method to track the truth. The former leads to the *distance minimizing belief binarization rules (DM rules)* and the latter to the *expected (epistemic) utility maximizing belief binarization rules (EUM* *rules).*

The DM rules determine the belief core B that minimizes the distance from a given probability function P. How can we measure the distance from a probability function to a belief core? We will use divergences between probability functions. To this end, we identify a belief core B with the uniform distribution $U(B)$ on B so that the distance from $U(B)$ to B is zero. This assumption leads to the suspension principle that a belief binarization rule should map $U(B)$ to B. When B is the singleton set $\{w\}$ of a world w, $U(B)$ represents the probabilistically certain belief that w is the actual world. Thus, the resulting belief core should be $\{w\}$. When B is not a singleton, $U(B)$ represents the probabilistically certain belief that any world outside B is not the actual world. Therefore, we can exclude those worlds. Moreover, since $U(B)$ is uniformly distributed over worlds in B, it represents that the agent has no information about which world in B is the actual world. Thus, the agent should suspend judgment about whether the actual world is in any strict subset of B, and thereby B should be the belief core. This is a demanding condition when suspending judgments is allowed, i.e., when completeness of binary beliefs — every event should be believed or disbelieved — is not required. This is so weak as to be satisfied by most threshold-based rules in Chapter 3.

Now our problem is how we can measure the distance from a probability function to the uniform distribution on a non-empty set of possible worlds. We will examine Bregman divergences. They are the most commonly used distance measures between probability functions, and thus we can adopt abundant existing research results. In epistemic decision theory, Bregman divergences are used to provide justifications for epistemic norms such as probabilism since Bregman divergences are related to the expected inaccuracy of credences. We will utilize Bregman divergences more directly as a distance measure for the DM rules, and we call the rules *DM(Bregman).* The most relevant feature of Bregman divergences regarding our study is that minimizing the Bregman divergence from P amounts to the minimization of the expected Bregman divergence with respect to P. This means that getting close to a given probability leads to the same result as expecting to get close to the truth with respect to the probability. Roughly speaking, distance minimization has the same effect as expected distance minimization. Since the distance between a world and a probability function can be viewed as an epistemic utility (disvalue) of the probability function at the world, DM(Bregman) can be viewed as an EUM rule. What is more, the epistemic utility of a probability function at a world given by the Bregman divergence from the world to the probability function is a strictly proper score, which leads to the following binarization rule.

We call the EUM rules with a strictly proper score EUM(SP). Strictly proper scores are the most typical inaccuracy measures in epistemic decision theory and are the most commonly used utilities in many other areas, such as probabilistic forecasting and belief elicitation contexts. Strict propriety requires that the expected utility of a probability function with respect to P is uniquely maximized at P. So to obtain maximal expected utility, we have to choose P when P is given, which implies the suspension principle. This indicates that EUM(SP) rules can be regarded as DM rules. Moreover, it is well known that expected strictly proper scores generate Breg-

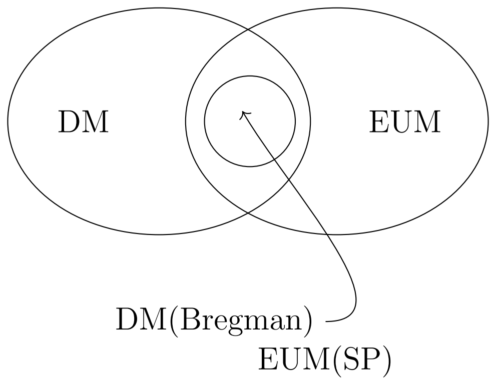


Figure 4.2: DM and EUM rules

man divergences. Thus, EUM(SP) rules are DM(Bregman) rules. This shows that EUM(SP) rules, or equivalently DM(Bregman) rules are the way to get close to the probability and to expectedly get close to the truth at the same time.

Although the EUM(SP) rules are the most representative EUM rules, it is worth studying the EUM rules in general. In epistemic decision theory, minimizing the expected inaccuracy of credences is the most common strategy to justify epistemic norms. We, in contrast, apply the minimization of the expected inaccuracy of probability functions, as decision rules, directly to belief binarization problems. So it is used to determine the belief core that is expected to get close to the truth. Here, one natural question arises: when can a belief core determined in this manner minimize the distance from the probability function as well? In other words, under what conditions on utilities can EUM rules be represented by DM rules? Inversely, under what conditions on distances can DM rules be represented by EUM rules? These questions will generalize the relationship between DM(Bregman) and EUM(SP). The relationships between the rules discussed so far are illustrated in Figure 4.2.

Now let us turn to the DM and EUM rules' geometrical features. Using the suspension principle, we can easily check whether a binarization rule is a DM rule or not, and see that many threshold-based rules introduced in Chapter 3 ($\mathrm{HT}^r(\mathrm{S})$, CS^s, CCT^g, HT^r) can be interpreted as DM rules. Can these rules be interpreted as EUM rules as well? This question is related to the convexity norm discussed in Section 3.5. We will show that every EUM rule satisfies C-convexity and R-convexity, which is derived from the linearity of the expectation operator. Since the above threshold-based rules do not satisfy C-convexity or R-convexity, we can conclude that they are not EUM-rationalizable, which means that they cannot be EUM rules.

Last but not least, the question is raised about how we can apply distance- and utility-based belief binarization rules to heterogeneous belief aggregation problems. As illustrated in Figure 4.1, the DM rules and EUM rules discussed above will be combined with an opinion pooling procedure such as linear pooling (LP) or geometric pooling (GP). We will also suggest direct heterogeneous belief aggregation rules based on some individual distances or utility functions. These two categories will be compared in terms of commutativity. Moreover, we will formulate some properties of distance- and utility-based heterogeneous belief aggregation, such as strong and weak unanimity and study how they relate to each other and convexity. In addition, we will examine whether certain specific combinations of LP or GP with some EUM or DM rules such as the DM rule with squared Euclidean distance and the one with Kullback-Leibler divergence satisfy them.

The most challenging part of this study will be the following part. First, we will represent probability functions not only in a probability simplex but also in De Finetti's coherent polytope. In the former approach, each world's probability (we will assume finite worlds) will be relevant, while the probabilities of some focused events will be relevant in the latter approach. Second, we will refine the conventional definition of Bregman divergence: to employ Bregman divergence for belief binarization problems, we want to allow infinite divergence on certain regions of boundaries of a probability simplex or De Finetti's coherent polytope. Third, we do not assume scoring rules to be additive. The above introduced rules and claims will be explicated and proved based on this technical setting. This will enrich the study of heterogeneous belief aggregation and contribute to other research areas. For example, we could apply this study to a justification of Bayesian conditionalization or linear pooling. Moreover, we could extend our study to a justification of probabilism with non-additive scores. What is more, we could discuss distance- and utility-based opinion pooling, analogously to DM rules and EUM rules.

The rest of this chapter is organized as follows: in the remainder of this section, we will review scoring rules and Bregman divergences, and recall the formal settings and definitions relevant to this chapter. In Section 4.2, we will explicate the DM rules and show that the suspension principle characterizes them. Moreover, we will suggest the refined definition of Bregman divergence and employ it to prove that DM(Bregman) is represented by the expected distance minimization. In Section 4.3 , we formulate the EUM rules with strictly proper scores and prove that EUM(SP) can be represented by DM(Bregman). Furthermore, we illustrate some examples of EUM rules and find the conditions for an EUM rule to be a DM rule and for a DM rule to be an EUM rule. In addition, we prove that every EUM rule is convex. In Section 4.4, which is based on joint work with Chisu Kim, we will combine distance- and utility-based binarization with opinion pooling and compare these combined rules with direct heterogeneous belief aggregation rules. We will also investigate properties of distance- and utility-based heterogeneous belief aggregation and examine whether certain specific rules satisfy them. Section 4.5 concludes this chapter.

4.1.2 Related Work

Scoring Rules Suppose that a meteorologist were to make the following probabilistic prediction about whether a typhoon would hit the country in a week: the typhoon will hit the country with a probability of 0.3. Furthermore, it turns out that the typhoon passes by the country. How can the quality of the meteorologist's prediction be assessed? Informally, she can be told that she was not wrong, but she was not wholly correct, and it could be said that she predicted close to the truth. Everybody would agree that the most crucial thing in evaluating her prediction is how accurate her prediction was, that is, how close it was to the truth. For this purpose, *scoring* *rules* have been devised to evaluate not only a probability estimate of some future event, but also any probability function over an algebra.

We now give a formal definition of scoring rules. Let W be a nonempty finite set of possible worlds and $\mathbb{P}(W)$ be a set of all probability distributions on W. A scoring rule is a function $S : W \times \mathbb{P}(W) \to \mathbb{R} \cup \{-\infty\}$ or $S : W \times \mathbb{P}(W) \to \mathbb{R} \cup \{\infty\}$ such that $S(w, P)$ is the score (reward/penalty) assigned to P when the realized outcome is w. It is natural to introduce the expected score to relate scoring rules with not only realized worlds but also probability functions over the set of all possible worlds: the expected score of Q with respect to P is defined by

$$\sum_{w \in W} P(w)S(w, Q)(=: \mathbb{E}_{w\sim P}[S(w, Q)])$$

One of the special cases of the expected score is called the Savage representation of a scoring rule S for P, which is defined by $\mathbb{E}_{w\sim P}[S(w, P)]$.[1]

Among many scoring rules, strictly proper scores have been extensively studied. The formal definition of strict proper scores is given as follows: a scoring rule is called strictly proper if the expected score of Q with respect to P is uniquely maximized when $Q = P$. Formally, S is strictly proper iff for all $P, Q \in \mathbb{P}(W)$

$$\mathbb{E}_{w\sim P}[S(w, Q)] \leq \mathbb{E}_{w\sim P}[S(w, P)]$$

and the equality holds only for $P = Q$.

The interpretation of the 'propriety' depends on how P is interpreted in $\mathbb{E}_{w\sim P}[S(w, Q)]$. In probabilistic forecasting (Brier 1950), P can be understood as an unknown objective chance, and Q as an expert's prediction. Then proper scoring rules can be used to track the objective chances. In belief elicitation (Schlag et al. (2015)), P refers to an expert's true belief, and Q to the expert's reported belief. Then proper scoring rules can be used to incentivize truthful reporting; they encourage the agents to reveal their true credal state in belief elicitation contexts.

In epistemic decision theory, proper scoring rules have been mainly utilized to justify probabilism (Predd et al.(2009)) or other epistemic norms (Pettigrew (2016)). To this end, the epistemic performance of probability functions needs to be compared with other credence functions, and therefore, the underlying set of Qs in $\mathbb{E}_{w\sim P}[S(w, Q)]$ — $\mathbb{P}(W)$ — is extended to the set of all credence functions. Then the strict-propriety norm requires that a probabilistic belief P expects itself to have a greater utility than the utility that it expects any other credence function to have. Although strictly proper scores have played a pivotal role in the justification of probabilism, whether the propriety is fully justified remains controversial. However, our discussion bypasses this difficulty since we pay attention to probability functions in the belief binarization context.

Let us introduce two mainly studied proper scoring rules: the Brier and the log score. The Brier score is defined by

$$S_{Brier}(w, P) = 2P(w) - \sum_{w \in W} P(w)^2$$

[1] See Savage (1971).

; the log score is defined by

$$S_{log}(w, P) = \log P(w)$$

The Savage representations of the Brier score and the log score are as follows:

$$\mathbb{E}_{w\sim P}[S_{Brier}(w, P)] = \sum_{w\in W} P(w)^2$$
$$\mathbb{E}_{w\sim P}[S_{log}(w, P)] = \sum_{w\in W} P(w)\log P(w)$$

Bregman Divergences How can we measure a distance between two probability functions? Since there are too many options[2], the most general definitions would be a good starting point. Arguably, the most general distance measures are divergences, defined as follows: a divergence is a function $D : \mathbb{P}(W) \times \mathbb{P}(W) \to [0, \infty]$ such that $D(P, Q) = 0$ iff $P = Q$. Note that it does not need to satisfy symmetry and triangle inequality. Let us mention two different divergences: the squared Euclidean distance is defined by

$$D_{SE}(P, Q) = \sum_{w\in W} (P(w) - Q(w))^2$$

; the Kullback-Leibler divergence is defined by

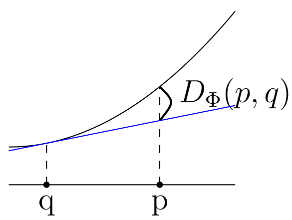
$$D_{KL1}(P, Q) = \sum_{w\in W} P(w) \log \frac{P(w)}{Q(w)}$$

if for all $w, P(w) \neq 0$ implies $Q(w) \neq 0$, o/w ∞.

Among many divergences, Bregman divergences are one of the extensively studied ones. They are general enough to include the squared Euclidean distance and Kullback-Leibler divergence, and have a close relationship with strictly proper scores, as we will see later. The definition of Bregman divergence is as follows: D is a Bregman divergence iff there is a differentiable, bounded and strictly convex function Φ such that

$$D_\Phi(p, q) = \Phi(p) - \Phi(q) - \nabla\Phi(q) \cdot (p - q)$$

for all p, q in a given convex subset $X \subseteq \mathbb{R}^m$. We call Φ the generating function of Bregman divergence D_Φ.



[2]For different kinds of distance measures between probability functions, see Deza & Deza (2016)

The above figure shows a convex function and the tangent line to the convex function at q. According to the Bregman divergence generated by the convex function, the distance from p to q is the difference between the function value at p and its tangent line approximation at q.

Let us see what the generating functions of $D_{SE}(P,Q)$ and $D_{KL1}(P,Q)$ are: for $D_{SE}(P,Q)$,

$$\Phi(P) = \sum_{w \in W} P(w)^2$$

; for $D_{KL1}(P,Q)$,

$$\Phi(P) = \sum_{w \in W} P(w) \log P(w)$$

Remember that the first one is the Savage representation of the Brier scoring rule, and the second one is the one of the log scoring rule. It is not a coincidence, since the Savage representation of strictly proper scores can be the generating function of a Bregman divergence.[3] There are different representation theorems between proper scoring rules and Bregman divergences (McCarthy (1956), Savage (1971), Gneiting & Raftery (2007)).

4.1.3 Distance- and Utility- based Rules

This chapter introduces distance- and utility-based heterogeneous belief aggregation. We begin by recalling the formal definitions of opinion pooling, belief binarization, and heterogeneous belief aggregation. Throughout this chapter, we assume that W is a finite non-empty set of possible worlds and denote by $\mathcal{P}(W)$ the powerset of W as in Chapter 3 so as to well-define the probability of a world. Let $N := \{1, ..., n\} (n \geq 2)$ be a set of individuals and $\vec{P}$ denote a profile $(P_1, ..., P_n)$ of individual probability functions P_i on $(W, \mathcal{P}(W))$. An *opinion pooling function* (OP) f is a function that takes as input the individual probabilities $\vec{P}$ and outputs the group's probability $f(P)$ on $(W, \mathcal{P}(W))$. And a *binarization rule* (BR) G is a function that takes a probability P on $(W, \mathcal{P}(W))$ and returns a binary belief $G(P)$, which is a function from $\mathcal{P}(W)$ to $\{0, 1\}$. Lastly, a *heterogeneous aggregator* (HA) F is a function that assigns to $\vec{P}$ in a given domain a binary belief $F(\vec{P})$.

As explained at the beginning of Chapter 3, we categorize heterogeneous belief aggregation into two groups: (1) collective belief binarization combined with an opinion pooling function and (2) direct rules that do not go through opinion pooling. In this chapter, we will address each group based on distance minimization and epistemic utility maximization. In Section 4.2 and 4.3, we restrict our attention to distance- and utility-based belief binarization ((1)) and in Section 4.4, we will address distance- and utility-based heterogeneous belief aggregation ((2) and the combination of (1) and opinion pooling).

[3] Note that $\Phi(P) = \mathbb{E}_{w \sim P}[S(w, P)]$ and $D_\Phi(P,Q) = \mathbb{E}_{w \sim P}[S(w,P)] - \mathbb{E}_{w \sim P}[S(w,Q)]$.

$$
\begin{tikzcd}
\text{Individuals' probabilistic Beliefs } \vec{P} \arrow[r, "\text{Opinion}", "\text{Pooling } f"'] \arrow[dr, "\text{(2) Direct Rules (Dist.Min/Util.Max)}"'] & \text{Group's probabilistic Belief } f(\vec{P})(=: P) \arrow[d, "\text{(1) Binarization } G\text{(Dist.Min/Util.Max)}"] \\
 & \text{Group's binary Belief } F(\vec{P})
\end{tikzcd}
$$

Figure 4.3

4.2 The Distance Minimizing Binarization Rules

In this and the next section, we focus only on belief binarization rules. As discussed in Section 3.4, most binarization rules suggested in belief binarization literature are based on some kinds of thresholds. We now propose novel *binarization rules based on distance and epistemic utility*. First of all, we suggest distance minimizing binarization rules (DM rules) and then explore the DM rules with Bregman divergence (DM(Bregman) rules). It is well-known that Bregman divergences and strictly proper scores are representable by each other. We will prove this fact in our framework and show that the DM(Bregman) rules can be represented by expected utility maximization with strictly proper scoring rules (EUM(SP) rules) and vice versa, which inspires us to define expected utility maximizing binarization rules (EUM rules) in Section 4.3.

We begin with the following assumptions that we will use throughout this chapter. First, binarization rules are presupposed to be *rational* as the threshold-based binarization rules introduced in Definition 3.17 in Section 3.4. In other words, every resulting binary belief is consistent (the belief set does not entail a contradiction) and deductively closed (the belief set contains all its logical consequences) so that it has *a non-empty belief core* whose supersets are exactly the believed events.[4] Therefore, we assume, here and subsequently, that a BR G is rational and regard $G(P)$ as a non-empty subset B of the set W of possible worlds.

Second, we will address distance minimization and expected utility maximization (or expected inaccuracy minimization) that take the following form

$$\operatorname*{argmin}_B g(P, B)$$

where g is an extended — i.e., ∞ is allowed as the output value — real-valued function of P and B, and $\operatorname{argmin}_B g(P, B)$ is the set of the non-empty subsets B of W that minimize the value $g(P, B)$. Accordingly, it is natural that we let binarization rules

[4]Recall Definition 3.5 in Section 3.2.2. Let $Bel : \mathcal{P}(W) \to \{0, 1\}$ be a binary belief on a finite space $(W, \mathcal{P}(W))$. Bel is rational iff Bel is consistent ($\bigcap Bel^{-1}(1) \neq \emptyset$) and deductively closed ($Bel^{-1}(1)$ contains W and it is closed under intersection and superset). See Definition 3.15 in Section 3.3 and Definition 3.3 in Section 3.2.1 as well.

$$
\begin{tikzcd}
\begin{array}{c}\text{Group's}\\ \text{probabilistic Belief}\\ P\end{array} \arrow[d, "\text{Rational Binarization } G"] \\
\begin{array}{c}\text{Group's}\\ \text{binary Belief}\\ G(P) =: B(\neq \emptyset, \subseteq W)\end{array}
\end{tikzcd}
$$

Figure 4.4

minimizing some real-valued function be a *correspondence*, which allows multiple outputs. We can later combine this with some tie-breaking rule to choose one belief core. Formally speaking, we regard a BR G on $(W, \mathcal{P}(W))$ as a correspondence from a given domain to $\mathcal{P}(W) \setminus \{\phi\}$, which is a function that takes as input a probability function P on $(W, \mathcal{P}(W))$ and outputs a set of non-empty subsets B of W. In the case where the output set is a singleton, we abbreviate $\operatorname{argmin}_B g(P, B) = \{B'\}$ by $\operatorname{argmin}_B g(P, B) = B'$.

Third, we will assume that a BR G has the universal domain $\mathbb{P}(W)$, which denotes the set of all probability functions on $(W, \mathcal{P}(W))$. Taken the above three assumptions together, we obtain the following: a binarization rule is a correspondence from $\mathbb{P}(W)$ to $\mathcal{P}(W) \setminus \{\emptyset\}$.

4.2.1 The DM Rules and the Suspension Principle

In this section, we define distance minimizing binarization rules and characterize them. For this purpose, we need to measure the distance between the input of a BR G — a probability function P on $(W, \mathcal{P}(W))$ — and a subset B of W. Our first main idea is to employ a *divergence* on a convex subset of $\mathbb{R}^m$ defined as follows:

Definition 4.1 (Divergence)**.** *Let X be a convex subset of $\mathbb{R}^m$. We call a function $d : X \times X \to [0, \infty]$ a divergence on X in $\mathbb{R}^m$ when $d(x, y) \geq 0$ where the equality holds iff $x = y$ for all $x, y \in X$.*

To this end, we need to represent probability functions P and subsets B in $\mathbb{R}^m$. For probability functions, we could deploy some typical methods to represent probability functions in $\mathbb{R}^m$. However, how can we represent a subset B in $\mathbb{R}^m$? Our second main idea is to identify it with the *uniform distribution* $U(B)$ on B — the probability distribution that assigns the same probability to each world in B and 0 to other worlds. It is plausible because when the input of a BR G is $U(B)$, B is the most natural belief binarization result, and thus we want to set the distance between $U(B)$ and B equal to 0.

Representations of Probabilities What remains in order to measure distance is to introduce how to represent probability functions in $\mathbb{R}^m$. There have been two approaches. In some contexts, such as judgment aggregation or epistemic utility theories, propositions are given at the outset, and only their probabilities are relevant. In other contexts, such as binarization theories or statistics, probabilities of worlds are used (under the assumption that there exist finite worlds). We include both approaches.

Recall that W is a finite non-empty set of possible worlds, and $\mathcal{P}(W)$ is the powerset-algebra, to whose elements probabilities are assigned. The first approach is to represent a probability function P by a point p in $\mathbb{R}^{|W|}$ where

$$p = (P(w))_{w \in W} \tag{4.1}$$

According to this representation method, we can represent an omniscient credence function V_w at $w \in W$ — assigning 1 to w — by a point v_w on $\{0,1\}^{|W|}$ where the w'-th coordinate is the following.

$$(v_w)_{w'} = V_w(w') = \mathbb{1}_{w=w'}$$

where $\mathbb{1}_{w=w'} = 1$ if $w = w'$, o/w $\mathbb{1}_{w=w'} = 0$. Then the set of the representation points of all probability distributions, denoted by $\triangle^W$, can be represented by the convex hull[5] of the representation points of all omniscient credence functions as follows:

$$\triangle^W = Conv(\{v_w \in \{0,1\}^{|W|} | w \in W\}) \subseteq \mathbb{R}^{|W|}$$

since $p = \sum_{w \in W} P(w) v_w$. Note that $\triangle^W$ is the same as the usual $|W|-1$-dimensional probability simplex, whose vertexes are the representation points of all omniscient credence functions. We say that p is the representation point of P in $\triangle^W$ if (4.1) holds.

We now move to the second approach. We introduce a non-empty subset $\mathcal{F}$ of $\mathcal{P}(W)$, and call it *the set of focused events*. Even though input probabilities and output binary beliefs are functions from $\mathcal{P}(W)$, there can be some situations where we are only interested in the focused events in $\mathcal{F}$. In this case, we represent probability functions $\mathbb{R}^{|\mathcal{F}|}$ and measure distances between them in this space. This approach can be used for a generalized agenda that is not required to be an algebra, such as agendas in judgment aggregation, epistemic decision theory and the basic setting in Chapter 2. Note that each input and output of a BR is assumed to be a function from the general agenda in Chapter 2, while they are functions from $\mathcal{P}(W)$ in this chapter.[6] However, this assumption of this chapter can be relaxed later.[7] [8]

[5]Let $R \subseteq \mathbb{R}^m$. The convex hull of R: $Conv(R) = \{z \in \mathbb{R}^m | \ z = \alpha x + (1-\alpha)y$ for some $x, y \in R, \alpha \in [0,1]\}$

[6]We need this condition to well define expected epistemic utility in EUM rules, where we need to calculate a probability of a singleton world unless we have IER (Invariant expectation under the same representation in Definition 4.9)

[7]The condition of IOR (Invariance under the same ouput-representation in Lemma 4.1), IIR (Invariance under the same input-representation in Definition 4.4), IER (Invariant expectation under the same input-representation in Definition 4.9) will play the role of relaxing this assumption.

[8]There can be another application of our setting: $\mathcal{F}$ can be interpreted as the set of the basic events, called the premises in Dietrich & List (2017b).

Now we are ready to formulate the second approach: we represent a probability function P by a point p in $\mathbb{R}^{|\mathcal{F}|}$ where

$$p = (P(A))_{A \in \mathcal{F}} \tag{4.2}$$

Thus, an omniscient credence function V_w at $w \in W$ is represented by a point v_w on $\{0, 1\}^{|\mathcal{F}|}$ where the A-th coordinate is the following.

$$(v_w)_A = V_w(A) = \mathbb{1}_{w \in A}$$

where $\mathbb{1}_{w \in A} = 1$ if $w \in A$, o/w $\mathbb{1}_{w \in A} = 0$. Moreover, we have

$$\triangle^{\mathcal{F}} = Conv(\{v_w \in \{0, 1\}^{|\mathcal{F}|} | w \in W\}) \subseteq \mathbb{R}^{|\mathcal{F}|}$$

since $p = \sum_{w \in W} P(w) v_w$ where $\triangle^{\mathcal{F}}$ denotes the set of the representation points of all probability distributions in this approach. Note that $\triangle^{\mathcal{F}}$ is a 0/1 polytope in $\mathbb{R}^{|\mathcal{F}|}$ (a polytope whose vertexes are on $\{0, 1\}^{|\mathcal{F}|}$). We say that p is the representation point of P in $\triangle^{\mathcal{F}}$ if (4.2) holds.

It is interesting to compare two approaches. In the case where $\mathcal{F} = \{\{w\} \in \mathcal{P}(W) | w \in W\}$, both ways are the same. If $|\mathcal{F}| < |W|$, one point in $\triangle^{\mathcal{F}}$ represents several distinct probability distributions, i.e., a probability distribution is not uniquely determined by a point: a point in $\triangle^{\mathcal{F}}$ represents a convex set of probabilities. It is because if P, P' have the same representation p, then for all $A \in \mathcal{F}$ we have $P(A) = P'(A) = \alpha P(A) + (1 - \alpha) P'(A)$ for all $\alpha \in [0, 1]$, which means that any linear combination of them has the same representation. Therefore, we can regard a point in $\mathcal{F}$ as a convex set of probabilities.

Note that many definitions, theorems, and statements in this chapter will be formulated using not only the representations in $\triangle^W$ but also the ones in $\triangle^{\mathcal{F}}$. To express this, we will use $\triangle^M$. Hence M is considered to be W throughout this chapter or to be $\mathcal{F}$ throughout this chapter.

Representations of Belief Cores Now let us turn to the representations of the uniform distribution $U(B)$ on a non-empty belief core $B \in \mathcal{P}(W) \setminus \{\emptyset\}$. Using the above approaches, they can also be represented in $\triangle^M$. We will denote by U^M the set of the representation points b of the uniform distributions $U(B)$ on all non-empty belief cores B.

With this in place, we will, hereafter, let a BR G be a correspondence from $\mathbb{P}(W)$ to U^M. In other words, $G(P)$ refers to a set of points b in U^M. Since we address binarization rules based on distance minimization or expected utility maximization, G takes the form of

$$\underset{b}{\operatorname{argmin}}\, g(P, b)$$

which is the set of the points b in U^M that minimize $g(P, b)$, where g is an extended real-valued function. Note that the uniform distribution on a belief core might not be uniquely determined by a point in $\triangle^{\mathcal{F}}$ as explained above, which can lead to the under-determination of a belief core. However, even if a point that represents several belief cores is selected as a group's belief, we could combine it with a tie-breaking rule later. Moreover, the following lemma shows that even though different belief cores give us different belief sets, their belief and disbelief in any event in $\mathcal{F}$ are the same if their representation points in $\triangle^{\mathcal{F}}$ are the same. It means that the focused events that should be believed are independent of the tie-breaking rule.

Lemma 4.1 (Invariance under the Same Ouput-representation (IOR))**.** *Let B and B' be non-empty subsets of W. If the uniform distributions on them are represented by the same point in $\triangle^{\mathcal{F}}$, i.e.,*

$$b = b'$$

then for all $A \in \mathcal{F}$,

$$B \subseteq A \textit{ iff } B' \subseteq A \textit{ and } B \subseteq \overline{A} \textit{ iff } B' \subseteq \overline{A}$$

where $\overline{A}$ is the complement of A.

Proof. Let $p_A := P(A)$ for all $A \in \mathcal{F}$ and for all $P \in \mathbb{P}(W)$. Let us prove the following first:

$$B \subseteq A \text{ iff } b_A = 1 \text{ and } B \subseteq \overline{A} \text{ iff } b_A = 0 \tag{4.3}$$

In $\triangle^{\mathcal{F}}$, $w \in A$ iff $(v_w)_A = 1$ for all $w \in W$ and for all $A \in \mathcal{F}$. Thus, $B \subseteq A$ means that for all $w' \in B$, $(v_{w'})_A = 1$, which is equivalent to $uni(B)_A(= \sum_{w' \in B} \frac{1}{|B|}(v_{w'})_A)$ $= 1$. Similarly, $w \in \overline{A}$ iff $(v_w)_A = 0$ for all $w \in W$ and for all $A \in \mathcal{F}$. Thus $B \subseteq \overline{A}$ means that for all $w' \in B$, $(v_{w'})_A = 0$, which is equivalent to $b_A = 0$.[9] Thus, by (4.3), since $b = b'$ means that for all $A \in \mathcal{F}$, $b_A = b'_A$, the claim follows. □

The statement (4.3) gives us a geometrical intuition about the binary belief corresponding to $b \in U^M$ assigned to the focused events A in $\mathcal{F}$. If $b_A = 1$, then A is believed; if $b_A = 0$, then $\overline{A}$ is believed; if $b_A \neq 0, 1$, then neither A nor $\overline{A}$ is believed. This explains why binary beliefs in the focused events in $\mathcal{F}$ are invariant under the same output-representation (IOR).

Lastly, before formalizing the definition of distance minimizing binarization rules, let us illustrate some examples of belief binarization problems.

Example 4.1 (Belief Binarization Problems)**.** *Consider the following four binarization problems where the set W of possible worlds or the pair of the set $\mathcal{F}$ of focused events and W are given:*

(1) $W = \{w_1, w_2, w_3\}$

(2) $\mathcal{F} = \{[a_1], [a_2]\}$, $W = \{w_1(\models a_1, a_2), w_2(\models a_1, \neg a_2), w_3(\models \neg a_1, a_2), w_4(\models \neg a_1, \neg a_2)\}$

(3) $\mathcal{F} = \{[a_1], [a_1 \wedge a_2]\}$, $W = \{w_1(\models a_1, a_2), w_2(\models a_1, \neg a_2), w_3(\models \neg a_1, a_2), w_4(\models \neg a_1, \neg a_2)\}$

[9] The following is another proof for the first part of (4.3) . If $B \subseteq A$, then $uni(B)_A = U(B)(A) = \sum_{w \in A} U(B)(w) = \sum_{w \in A \cap B} U(B)(w) = \sum_{w \in B} U(B)(w) = 1$. If $B \nsubseteq A$, then $\sum_{w \in A \cap B} U(B)(w) \neq \sum_{w \in B} U(B)(w)$. Thus $b_A \neq 1$.

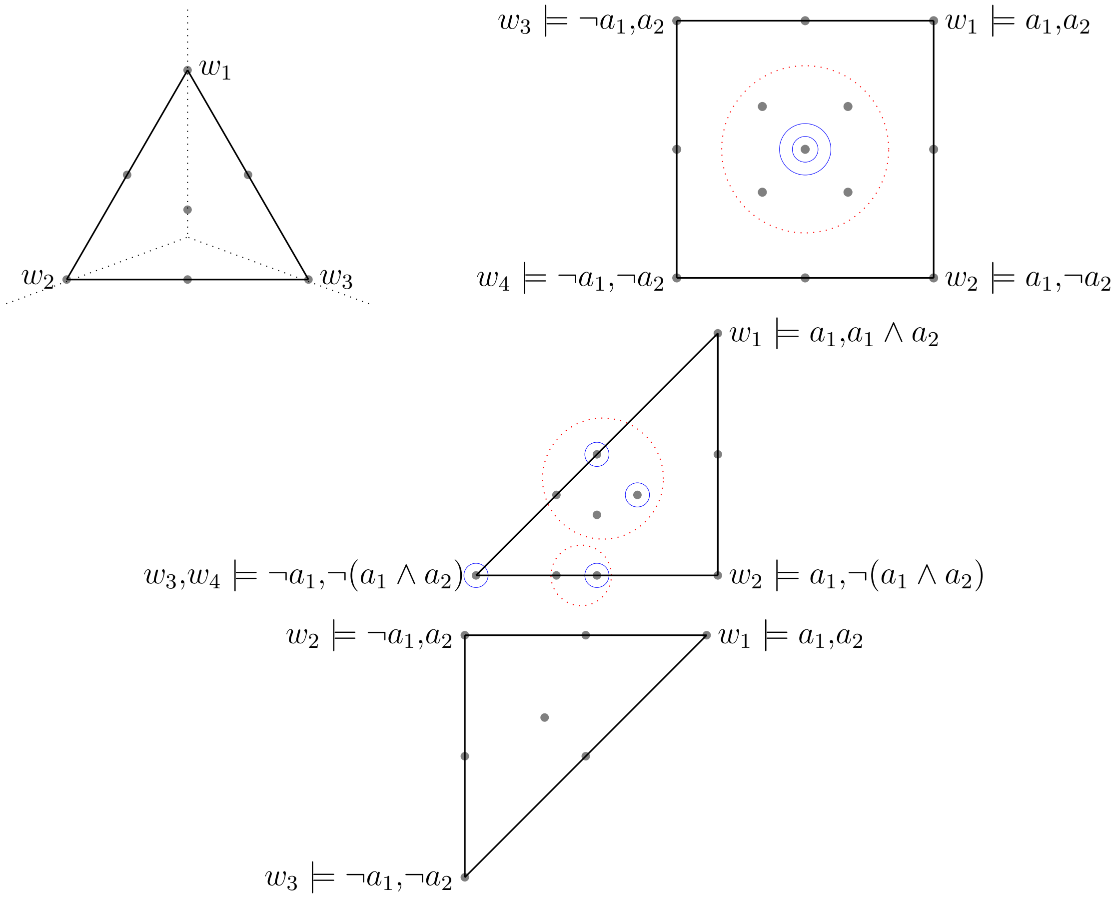


Figure 4.5: Belief Binarization Problems

(4) $\mathcal{F} = \{[a_1], [a_2]\}$, $LC = \{a_1 \to a_2\}$, $W = \{w_1(\models a_1, a_2), w_2(\models a_1, \neg a_2), w_3(\models \neg a_1, \neg a_2)\}$

where a_1 *and* a_2 *are atomic formulas in the standard propositional logic, and for any formula* ϕ, $[\phi]$ *is the set of the valuations under which* ϕ *holds. We write* $w \models \phi_1, \phi_2$ *when* ϕ_1 *and* ϕ_2 *hold under the valuation* w*. In (2), (3), and (4), each set of possible worlds is given as the set of all logically possible valuations. In the case with LC (short for logical constraints) like (4), logically possible valuations mean the valuations under which the formulas in LC hold.*

Figure 4.5 depicts (1)/(2)/(3)/(4) in $\triangle^W/\triangle^{\mathcal{F}}/\triangle^{\mathcal{F}}/\triangle^{\mathcal{F}}$, respectively. Each point represents the uniform distribution on a belief core. Some of them represent several uniform distributions, e.g., the central point in the upper right figure represents three different uniform distributions. The solid blue circles express that there are extra uniform distributions. Note that the points that are surrounded by a dotted red circle give us the same beliefs/disbeliefs of the focused events in $\mathcal{F}$ by the statement (4.3) in the proof of Lemma 4.1.

The DM rules and the Suspension Principle Now we deploy a divergence d on $\triangle^M$ and the representation methods discussed above in order to formulate distance minimizing binarization rules (DM rules) as follows.

Definition 4.2 (Distance Minimization rule (DM rule)). *A BR G is a distance minimization rule (DM rule) in $\triangle^M$ iff there is a divergence d on $\triangle^M$ such that*

$$G(P) = \operatorname*{argmin}_b d(p, b)$$

for all $P \in \mathbb{P}(W)$ and its representation point p in $\triangle^M$.

So a DM rule with a divergence d on $\triangle^M$ is a correspondence that takes as input any probability function P and outputs the points b in U^M that minimize the distance from p.

The following theorem states that the DM rules are characterized by a specific epistemic principle, what is called the suspension principle. The suspension principle requires that if there is no reason to prefer one epistemically possible world over another — here, an epistemically possible world means a world with a non-zero probability —, we should not believe any event that excludes some epistemically possible worlds, and we should believe any event that includes all the epistemically possible worlds. In other words, we should suspend every strict subset of the set of all epistemically possible and equally likely worlds and believe all this set's supersets. Here is the formal definition of the suspension principle.

Definition 4.3 (Suspension Principle). *A BR G satisfies the suspension principle iff for all $P \in \mathbb{P}(S)$ and its representation point $p \in \triangle^M$, and for all $b \in U^M$,*

$$\textit{if } p = b \textit{ then } G(P) = b.$$

This means that if P is a uniform distribution on some worlds, then binarization rules should result in the belief core that consists of those worlds.

For the characterization theorem, we will need the following condition, which gives some control over the cases where points in $\triangle^M$ can represent several probability distributions.

Definition 4.4 (Invariance under the same input-representation (IIR)). *A BR G satisfies the invariance under the same input-representation (IIR) in $\triangle^M$ iff if $P, P' \in \mathbb{P}(W)$ have the same representation in $\triangle^M$, i.e.,*

$$\textit{if } p = p' (\in \triangle^M), \textit{ then } G(P) = G(P')$$

where p and p' are the representation points of P and P' in $\triangle^M$, respectively.

If G satisfies IIR in $\triangle^{\mathcal{F}}$, then the binarization results depend only on the probabilities of the focused events in $\mathcal{F}$. Thus the representation point p plays the role of the input of G. This amounts to dealing with a generalized agenda and a probabilistic belief — a function extendable to a probability function on the algebra generated by the generalized agenda — as in Chapter 2. Let us compare this with the invariance under the output-representation (IOR) in Lemma 4.1. IIR means that two probability functions with the same representation input-point give us the same output-point, which is associated with several belief cores. On the other hand, IOR in Lemma 4.1 shows that any output-point gives us the same belief or disbelief about the focused events in $\mathcal{F}$.

The following theorem says that the suspension principle characterizes the DM rules if G satisfies IIR.

Theorem 4.2 (Characterization of DM rule)**.** *A BR G is a DM rule in $\triangle^M$ iff*

(i) G satisfies IIR in $\triangle^M$ and

(ii) G satisfies the suspension principle.

Proof. ($\rightarrow$) Since G has the form of $G(P) = \mathrm{argmin}_b\, d(p, b)$ for some divergence d, (i) and (ii) hold.
($\leftarrow$) Let d be as follows:

$$d(p, b) := \begin{cases} \min\limits_{b' \in G(P)} d_E(p, b') & \text{if } b \in G(P) \\ \\ \max\limits_{p', b'} d_E(p', b') & \text{otherwise} \end{cases}$$

where d_E is the Euclidean distance. It is well defined thanks to (i). Then, if $p = b$ then $G(P) = b$ by (ii) and thus $d(p, b) = d_E(p, b) = 0$. If $p \neq b$, then $d_E(p, b) > 0$ even if $b \in G(P)$, and thus $d(p, b) \neq 0$ whether $b \in G(P)$ or not. Thus, d is a divergence in $\triangle^M$. Furthermore, we have $G(P) = \mathrm{argmin}_b\, d(p, b)$ by the construction of d. □

We remark that DM rules in $\triangle^W$ always satisfy IIR. Thus, a BR G is a DM rule in $\triangle^W$ iff G satisfies the suspension principle. From this theorem, it can be easily checked that the threshold rules HT^r(S), CS^s and CCT^g, introduced in Definition 3.17 in Section 3.4, can also be seen as a DM rule in $\triangle^W$, while gCS^s cannot.

The most natural DM rule can be given by restricting the squared Euclidean distance D_{SE} to $\triangle^W$. We call this rule DM(SE) rule in $\triangle^W$. Let us give an example.

Example 4.2. *Let $W := \{1, 2, 3\}$. Note that $\triangle^W$ is a 2-dimensional simplex. DM(SE) rule in $\triangle^W$ is illustrated in Figure 4.6. The dotted lines divide the simplex seven regions. Each region excluding the dotted lines is the preimage-region $G^{-1}(b)(= \{p \in \triangle^W | G(P) = b\})$ of the point inside the region under G.*

In the following section, we will generalize the above example to DM rules with a Bregman divergence.

4.2.2 The DM Rules with Bregman Divergences

This section focuses on DM rules with a Bregman divergence, called DM(Bregman) rules. One major advantage of Bregman divergences is that they have a close relationship with proper scores. In the following, we will revise and apply the existing

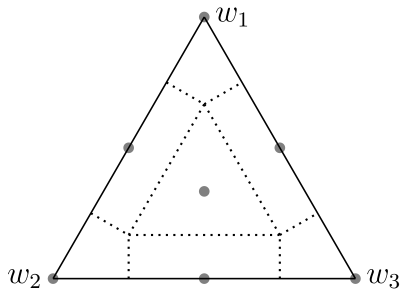


Figure 4.6: DM(SE)

research on the relation between Bregman divergences and proper scores to our belief binarization problem: firstly, we will refine the definition of Bregman divergence and secondly, we will show that DM(Bregman) rules can be viewed as a special kind of expected score maximization rules.

Refined Bregman Divergence First of all, we refine the definition of Bregman divergence in order to employ it for the belief binarization problems. In the typical definition of Bregman divergence in most of the literature, the domain of the second argument is an open set — e.g., the (relative) interior of a probability simplex $\triangle^W$ or $\mathbb{R}^m$ — or the value of Bregman divergence cannot be infinity. For our purposes, however, we need to define it in a closed set $\triangle^M$, because many uniform distributions on belief cores are located on the boundary of the set. Furthermore, we should allow infinity as a possible value of divergence, because we want to embrace some asymptotically divergent distance measures like the Kullback-Leibler divergence. For this reason, we need to extend the definition of Bregman divergence with *infinity* to the (relative) *boundary* of $\triangle^M$.

Considering that our definition includes infinity on the boundary, it can be compared with the definitions in Adamcik (2014a) and Pettigrew (2016). In Adamcik (2014a), Bregman divergence is defined in $\triangle^W$ and its value can be infinity.[10] Our definition is more general because it embraces not only Bregman divergences in $\triangle^W$ but also the ones in $\triangle^{\mathcal{F}}$. In Pettigrew (2016), Bregman divergence is defined in $\triangle^{\mathcal{F}}$ and its value can be infinity. However, the additivity of Bregman divergence is supposed and thus only one-dimensional Bregman divergences are addressed. We want to develop our definitions and theorems more generally without the assumption of additivity. In summary, our definitions and theorems work *not only in a simplex* $\triangle^W$ *but also in a general 0/1-polytope* $\triangle^{\mathcal{F}}$ dealing with not only additive but also *non-additive* divergence.

Let us explain the main features that we want to have in our definition in more detail. We want to extend the definition of finite Bregman divergence in the (relative) interior of $\triangle^M$ to the boundary, where we allow infinity. However, we do not want infinite distance all over the boundary, but we want to regulate where it should be finite and where infinite. For example, consider two points on the (relative) boundary

[10] In $\triangle^W$, our definition looks simpler, but we can easily prove that Adamcik's definition coincides with ours except that we have a continuity condition.

that lie in the (relative) interior of the same lower dimensional face. We want the divergence between them to be finite, just as the one between two points in the (relative) interior of the 0/1-polytope. Moreover, we want the divergence to be continuous in the region where it should be finite, just as the Bregman divergence is continuous in the (relative) interior. On top of that, we want to keep main properties — e.g., the relation between distance minimization and expected score maximization[11] — and well-known examples of Bregman divergence — e.g., the squared Euclidean distance and Kullback-Leibler divergence —, which enables one to use the existing results about Bregman divergences.

We begin by providing a way to denote lower dimensional faces not only in the simplex $\triangle^W$ but also in the 0/1-polytope $\triangle^{\mathcal{F}}$. In $\triangle^W$, we can denote by $Supp(Q)(:= \{w \in W | Q(w) \neq 0\})$ the set of the worlds corresponding to all vertexes of the lowest dimensional face on which $q(\in \triangle^W)$ lies. We need to extend this notion to indicate the faces of $\triangle^{\mathcal{F}}$, where a point in $\triangle^{\mathcal{F}}$ can represent multiple probabilities.

Definition 4.5 (Maximal Support, $\triangle_q$ and $\mathbb{F}_q$). *Let $p, q \in \triangle^M (\subseteq \mathbb{R}^m)$.*

(1) The maximal support of q is defined by

$$MSupp(q) := \bigcup_Q Supp(Q)$$

where Qs are the probability distributions represented by q, and $Supp(Q) := \{w \in W | Q(w) \neq 0\}$.

(2) The lowest dimensional face on which q lies is

$$\triangle_q := Conv(MSupp(q)) = \{x \in \triangle^M | MSupp(x) \subseteq MSupp(q)\}$$

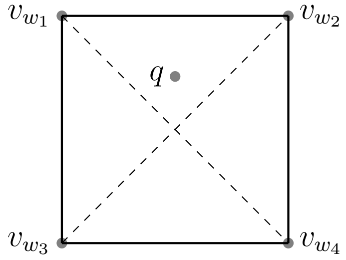
(3) The (disjoint) union of the relative interiors of all faces that p lies on is

$$\mathbb{F}_p := \{x \in \triangle^M | MSupp(p) \subseteq MSupp(x)\} = \{x \in \triangle^M | p \in \triangle_x\}$$

So $w \in MSupp(q)$ means that there exists a probability function Q represented by q such that $Q(w) \neq 0$ — i.e., a probability represented by q assigns to w a non-zero probability. Geometrically speaking, the maximal support of q is the set of the worlds corresponding to all vertexes of the lowest dimensional face on which q lies. In Figure 4.7, we give examples of maximal support, which will provide a geometric intuition behind it.

Using the notion of maximal support, we can designate the lowest dimensional face $\triangle_q$ on which q lies, which is the convex hull of the maximal support of q. We can easily check that $\triangle_q$ is a (sub-)0/1-polytope that is the set of points whose maximal support is a subset of q's maximal support. In contrast, $\mathbb{F}_p$ is the set of points whose maximal support is a superset of q's maximal support. In Figure 4.8, we give examples of the above-introduced notions. In the left figure, $\triangle_q$ is depicted, and in the right figure, $\mathbb{F}_p$ is displayed.

[11] It will be shown in Theorem 4.4 below.



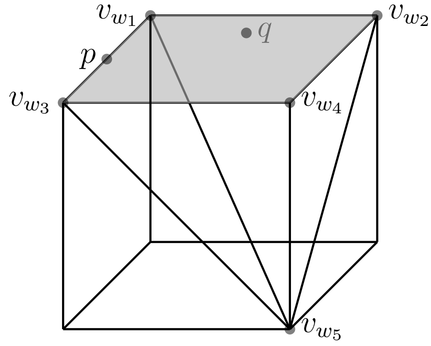
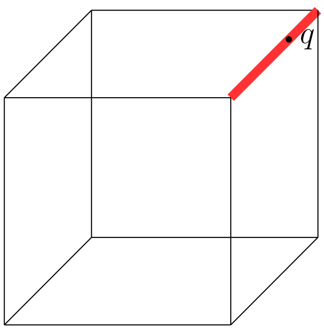


Figure 4.7: In the left figure, q is the representation of Q_1, Q_2 and Q_3 where $Supp(Q_1) = \{w_1, w_2, w_3\}$, $Supp(Q_2) = \{w_1, w_2, w_4\}$, and $Supp(Q_3) = \{w_1, w_2, w_3, w_4\}$. Thus, $MSupp(q) = \{w_1, w_2, w_3, w_4\}$. In the right figure, $MSupp(p) = \{w_1, w_3\} \subseteq MSupp(q) = \{w_1, w_2, w_3, w_4\}$



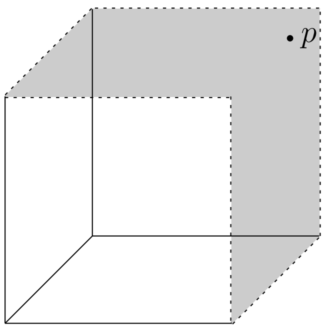


Figure 4.8: When $\triangle^M = [0, 1]^3$, $\triangle_q$ is the thick red line including the end points and $\mathbb{F}_p$ is the union of the relative interior($(0, 1)^3$) of $\triangle^M$ and the grey area excluding the dotted boundary.

In the following remark, we gather some useful results that can be easily checked. This will increase the understanding of the new notions. Recall that $ri(X)$ is the relative interior of X ($\subseteq \mathbb{R}^m$).[12]

Remark 4.1. *Let $q \in \triangle^M$ and $MSupp(q) = \{w_1, ..., w_k\}$.*

(1) There is a weighting vector $(\lambda_i)_i \in (0, 1]^k$ such that $q = \sum_{i=1}^{k} \lambda_i v_{w_i}$.

(2) If $q \in ri(\triangle^M)$, then $MSupp(q) = W$ and $\triangle_q = \triangle^M$.

(3) $ri(\triangle_q) = \{p \in \triangle^M | MSupp(p) = MSupp(q)\}$

(4) $q \in ri(\triangle_q)$.

(5) $MSupp(p) \subseteq MSupp(q)$ iff $p \in \triangle_q$ iff $q \in \mathbb{F}_p$.

(6) If $MSupp(p) \subseteq MSupp(q)$, then $\mathbb{F}_q \subseteq \mathbb{F}_p$.

Now we are ready to formulate our definition of Bregman divergence D. We modify its typical definition to the extent that $D(p, q)$ is *finite and continuous* except possibly that it is infinite if $q \notin \mathbb{F}_p$ — i.e., if there exists a world to which a probability represented by p assigns a non-zero value but every probability represented by q assigns zero. In other words, $D(p, q)$ is finite and continuous so far as

$$q \in \mathbb{F}_p$$

in other words, all worlds epistemically possible according to some probability represented by p are also epistemically possible according to some probability represented by q.

Definition 4.6 (Refined Bregman divergence)**.** *$D : \triangle^M \times \triangle^M \to [0, \infty]$ is a Bregman divergence in $\triangle^M$ iff there is a continuous, bounded and strictly convex function $\Phi : \triangle^M \to \mathbb{R}$ satisfying the following: for all $p, q \in \triangle^M$,*

(i) if $q \in \mathbb{F}_p$, then the directional derivative[13] $\nabla_{p-q}\Phi(q)$ exists, being finite and continuous in q, and

$$D(p, q) = \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)$$

(ii) otherwise,

$$D(p, q) = \lim_{\substack{x \to q \\ : x \in \mathbb{F}_p}} D(p, x)$$

which exists, infinity being allowed as limits.

---

[12] For the formal definition of the relative interior, see Appendix or the definitions at the end of Section 3.4.

[13] Let $f : dom(f)(\subseteq \mathbb{R}^n) \to \overline{\mathbb{R}}$, $x \in int(dom(f))$ and $v \in \mathbb{R}^n$. f has a *derivative in the direction* v at x if the following limit exists(finite):

$$\lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$

which is denoted by $\nabla_v f(x)$.

This definition is compared with the conventional ones of Bregman divergence on $ri(\triangle^M)$ as follows: The domain of D is not $ri(\triangle^M)$ but a closed set $\triangle^M$, and its codomain includes infinity. As usual, Bregman divergence is defined in terms of a convex function Φ, which is called a *Bregman divergence generator*. What distinguishes our definition from the conventional ones is part (i) and (ii). In the conventional definition, part (i) is applied in $ri(\triangle^M)$, which is the whole domain in the conventional ones, and thus (ii) is not needed. By contrast, we apply part (i) to the region where $q \in \mathbb{F}_p$, and we extend this continuously, infinity being allowed as limits, to the rest of the domain.

Note that we use the directional derivative instead of the gradient in the conventional definition. It is because we need to define divergence not only in the interior but also on the boundary, where gradients are not well defined. In the interior of the whole 0/1-polytope $\triangle^M$, it holds that $\nabla_{p-q}\Phi(q) = \nabla\Phi(q) \cdot (p-q)$ because Φ is differentiable from part (i) by the convexity of Φ — in $ri(\triangle^M)$, the differentiability follows from the existence of the finite directional derivative since Φ is convex (see Theorem 25.2 in Rockafellar (1970)). In the interior of any lower dimensional face $\triangle_q$, we could also say that $\Phi \restriction ri(\triangle_q)$ (the restriction of Φ to $ri(\triangle_q)$) is differentiable in the sense that it is differentiable in the lower dimensional space (note that the affine hull of $\triangle_q - q(:=\{x-q|x \in \triangle_q\}$) is a subspace of $\mathbb{R}^m$). In this sense, we could conclude that our definition of $D(p,q)$ coincides with the conventional one not only for $p, q \in ri(\triangle^M)$ but also for p, q in the relative interior $ri(\triangle_q)$ of any lower dimensional face.

Now let us suggest a natural alternative way to extend the conventional definition of Bregman divergence to the boundary and compare this with our definition. Instead of part (i) and (ii), we could have defined Bregman divergence when q is in $ri(\triangle^M)$ and extend it to the boundary as follows:

(*) For all $p, q \in \triangle^M$, if $q \in ri(\triangle^M)$, then $\nabla_{p-q}\Phi(q)$ exists, being finite[14], and $D(p,q) = \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)$, otherwise $D(p,q) = \lim_{\substack{x \to q \\ :x \in ri(\triangle^M)}} D(p,x)$ which exists, infinity being allowed as limits.

Even though this definition gives us one way to extend Bregman divergence, it only follows that the divergence is finite and continuous in $ri(\triangle^M)$. For example, it does not guarantee finite and continuous divergence between two points in the relative interior of the same face on the boundary. Figure 4.9 shows the cases that could arise if we defined Bregman divergence according to (*), which we want to avoid (We will see later that if we do not prevent this cases, then we cannot prove the relation between Bregman divergences and proper scores in Theorem 4.4).

For this reason, we want to extend the region where Bregman divergence should be finite and continuous — from $ri(\triangle^M)$ to $\mathbb{F}_p$. Notice that $q \in \mathbb{F}_p$ iff

$$MSupp(p) \subseteq MSupp(q)$$

This means that if any world is epistemically possible according to a probability

[14]This is equivalent to the continuous-differentiability of Φ in $ri(\triangle^M)$ since Φ is convex (see Theorem 25.2, Theorem 25.5 and Corollary 25.5.1 in Rockafellar (1970)). Therefore, $\nabla_{p-q}\Phi(q)$ is continuous in q.

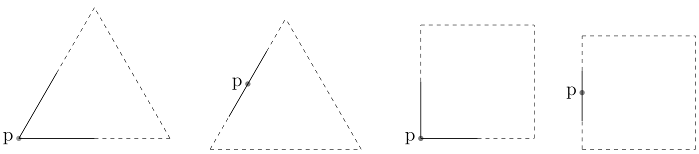


Figure 4.9: The dashed lines including the end points in each polytope represent where $D(p,\cdot)$ might be infinite and the complement of them is where $D(p,\cdot)$ is finite and continuous according to (*). We want to avoid these cases.

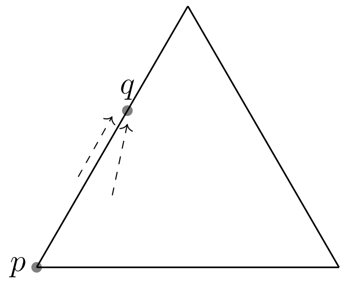


Figure 4.10: Each arrow represents x approaching q. The left arrow represents the case (a) and both arrows represent the case (b). Both limits should exist and be the same as $\nabla_{p-q}\Phi(q)$.

represented by p, then the world is epistemically possible according to a probability represented by q as well. Loosely speaking, we want Bregman divergences to be finite as far as q does not exclude any world that p does not exclude.

Now let us comment on the continuity condition in (i) and the limit condition of (ii). The continuity condition in (i) includes not only the case (a) where x approaches q only through $ri(\triangle_q)$, but also the more general case (b) where x approaches q through all faces that include q, and thus through $\mathbb{F}_p$. (See Figure 4.10.) The case (a) means that $\nabla_{p-\cdot}\Phi \upharpoonright ri(\triangle_q)(\cdot)$ is continuous at q. This is always the case because $\Phi \upharpoonright ri(\triangle_q)$ is a convex function and thus continuously differentiable.[15] The case (b) means that $\nabla_{p-\cdot}\Phi \upharpoonright \mathbb{F}_p(\cdot)$ and thus $\nabla_{p-\cdot}\Phi(\cdot)$ is continuous at $q \in \mathbb{F}_p$, which ensures the continuity between different faces as well.[16] In the limit condition of (ii), we impose the restriction on the sequence x approaching q: $x \in \mathbb{F}_p$, in order to guarantee the existence of directional derivative $\nabla_{p-x}\Phi(x)$. Since we do not have $\lim_{x\to q} D(p,x) = D(p,q)$, even though $D(p,q) = \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} D(p,q)$, we might not have the continuity of the whole space. (See Figure 4.11.) It is natural because there can be regions where the divergence takes the value infinity. From part (i) and part (ii) together, we have a kind of continuity in $\triangle^M$ to the extent that

$$D(p,q) = \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} D(p,x)$$

[15]See Theorem 25.5 and Corollary 25.5.1 in Rockafellar (1970).

[16]The continuity condition in (i) can be formulate as $\nabla_{p-q}\Phi(q) = \lim_{x\to q}\nabla_{p-x}\Phi(x)$. One may wonder whether we need to impose a restriction on x such that $\nabla_{p-q}\Phi(q) = \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}}\nabla_{p-x}\Phi(x)$ so that $\nabla_{p-x}\Phi(x)$ exists. The restriction is redundant because for x to approach to q, x in a neighborhood of q should lie on the faces $\mathbb{F}_q$ that q belongs to. It means that $MSupp(q) \subseteq MSupp(x)$. Since $MSupp(p) \subseteq MSupp(q)$, we have $MSupp(p) \subseteq MSupp(x)$. Moreover, note that $\lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}}\nabla_{p-x}\Phi(x) = \lim_{\substack{x\to q\\ :x\in ri(\triangle^M)}}\nabla_{p-x}\Phi(x)$ since the continuity in the relative interior of any lower dimensional face follows from Φ being convex as mentioned above.

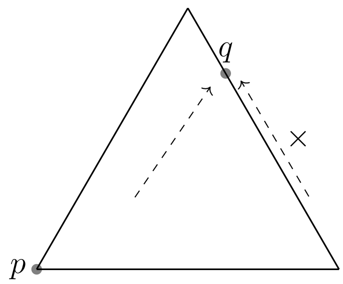


Figure 4.11: Each arrow represents x approaching q. The left arrow satisfies $MSupp(p) \subseteq MSupp(x)$ and $D(p, q)$ is defined as its limit, which might be different from the limit of the right arrow that, indeed, might not exist.

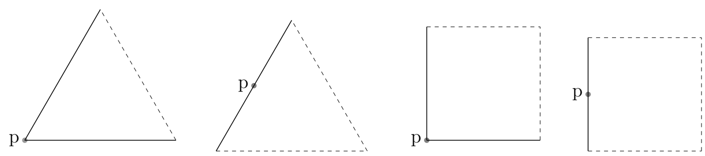


Figure 4.12: The dashed lines including the end points in each polytope represent where $D(p, \cdot)$ might be infinite. The complement of them is $\mathbb{F}_p$ where $D(p, \cdot)$ is finite and continuous.

for all $p, q \in \triangle^M$. Note that $x \in \mathbb{F}_p$ is a necessary and sufficient condition that $D(p, x)$ can be defined without using limits.

To summarize, we accomplished all of our aims presented at the beginning of this section:

- $D(p, \cdot)$ is finite and continuous not only in $ri(\triangle^M)$ but also in $\mathbb{F}_p$ (see Figure 4.12),

while keeping main properties and well-known examples of Bregman divergence:

- the definition is the same as the conventional ones not only in $ri(\triangle^M)$ but also in $ri(\triangle_q)$ for any q: if $p \in \triangle_q$, then $D(p, \cdot) \restriction ri(\triangle_q)$ is a Bregman divergence generated by $\Phi \restriction ri(\triangle_q)$ and

- our definition includes the squared Euclidean distance and Kullback-Leibler divergence etc. (all in Table 1 in Banerjee et al. (2005)) and the divergences defined with continuous proper scores in Theorem 4.9, to be proven soon.

Representation of DM(Bregman) by Expected Distance Minimization Now we employ our refined Bregman divergence for DM rules. A DM(Bregman) rule is the DM rule with a Bregman divergence D in $\triangle^M$, which has the following form: for all $P \in \mathbb{P}(W)$

$$G(P) = \operatorname*{argmin}_b D(p, b)$$

where p ($\in \triangle^M$) is the representation point of P.

Now we will prove that the DM(Bregman) rule can be viewed as a minimization of the expected divergence between the two points: the point corresponding to a world and the point corresponding to a belief core. To show this, we will need the following lemma, which says that the directional derivative of a convex function is linear, if it exists and finite.

Lemma 4.3. *Let $\Phi : \triangle^M \to \mathbb{R}$ be a convex function. The following statements are equivalent:*

(i) For all $p, q \in \triangle^M$ such that $q \in \mathbb{F}_p$, the directional derivative $\nabla_{p-q}\Phi(q)$ exists and is finite.

(i′) For all $q \in \triangle^M$ there exists $f \in \mathbb{R}^m$ such that for all $p \in \triangle_q$

$$\nabla_{p-q}\Phi(q) = f \cdot (p - q)$$

Proof. (i′)→(i): straightforward. (i)→(i′): Suppose that f is a subgradient[17] of ϕ at q. As $q \in ri(\triangle_q)$, the existence of a subgradient is guaranteed by the convexity of ϕ and Theorem 23.4 in Rockafellar (1970). For $h(> 0)$, from the definition of subgradient we have that $\Phi(q + h(p - q)) \geq \Phi(q) + f \cdot h(p - q)$ for all $p \in \triangle_q$, that is,

$$\frac{\Phi(q + h(p - q)) - \Phi(q)}{h} \geq f \cdot (p - q)$$

For $h(> 0)$ that is small enough that $q - h(p-q) \in \triangle_q$ (Such h exists since $q \in ri(\triangle_q)$.), we have that $\Phi(q - h(p - q)) \geq \Phi(q) - f \cdot h(p - q)$ for all $p \in \triangle_q$, that is,

$$\frac{\Phi(q) - \Phi(q - h(p - q))}{h} \leq f \cdot (p - q)$$

Since Φ has the finite directional derivative at q in the direction of $p - q$, with $h \to 0$, we get $\nabla_{p-q}\Phi(q) = f \cdot (p - q)$.[18] □

Now, our aim is to prove that the DM(Bregman) rule can be represented by a decision rule that minimizes expected distance from the point v_w ($\in \triangle^M$) corresponding to a world $w \in W$, which is the representation point of the omniscient credence function V_w at $w \in W$ (Recall that when $\triangle^M = \triangle^W$, $(v_w)_{w'} = V_w(w') = \mathbb{1}_{w=w'}$, and when $\triangle^M = \triangle^{\mathcal{F}}$, $(v_w)_A = V_w(A) = \mathbb{1}_{w \in A}$). Although our proof runs along similar lines as the proofs of Theorem 1 in Banerjee et al. (2005) and Theorem 2 in Adamcik (2014a), subtle adjustments are necessary for our belief binarization problem. First, our refined Bregman divergence is defined not only in $\triangle^W$ but also in $\triangle^{\mathcal{F}}$ as well. Second, the refined Bregman divergence is defined neither in $\mathbb{R}^m$ nor in an open convex subset, but in a closed convex subset $\triangle^M$. Third, we allow infinity as a value of divergence.

[17] Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be a convex function. $f \in \mathbb{R}^n$ is called a subgradient of Φ at $y \in \mathbb{R}^n$ if

$$f \cdot (x - y) \leq \Phi(x) - \Phi(y) \text{ for all } x \in \mathbb{R}^n$$

[18] Notice that for any subgradients f and f' of Φ at q, we have $f \cdot (p - q) = f' \cdot (p - q)$ for all $p \in \triangle_q$. It shows the uniqueness of the subgradient of $\Phi \restriction \triangle_q$ at q, which indicates differentiability at q.

Now let us prove the following theorem, which shows that we can represent the DM(Bregman) rule by a decision rule minimizing expected divergence, which we will call EUM(SP). In the next section, we will explain the reason why we call it that. Throughout, we denote the expectation of g with respect to a probability distribution $P \in \mathbb{P}(W)$ by $\mathbb{E}_{w\sim P}[g(w)]$ where $g : W \to \mathbb{R} \cup \{\infty\}$ or $g : W \to \mathbb{R} \cup \{-\infty\}$. Note that $\mathbb{E}_{w\sim P}[g(w)] = \sum_{w\in W} P(w)g(w)$.

Theorem 4.4 (Representation of DM(Bregman) by EUM(SP))**.** *Let D be a Bregman divergence in $\triangle^M$. Then*

$$\operatorname*{argmin}_b D(p,b) = \operatorname*{argmin}_b \mathbb{E}_{w\sim P}[D(v_w,b)]$$

for all $p \in \triangle^M$ and any probability $P \in \mathbb{P}(W)$ represented by p.

Proof. We will prove the stronger claim that

$$D(p,q) = \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)] \tag{4.4}$$

for all $p, q \in \triangle^M$ and p's any extension $P \in \mathbb{P}(W)$. First, assume that $q \in \mathbb{F}_p$. Then not only the left-hand side but also the right-hand side are finite because

$$\mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)] = \sum_{w\in Supp(P)} P(w)D(v_w,q) - \sum_{w\in Supp(P)} P(w)D(v_w,p)$$

and for all $w \in Supp(P)(\subseteq MSupp(p) \subseteq MSupp(q))$, $D(v_w,q)$ and $D(v_w,p)$ are finite. Let $D(p,q) = \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)$. Then

$$\begin{aligned}
&\mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)]\\
&= \mathbb{E}_{w\sim P}[\Phi(v_w) - \Phi(q) - \nabla_{v_w-q}\Phi(q)] - \mathbb{E}_{w\sim P}[\Phi(v_w) - \Phi(p) - \nabla_{v_w-p}\Phi(p)]\\
&= \Phi(p) - \Phi(q) - \mathbb{E}_{w\sim P}[\nabla_{v_w-q}\Phi(q)] + \mathbb{E}_{w\sim P}[\nabla_{v_w-p}\Phi(p)]\\
&= \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)
\end{aligned}$$

The last equality follows from the fact that

$$\mathbb{E}_{w\sim P}[\nabla_{v_w-q}\Phi(q)] = \mathbb{E}_{w\sim P}[f\cdot(v_w-q)] = f\cdot\mathbb{E}_{w\sim P}[(v_w-q)] = f\cdot(\mathbb{E}_{w\sim P}[v_w]-q)$$

where f is a (sub)gradient at q and $\mathbb{E}_{w\sim P}[\vec{g}(w)] = (\mathbb{E}_{w\sim P}[g_i(w)])_{i\le m}$ for $\vec{g} : W \to \mathbb{R}^m$. Since $w \in MSupp(q)$ for all $w \in Supp(P)$, this holds by the linearity of expectation and the linearity of a directional derivative of a convex function. (See Lemma 4.3, or Theorem 25.2 in Rockafellar (1970)) Our claim holds from

$$\mathbb{E}_{w\sim P}[v_w] = \sum_{w\in W} P(w)v_w = p$$

Next, assume that $q \notin \mathbb{F}_p$. Then

$$\begin{aligned}
D(p,q) &= \lim_{x \to q : x \in \mathbb{F}_p} D(p,x) \\
&= \lim_{x \to q : x \in \mathbb{F}_p} (\mathbb{E}_{w \sim P}[D(v_w, x)] - \mathbb{E}_{w \sim P}[D(v_w, p)]) \\
&= \lim_{x \to q : x \in \mathbb{F}_p} \sum_w P(w) D(v_w, x) - \mathbb{E}_{w \sim P}[D(v_w, p)] \\
&= \sum_{w \in Supp(P)} P(w) \lim_{\substack{x \to q \\ : MSupp(p) \subseteq MSupp(x)}} D(v_w, x) - \mathbb{E}_{w \sim P}[D(v_w, p)] \\
&= \mathbb{E}_{w \sim P}[D(v_w, q)] - \mathbb{E}_{w \sim P}[D(v_w, p)]
\end{aligned}$$

The forth equality holds since $P(w), D(v_w, x) \geq 0$. Let us explain why the last equality holds: For any $w \in Supp(P)$, thus for any w such that $\{w\} = MSupp(v_w) \subseteq MSupp(p)$, $\lim_{\substack{x \to q \\ : MSupp(p) \subseteq MSupp(x)}} D(v_w, x)$ exists because $\{x \in \triangle^M | MSupp(p) \subseteq MSupp(x)\} \subseteq \{x \in \triangle^M | MSupp(v_w) \subseteq MSupp(x)\}$.[19] Moreover,

$$\lim_{\substack{x \to q \\ : MSupp(p) \subseteq MSupp(x)}} D(v_w, x) = D(v_w, q)$$

since

$$\lim_{\substack{x \to q \\ : MSupp(v_w) \subseteq MSupp(x)}} D(v_w, x) = D(v_w, q)$$

□

It is worth noting that we could not prove this theorem for our belief binarization problem, if the definition of Bregman divergence guaranteed finiteness only in $ri(\triangle^M)$. Suppose that we try to prove the first part of (4.4) on the assumption that $q \in ri(\triangle^M)$, and we use a definition of Bregman divergence that might yield the cases in Figure 4.9, e.g., the definition with (*)(p.134) instead of (i) and (ii) in Definition 4.6. The left-hand side $D(p, q)$ is finite for $q \in ri(\triangle^M)$. However, $D(v_w, p)$ might not be finite even though $w \in Supp(P)$, and thus the right-hand side of (4.4) might not be finite.

The distance between v_w and b can be thought of as a utility in epistemic decision theory — an epistemic disvalue. Thus, DM(Bregman) rule can be seen as a decision rule *maximizing expected utility*. This inspires us to define a new rule that applies epistemic decision theory directly to binarization problems.

[19] For any $r \in \triangle^M$, in the case where $MSupp(r) \not\subseteq MSupp(q)$, $\lim_{x \to q} D(r, x)$ might not exist if we do not impose the condition about the sequence that $MSupp(r) \subseteq MSupp(x)$. With this condition we can define $D(r, x)$ without using limit.

4.3 The Expected Utility Maximizing Binarization Rules

4.3.1 The EUM Rules with Strictly Proper Scores

In this paper, we assume that our perfectly rational agent's qualitative beliefs ought to be consistent and deductively closed. This assumption enables us to think of an epistemic utility of a qualitative belief state at a world as a function of belief cores and worlds. Furthermore, since belief cores, in our set-up, can be represented by uniform distributions on them, we can make use of the scoring rules that have been used to measure epistemic performances of credal states. With this in place, we will propose a rule to optimize expected utility with respect to the probability function given by the agent's credal state, in order to select a belief core that is well matched with the agent's credal state. Now let us give a formal definition of *expected utility maximization rule (EUM rule).*

Definition 4.7 (Expected Utility Maximization rule (EUM rule))**.** *A BR G is an expected utility maximization rule in $\triangle^M$ iff there is a utility function $u : W \times U^M \to \mathbb{R} \cup \{-\infty\}$ satisfying*

$$G(P) = \underset{b}{\operatorname{argmax}}\, \mathbb{E}_{w \sim P}[u(w, b)]$$

for all $P \in \mathbb{P}(W)$.

In this section, we restrict our focus to already well-developed epistemic utility functions, namely proper scores (in the next section, we will consider more general versions of utility functions). It will be useful in order to investigate the relation between EUM rules and DM(Bregman) rules introduced in the previous section. Put differently, we shall consider the case where

$$u(w, b) := -I(w, b)$$

for some continuous strictly proper (SP) score $I : W \times \triangle^M \to [0, \infty]$, to be defined below.

Continuous Strictly Proper Score Now we define a continuous strictly proper score in our setting where probability distributions are represented in $\triangle^M$ and infinity is allowed. We include infinity as a value of scores on some region of boundary, where we do not demand continuity. Thus, when we talk about continuity of a score including infinity, we also need to regulate the region where it can be infinite.

Definition 4.8 (Continuous Strictly Proper (SP) Score)**.** *Let I be a function $I : W \times \triangle^M \to [0, \infty]$.*

(1) I is continuous iff for all $w \in W$ and $q \in \triangle^M$,

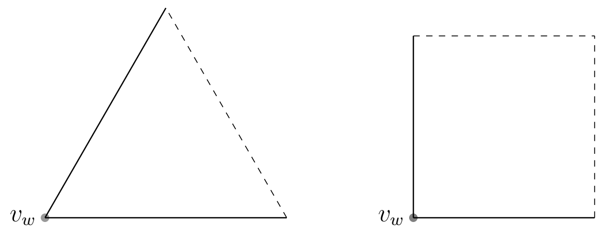


Figure 4.13: The dashed lines including the end points in each polytope represent where a continuous score $I(w, \cdot)$ might be infinite. The complement of them is $\mathbb{F}_{v_w}$ where $I(w, \cdot)$ is finite and continuous.

(i) if $q \in \mathbb{F}_{v_w}$ *then* $I(w, q)$ *is finite and continuous in* q[20]

(ii) otherwise, I *is extended to a continuous function that might be infinite, meaning that*

$$I(w, q) = \lim_{\substack{x \to q \\ : x \in \mathbb{F}_{v_w}}} I(w, x)$$

which exists, infinity being allowed as limits.

(2) I *is called a strictly proper score (SP) iff*

$$\operatorname*{argmin}_{q \in \triangle^M} \mathbb{E}_{w \sim P}[I(w, q)] = p$$

for all $P \in \mathbb{P}(W)$ *and its representation point* p*.*

According to our definition above, continuous scores have the following features (see Figure 4.13):

- $I(w, \cdot)$ is finite not only in $ri(\triangle^M)$ but also in $\mathbb{F}_{v_w}$
- $I(w, \cdot)$ is continuous in $\mathbb{F}_{v_w}$

Notice that $q \in \mathbb{F}_{v_w}$, i.e., $w \in MSupp(q)$ says that a probability represented by q assigns to w a positive value. It means that w is epistemically possible from the point of view of a probability represented by q. In this case, we demand that $I(w, q)$ should not receive infinite score.

However, the first feature cannot be regarded as distinct from other strict proper scores; the following lemma shows that if I is strictly proper, then we can derive that $I(w, \cdot)$ is finite in $\mathbb{F}_{v_w}$.[21]

Lemma 4.5. *Let* $I : W \times \triangle^M \to [0, \infty]$ *be a strictly proper score. Then,* $I(w, \cdot)$ *is finite in* $\mathbb{F}_{v_w}$

[20] This can be formulated as $I(w, q) = \lim_{\substack{x \to q \\ : x \in \mathbb{F}_{v_w}}} I(w, x) = \lim_{x \to q} I(w, x)$, since in the neighborhood around q, we have $x \in \mathbb{F}_q$, i.e., $MSupp(q) \subseteq MSupp(x)$, from which follows that $w \in MSupp(x)$. Note that $\lim_{\substack{x \to q \\ : x \in \mathbb{F}_{v_w}}} I(w, x) = \lim_{\substack{x \to q \\ : w \in \triangle^M}} I(w, x)$, since it is finite in $\mathbb{F}_{v_w}$. We can prove this in a similar way to continuous extension theorems.

[21] This is the same as the notion of regular score in Definition 2 in Gneiting & Raftery (2007).

Proof. Since $\mathbb{E}_{w\sim P}[I(w,p)] < \mathbb{E}_{w\sim P}[I(w,q)]$ for all $q \neq p$, $\mathbb{E}_{w\sim P}[I(w,p)]$ should be finite for all $P \in \mathbb{P}(W)$ and $p \in \triangle^M$ such that p is a representation point P.[22] Thus, for all $w \in Supp(P)$, $I(w,p)$ is finite for all P represented by p. Therefore, for all $w \in MSupp(P)$, $I(w,p)$ is finite. □

Invariant Expectation under the Same Input-representation Notice that the expectation value in the EUM rules depends not only on the point p in $\triangle^M$ but also on the probability distribution P, in contrast to the divergence in DM rule. Thus, to see the connection between the EUM rules and DM rules, we need the following requirement, which is relevant in $\triangle^{\mathcal{F}}$ where a point p might represent several probability distributions.

Definition 4.9 (Invariant Expectation under the Same Input-representation (IER))**.** *A function $I : W \times \triangle^M \to [0,\infty]$ has an invariant expectation under the same input-representation (IER) iff if $P, P' \in \mathbb{P}(W)$ has the same representation in $\triangle^M$, i.e., $p = p'(\in \triangle^M)$, then*

$$\mathbb{E}_{w\sim P}[I(w,q)] = \mathbb{E}_{w\sim P'}[I(w,q)]$$

for all $q \in \triangle^M$.

IER has a close relationship with IIR in Definition 4.4. If a BR G is an EUM rule with I that satisfies IER, then G is invariant under the same representation (IIR). Although IER may seem very strong condition, it is actually a mild restriction because a large class of scores obey IER. Every scoring function defined in $\triangle^W$ satisfies IER. We can generalize this in $\triangle^{\mathcal{F}}$ as follows.

Lemma 4.6. *Let $I : W \times \triangle^{\mathcal{F}} \to [0,\infty]$ be a scoring function. I satisfies IER if I is a partition-wise score, i.e., there is a partition $A_1, ..., A_k$ of W such that (i) $A_1, ..., A_k \in \mathcal{F}$ and (ii) for all $w, w' \in A_i$, $I(w,q) = I(w',q)$ for all $i \leq k$ and $q \in \triangle^{\mathcal{F}}$.*

Proof. Since $\mathbb{E}_{w\sim P}[I(w,q)] = \sum_{w\in W} P(w)I(w,q) = \sum_{i\leq k} P(A_i)I(w_i,q)$ where $w_i \in A_i$, we have $\mathbb{E}_{w\sim P}[I(w,q)] = \sum_{i\leq m} p_{A_i} I(w_i,q) = \mathbb{E}_{w\sim P'}[I(w,q)]$. □

There are other ways to satisfy IER. Any additive scores defined in $\triangle^{\mathcal{F}}$ also enjoy IER as the following lemma shows.

Lemma 4.7. *If $I : W \times \triangle^{\mathcal{F}} \to [0,\infty]$ is additive, i.e., for all $w \in W$ and $p \in \triangle^{\mathcal{F}}$*

$$I(w,p) = \sum_{A\in\mathcal{F}} I_A((v_w)_A, p_A)$$

where for all $A \in \mathcal{F}$, $I_A : \{0,1\} \times [0,1] \to [0,\infty]$, then it has IER.

[22] In $\triangle^W$, this condition is the same as I being regular in Definition 1 in Gneiting & Raftery (2007), which is assumed in order to prove the relation between Bregman divergences and proper scores.

Proof. First compute the following.

$$
\begin{aligned}
\mathbb{E}_{w\sim P}[I(w,q)] &= \sum_{w\in W} P(w)I(w,q) = \sum_{w\in W} P(w)\sum_{A\in\mathcal{F}} I_A((v_w)_A, q_A)\\
&= \sum_{A\in\mathcal{F}}\left(\sum_{w\in A} P(w)I_A(1,q_A) + \sum_{w\notin A} P(w)I_A(0,q_A)\right)\\
&= \sum_{A\in\mathcal{F}}\Big(p_A I_A(1,q_A) + (1-p_A)I_A(0,q_A)\Big) \qquad (4.5)
\end{aligned}
$$

Thus, if $p = p'$, then $\mathbb{E}_{w\sim P}[I(w,q)] = \mathbb{E}_{w\sim P'}[I(w,q)]$. $\square$

From (4.5) in the above proof, we can easily check the following as well.

Remark 4.2. *If an additive score I is event-wise strictly proper (E-SP), i.e.,*

$$\operatorname*{argmin}_{q_A\in[0,1]}\Big(p_A I_A(1,q_A) + (1-p_A)I_A(0,q_A)\Big) = p_A$$

for all $A \in \mathcal{F}$, then I is strictly proper.

Representation of EUM(SP) by DM(Bregman) Lastly, we need the following lemma and its corollary for the proof of the main theorem of this section. We are dealing with continuity not only in $\triangle^W$ but also in $\triangle^{\mathcal{F}}$. The following lemma enables one to find a continuous function assigning $P \in \mathbb{P}(W)$ to $p \in \triangle^{\mathcal{F}}$. It is interesting in itself, because the lemma shows a 'continuous' relation between a probability simplex and De Finetti's coherent polytope.

Lemma 4.8 (Continuous Selection)**.** *There is a continuous function taking any $p \in \triangle^{\mathcal{F}}$ and giving $P \in \mathbb{P}(W)$ where P has the representation p in $\triangle^{\mathcal{F}}$.*

Proof. Let $|\mathcal{F}| = m$ and $|W| = n$. Firstly, observe that we have a linear function $L : \triangle^W \to \triangle^{\mathcal{F}}$ that can be represented by a $m \times n$-binary-matrix as following:

$$
\begin{pmatrix}
(v_{w_1})_1 & (v_{w_2})_1 & \cdots & (v_{w_n})_1\\
(v_{w_1})_2 & (v_{w_2})_2 & \cdots & (v_{w_n})_2\\
\vdots & \vdots & \ddots & \vdots\\
(v_{w_1})_m & (v_{w_2})_m & \cdots & (v_{w_n})_m
\end{pmatrix}
\begin{pmatrix}
P(w_1)\\ P(w_2)\\ \vdots\\ P(w_n)
\end{pmatrix}
=
\begin{pmatrix}
p_1\\ p_2\\ \vdots\\ p_m
\end{pmatrix}
$$

Our aim is to find a continuous inverse of L. (See Figure 4.14)

First of all, we can triangulate $\triangle^{\mathcal{F}}$ in such a way that $\triangle^{\mathcal{F}}$ is a union of simplexes $\triangle_1$, ... , $\triangle_k$ and $\bigcup_{i=1}^{k} V(\triangle_i) = V(\triangle^{\mathcal{F}})$ where $V(\triangle)$ denotes the set of all vertexes of a polytope $\triangle$. It is always possible, because $\triangle^{\mathcal{F}}$ is a polytope. For a vertex v of $\triangle^M$ choose one of the omniscient probability measures V_w such that $L(V_w) = v$.

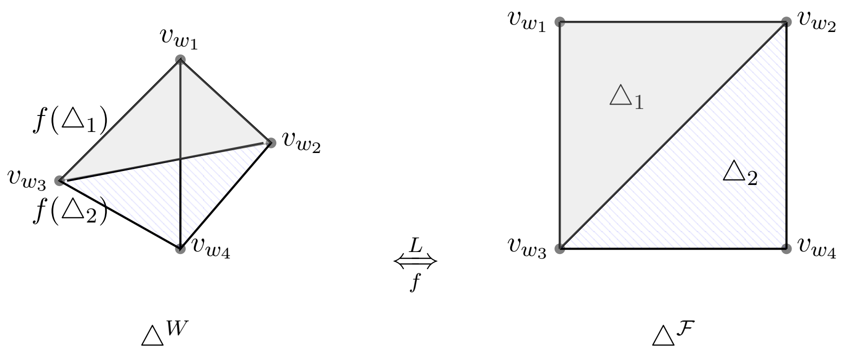


Figure 4.14

It is always possible because for any vertex v it holds that $\{w|v = v_w\} \neq \emptyset$. For any $p \in \triangle_i$ we can uniquely represent p by $p = \sum_{v \in V(\triangle_i)} \lambda_v v$ for some $(\lambda_v)_v$ such that $\sum_{v \in V(\triangle_i)} \lambda_v = 1$ and $\lambda_v \geq 0$. Then we can define a function f_i from $\triangle_i$ to $\triangle^W$ such that $f_i(p) = \sum_{v \in V(\triangle_i)} \lambda_v L^{-1}(v)$ where $L^{-1}(v)$ denotes the selected omniscient probability measure. Observe that f_i is continuous. Note that for any $q \in \triangle_i \cap \triangle_j$, $f_i(q) = f_j(q)$. Now, we can construct a unique map $f : \cup_{i=1}^k \triangle_i \to \triangle^W$ by gluing $f_1, f_2, ..., f_k$ such that $f|\triangle_i = f_i$ for all $1 \leq i \leq k$.

Let us check that f is continuous. Suppose that A is a closed subset of $\triangle^W$. Then $f^{-1}(A) = \bigcup_{i=1}^k f_i^{-1}(A)$. Since every $f_i^{-1}(A)$ is closed because of the continuity of f_i, and a finite union of closed sets is closed, it follows that $f^{-1}(A)$ is also closed.

It remains to show that $L(f(p)) = p$ for all $p \in \triangle^{\mathcal{F}}$. First, pick a $\triangle_i$ such that $p \in \triangle_i$. Let us compute.

$$\begin{aligned} L(f(p)) &= L(f(\sum_{v \in V(\triangle_i)} \lambda_v v)) = L(\sum_{v \in V(\triangle_i)} \lambda_v L^{-1}(v)) \\ &= \sum_{v \in V(\triangle_i)} \lambda_v L(L^{-1}(v)) = \sum_{v \in V(\triangle_i)} \lambda_v v = p \end{aligned}$$

where in the third equality, we used the linearity of L. □

Corollary 4.8.1. *Let $I : W \times \triangle^M \to [0, \infty]$ be a continuous score and $q \in \triangle^M$. Let f be a continuous function in the above lemma when $\triangle^M = \triangle^{\mathcal{F}}$. When $\triangle^M = \triangle^W$, let f be an identity function. Then $\mathbb{E}_{f(q)}[I(w, q)]$ is continuous at q.*

Proof. $\mathbb{E}_{f(q)}[I(w, q)] = \sum_w f(q)(w) I(w, q)$ and if $w \in Supp(f(q))$, then $w \in MSupp(q)$. Thus, for all $w \in Supp(f(q))$, $I(w, \cdot)$ is finite and continuous at q. Moreover, f is continuous and the projection on the w-th barycentric coordinate is continuous. Therefore, our claim holds. □

Now, let us show how EUM(SP) is related to DM(Bregman). This theorem shows that EUM(SP) can be represented by DM(Bregman) when I is a continuous strictly proper score with IER.

Theorem 4.9 (Representation of EUM(SP) by DM(Bregman))**.** *Let $I : W \times \triangle^M \to [0,\infty]$ be a continuous strictly proper (SP) score with IER. Then there is a Bregman divergence D in $\triangle^M$ such that*

$$\operatorname*{argmin}_b \mathbb{E}_{w\sim P}[I(w,b)] = \operatorname*{argmin}_b D(p,b)$$

for all $P \in \mathbb{P}(W)$ and its representation $p \in \triangle^M$.

Proof. For $p, q \in \triangle^M$, let us define a divergence as follows:

$$D(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)] \tag{4.6}$$

Since I satisfies IER, it is well defined and since I is SP, it is a divergence. We will show that it is a Bregman divergence with

$$\Phi(p) = -\mathbb{E}_{w\sim P}[I(w,p)]$$

Note that Φ is well-defined since I satisfies IER.

Claim (1): Φ is continuous, bounded and strictly convex on $\triangle^M$.
By IER and Corollary 4.8.1, Φ is continuous. Since $I(w,p)$ is finite for all $w \in MSupp(p)$, it is finite for all $w \in Supp(P)(\subseteq MSupp(p))$, and thus $\mathbb{E}_{w\sim P}[I(w,p)]$ is finite.
Now let us prove the strict convexity. For $p, q(\in \triangle^M)$ and $\lambda(\in [0,1])$ we have

$$\begin{aligned}
-\Phi(\lambda p + (1-\lambda)q) &= \mathbb{E}_{w\sim \lambda P+(1-\lambda)Q}[I(w, \lambda p + (1-\lambda)q)] \\
&= \lambda \mathbb{E}_{w\sim P}[I(w, \lambda p + (1-\lambda)q)] + (1-\lambda)\mathbb{E}_{w\sim Q}[I(w, \lambda p + (1-\lambda)q)] \\
&> \lambda \mathbb{E}_{w\sim P}[I(w,p)] + (1-\lambda)\mathbb{E}_{w\sim Q}[I(w,q)] \\
&= -\lambda\Phi(p) - (1-\lambda)\Phi(q)
\end{aligned}$$

The first equality holds by IER because $\lambda P + (1-\lambda)Q$ is one of probability distributions that are represented in $\triangle^M$ by $\lambda p + (1-\lambda)q$. The second equality comes from the linearity of expectation and the inequality in the third line holds because I is SP. Thus $\Phi(p)$ is strictly convex.

Claim (2): If $MSupp(p) \subseteq MSupp(q)$, then the directional derivative $\nabla_{p-q}\Phi(q)$ exists and is finite. Moreover $\nabla_{p-\cdot}\Phi(\cdot)$ is continuous at q.
We will show that

$$\nabla_{p-q}\Phi(q) = -\mathbb{E}_{w\sim P}[I(w,q)] + \mathbb{E}_{w\sim Q}[I(w,q)]$$

and it is finite and continuous in q. Assume that $q \in \mathbb{F}_p$. Note that there is enough small h such that $q + h(p-q), q - h(p-q) \in ri(\triangle_q)$ because $q \in ri(\triangle_q)$. For $h > 0$, let us compute.

$$
\begin{aligned}
&\frac{1}{h}[\Phi(q+h(p-q))-\Phi(q)]\\
=&-\frac{1}{h}[\sum_w (Q+h(P-Q))(w)I(w,q+h(p-q))-\sum_w Q(w)I(w,q)]\\
=&-\frac{1}{h}\sum_w (Q(w)+h(P(w)-Q(w)))[I(w,q+h(p-q))-I(w,q)]\\
&-\sum_w P(w)I(w,q)+\sum_w Q(w)I(w,q)
\end{aligned}
$$

The first equality holds by IER. The last equality holds since every term is finite because

$$Supp(P), Supp(Q) \subseteq MSupp(q+h(p-q)) = MSupp(q)$$

Since I is strictly proper, we know that

$$\sum_w (Q(w)+h(P(w)-Q(w)))[I(w,q+h(p-q)-I(w,q)] \leq 0$$

It implies that

$$\frac{1}{h}[\Phi(q+h(p-q))-\Phi(q)] \geq -\sum_w P(w)I(w,q)+\sum_w Q(w)I(w,q)$$

Similarly, for $h>0$, we have

$$\frac{1}{h}[\Phi(q)-\Phi(q-h(p-q))] \leq -\sum_w P(w)I(w,q)+\sum_w Q(w)I(w,q)$$

Notice that $\sum_w P(w)I(w,q)$ is continuous in q because for w such that $P(w)\neq 0$, we have $w \in Supp(P) \subseteq MSupp(q)$ and thus, $I(w,q)$ is continuous in q. By IER and Corollary 4.8.1, we also have that $\sum_w Q(w)I(w,q)$ is continuous in q. It implies that $\nabla_{p-q}\Phi(q)$ exists as desired. Note that

$$-\sum_w P(w)I(w,q)+\sum_w Q(w)I(w,q) = -\mathbb{E}_{w\sim P}[I(w,q)]+\mathbb{E}_{w\sim Q}[I(w,q)]$$

and it is finite and continuous in q as we indicated in the above.

Claim (3): For all $p,q \in \triangle^M$, $D(p,q)=\Phi(p)-\Phi(q)-\nabla_{p-q}\Phi(q)$ if $q\in\mathbb{F}_p$, otherwise $D(p,q)=\lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} D(p,x)$ which exists (infinity being allowed as limits).

First assume that $q\in\mathbb{F}_p$. By (2) we have

$$
\begin{aligned}
D(p,q) &= \mathbb{E}_{w\sim P}[I(w,q)]-\mathbb{E}_{w\sim P}[I(w,p)]\\
&= -\mathbb{E}_{w\sim P}[I(w,p)]+\mathbb{E}_{w\sim Q}[I(w,q)]+\mathbb{E}_{w\sim P}[I(w,q)]-\mathbb{E}_{w\sim Q}[I(w,q)]\\
&= \Phi(p)-\Phi(q)-\nabla_{p-q}\Phi(q)
\end{aligned}
$$

Otherwise, we need to show that

$$\lim_{x \to q : x \in \mathbb{F}_p} D(p,x) = \mathbb{E}_{w \sim P}[I(w,q)] - \mathbb{E}_{w \sim P}[I(w,p)]$$

Let us compute.

$$\begin{aligned}
\lim_{x \to q : x \in \mathbb{F}_p} D(p,x) &= \lim_{x \to q : x \in \mathbb{F}_p} (\mathbb{E}_{w \sim P}[I(w,x)] - \mathbb{E}_{w \sim P}[I(w,p)]) \\
&= \lim_{x \to q : x \in \mathbb{F}_p} \sum_w P(w) I(w,x) - \mathbb{E}_{w \sim P}[I(w,p)] \\
&= \sum_{w \in Supp(P)} P(w) \lim_{\substack{x \to q \\ : MSupp(p) \subseteq MSupp(x)}} I(w,x) - \mathbb{E}_{w \sim P}[I(w,p)] \\
&= \mathbb{E}_{w \sim P}[I(w,q)] - \mathbb{E}_{w \sim P}[I(w,p)]
\end{aligned}$$

The third equality holds because $P(w), I(w,x) \geq 0$. The forth equality holds since for $w \in Supp(P)$,

$$\lim_{\substack{x \to q \\ : MSupp(p) \subseteq MSupp(x)}} I(w,x) = I(w,q)$$

since

$$\lim_{\substack{x \to q \\ : MSupp(v_w) \subseteq MSupp(x)}} I(w,x) = I(w,q)$$

□

The next corollary follows from the above theorem.

Corollary 4.9.1. *(1) Let I be a continuous SP score in $\triangle^W$. Then $D(p,q) := \mathbb{E}_{w \sim P}[I(w,q)] - \mathbb{E}_{w \sim P}[I(w,p)]$ is a Bregman divergence in $\triangle^W$.*

(2) Let I be a continuous additive E-SP score in $\triangle^{\mathcal{F}}$. Then $D(p,q) := \mathbb{E}_{w \sim P}[I(w,q)] - \mathbb{E}_{w \sim P}[I(w,p)]$ is an additive Bregman divergence in $\triangle^{\mathcal{F}}$.

Proof. (1) IER always hold in $\triangle^W$. (2) Since I is additive, by Lemma 4.7 it has IER and since I is a E-SP score by Remark 4.2 it is SP. □

Let us compare our results with similar theorems in other literature. Gneiting & Raftery (2007) and Banerjee et al. (2005) showed similar results to Corollary 4.9.1(1). Gneiting & Raftery (2007) showed, in $\triangle^W$, the relation between regular proper scores and Bregman divergences. However, scores and divergences are not assumed necessarily to be continuous. Hence our proof shows more because we derive the continuity of Bregman divergence from the continuity of score. The theorem in Banerjee et al. (2005) is similar to the above Corollary (1) except that Bregman divergences are defined on $\mathbb{R}^m$ instead of $\triangle^W$ and they exclude infinity. By contrast, our proof shows how to handle infinity.

The relation between additive continuous SP scores and additive Bregman divergences in Predd et al. (2009) and Pettigrew (2016) is similar to Corollary 4.9.1(2). Since they are dealing with non-probabilistic credences as well, their result is stronger than ours in the sense that Bregman divergences are defined on $[0,1]^{\mathcal{F}}$ instead of $\triangle^{\mathcal{F}}$. However, they assume additivity and in this sense, our result is stronger.

Compared to other literature, Theorem 4.9 is more comprehensive in the sense that it provides proofs for the cases in $\triangle^W$ and for additive scores in $\triangle^{\mathcal{F}}$ at the same time. On top of that, Theorem 4.9 gives us the way to deal with non-additive scores in $\triangle^{\mathcal{F}}$, at the cost of the assumption of IER.

It is worth asking how our more complicated definition of Bregman divergence played out in the proof of the theorem. Recall that we could not have had Theorem 4.4, if we had used (*) (being finite and continuous in $ri(\triangle^M)$, p.134) instead of part (i) and (ii) in Definition 4.6. On the other hand, we could have had the same form of Theorem 4.9 with a weaker notion of Bregman divergence than ours. Theorem 4.9 tells more with our definition because we demand more in order to be our Bregman divergence and thus we proved more.

From Theorem 4.9 and Theorem 4.4, we have the following claims, which might be viewed as a converse of both theorems in certain conditions.

Corollary 4.9.2. *(1) Let $I : W \times \triangle^M \to [0,\infty]$ be a utility function. I satisfy IER and is continuous SP iff*

$$D_I(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$$

is a Bregman divergence.

(2) Let $D : \triangle^M \times \triangle^M \to [0,\infty]$ be a divergence and suppose that $I_D(w,q) := D(v_w,q)$. D is a Bregman divergence iff

$$D(p,q) = \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)]$$

and $I_D(w,q)$ satisfies IER and continuous in q.

Proof. (1) ($\to$) We can easily check this from the proof of Theorem 4.9.
($\leftarrow$) Since D_I is a divergence, I is SP. Let $p = v_w$ for any $w \in W$. Since $D_I(v_w,q) = I(w,q) - I(w,v_w)$ and D_I is continuous, I is continuous of q.
(2) ($\to$) We can easily check this from the proof of Theorem 4.4.
($\leftarrow$) Since D is a divergence, $D(v_w,q)$ is SP. Thus we can apply Theorem 4.9, and from its proof we know that there is a Bregman divergence d_D that is the same as D. $\square$

To summarize this and the last section, we proved, with our refined definitions, that the DM(Bregman) rules and the EUM(SP) rules have the same extension under certain conditions (IER):

- A strictly proper score I satisfying IER of an EUM rule can be extended to a Bregman divergence D_I such that $D_I(p,q) = \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$, and the DM rule with D_I generates the same results with the EUM rule.

- A Bregman divergence D of a DM rule can be restricted to a strictly proper score I_D such that $I_D(w,q) = D(v_w,q)$, and the EUM rule with I_D generates the same results with the DM rule.

4.3.2 The EUM Rules in General: Examples

In the previous section, we investigated the EUM rules with strictly proper scores on $W \times \triangle^M$. In this section, we deal with the EUM rules that do not presuppose the domain of the second argument of a utility function to be the set of all probability functions. A utility function can be defined, from the beginning, on the set of belief cores. By way of illustration, here is an example of a general EUM rule. This example is designed to include the EUM rule that represents DM(SE).

Example 4.3. *Let $W := \{1, 2, 3\}$ and G be an EUM rule in $\triangle^W$ with utility u. Assume that u is neutral, which means that $u(w, b) = u(\pi(w), b_\pi)$ where π is a permutation on W and $b_\pi := (b_{\pi(w)})_{w \in W}$. u can be given by the left part of Table 4.1. Assume $R, R', W, W' \in \mathbb{R}$ and $R > R' > 0 > -W' > -W$. Let $p = (p_1, p_2, p_3)$ and w.l.o.g assume that $p_1 \geq p_2 \geq p_3$. We can calculate $\mathbb{E}_{w \sim P}[u(w, b)]$ and conditions for $b \in G(P)$ as below.*

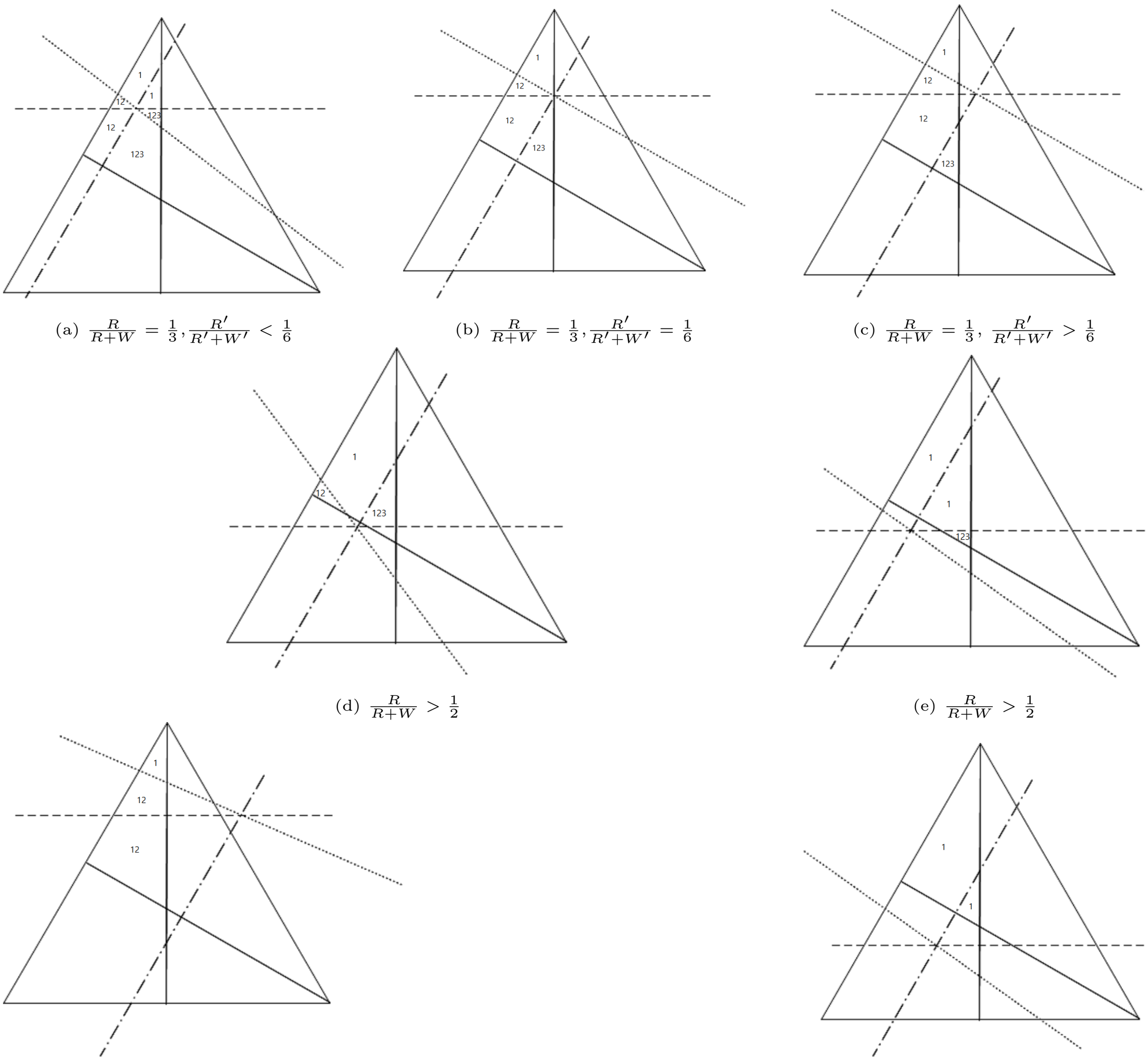
b	$u(w_1, b)$	$u(w_2, b)$	$u(w_3, b)$	$\mathbb{E}_{w \sim P}[u(w, b)]$	condition for $b \in G(P)$
$(1, 0, 0)$	R	$-W$	$-W$	$Rp_1 - W(1 - p_1)$	$p_1 \geq \frac{W}{R+W}$ and $p_1 \geq -\frac{R'+W'}{R+W}p_3 + \frac{R'+W}{R+W}$
$(\frac{1}{2}, \frac{1}{2}, 0)$	R'	R'	$-W'$	$R'(1 - p_3) - W'p_3$	$p_3 \leq \frac{R'}{R'+W'}$ and $p_1 \leq -\frac{R'+W'}{R+W}p_3 + \frac{R'+W}{R+W}$
$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	0	0	0	0	$p_1 \leq \frac{W}{R+W}$ and $p_3 \geq \frac{R'}{R'+W'}$

Table 4.1

These rules yield different results depending on the values of R, R', W',and W. In Figure 4.15, we selected some examples and illustrated where the preimage-regions of belief cores are located. In each probability simplex, a belief core is indicated by numbers — e.g., '12' means that $B = \{w_1, w_2\}$ — in the area where $p_1 > p_2 > p_3$. Since u is neutral, other areas are symmetrically the same . In Figure 4.15, we can observe that the preimage-region $\{p \in \triangle^W | G(P) = b\}$ of any $b \in U^W$ is connected and convex, which we will prove later. Moreover, from Theorem 4.2 follows that (a),(b),(c), and (d) are DM rules and (e), (f), and (g) are not DM rules. Note that the above EUM rule with $\frac{R}{R+W} = \frac{1}{3}$ and $\frac{R'}{R'+W'} = \frac{1}{6}$ (Figure 2.(b)) is the same as DM(SE) rule in Example 4.2.

The above example can be generalized to the cases where $|W| = m$ by the following definition.

Definition 4.10 (Informativeness-sensitive Utility)**.** *Let $|W| = m$. A utility function $u : W \times \triangle^W \to \mathbb{R} \cup \{-\infty\}$ is an informativeness-sensitive utility iff there are*



(a) $\frac{R}{R+W} = \frac{1}{3}, \frac{R'}{R'+W'} < \frac{1}{6}$ (b) $\frac{R}{R+W} = \frac{1}{3}, \frac{R'}{R'+W'} = \frac{1}{6}$ (c) $\frac{R}{R+W} = \frac{1}{3}, \frac{R'}{R'+W'} > \frac{1}{6}$

(d) $\frac{R}{R+W} > \frac{1}{2}$ (e) $\frac{R}{R+W} > \frac{1}{2}$

(f) $\frac{R}{R+W} < \frac{1}{2}, \frac{R'}{R'+W'} > \frac{1}{3}$ (g) $\frac{R}{R+W} > \frac{2}{3}$

Figure 4.15: EUM rules with Informativeness-sensitive utilities (EUM(I))

$R_1, ..., R_l, W_1, ..., W_{m-1}$ *such that* $R_1 > R_2 > ... > R_{m-1} > R_m = 0 > -W_{m-1} > ... > -W_2 > -W_1$ *and*

$$u(w, uni(B)) = \begin{cases} R_{|B|} & \textit{if} \quad w \in B \\ -W_{|B|} & \textit{if} \quad w \notin B \end{cases}$$

where $uni(B)$ *is the representation point in* $\triangle^W$ *of the uniform distribution* $U(B)$ *on* B.

Informativeness-sensitive utility takes cardinalities of belief cores into account, and therefore the EUM rule with this utility function, which we call EUM(I), can evaluate the truth-conductiveness and informativeness of belief cores simultaneously. This utility function gives a greater reward/penalty to braver belief cores when they turn out to be true/false, respectively, in order to evaluate epistetmically risky behavior in a balanced way. DM(SE) is the same as the EUM rule with the utility $u(w, b) := D_{SE}(v_w, uni(W)) - D_{SE}(v_w, b)$, and thus it can be seen as a EUM(I). This observation shows us another advantage of DM(SE), since it measures how informative binary beliefs are as well as how close binary beliefs are to the given probability, and to the truth.

Note that we set $u(w, uni(W)) = 0$ for all $w \in W$. Keeping this fixed, we may relax the condition on W_is. For example, there can be a neutral utility satisfying $-W_{m-1} = ... = -W_2 = -W_1$. However, the EUM rules with this kind of neutral utility cannot involve DM(SE) since it does not hold that $W_i = W_j$ for some i,j from $u(w, b) = D_{SE}(v_w, uni(W)) - D_{SE}(v_w, b)$.

4.3.3 Relation between the DM Rules and the EUM Rules

Now we investigate how to extend the representation theorems between the DM(Bregman) rules and EUM(SP) rules to general cases and under what conditions the DM rules and the EUM rules can represent each other. First, we shall find a sufficient and necessary condition for an EUM rule to be a DM rule, which is a generalization of Theorem 4.9 and Corollary 4.9.2 (1). As the following example shows, it is not the case that every EUM rule is a DM rule.

Example 4.4.

(1) Let $|W| = 2$ and G be an EUM rule with a utility u where u is an informativeness-sensitive utility with $R_1 > R_2 = 0 > -W_1$. If $R_1 > W_1$, then $b = (\frac{1}{2}, \frac{1}{2})$ never maximizes expected utility (see Easwaren (2016), Pettigrew (2017)), even when $P(w_1) = P(w_2) = \frac{1}{2}$. Thus, by Theorem 4.2, this EUM rule cannot be a DM rule.

(2) A utility function u is a simple iff there are $R, W > 0$ such that

$$u(w, uni(B)) = \begin{cases} R & \text{if} \quad w \in B \\ -W & \text{if} \quad w \notin B \end{cases}$$

where $uni(B)$ is the representation point of the uniform distribution $U(B)$ on B. This utility function is simple in the sense that the utilities do not depend on the cardinalities of belief cores. We can easily see that the whole set W always maximizes the EUM rule with a simple utility function, which implies that this rule cannot be a DM rule by Theorem 4.2. We denote this EUM rule by $EUM(S)$.

To see the connection between the EUM rules in general and the DM rules, we introduce strict propriety restricted to uniform distributions on belief cores.

Definition 4.11 (U-Strictly Proper Utility (U-SP))**.** *A utility function $u : W \times U \to \mathbb{R} \cup \{-\infty\}$ is U-strictly proper (U-SP) iff*

$$\operatorname*{argmax}_b \mathbb{E}_{w \sim U(B')}[u(w, b)] = uni(B')$$

for all $B' (\neq \emptyset) \subseteq W$.

This theorem shows under what conditions an EUM rule can be a DM rule.

Theorem 4.10 (Representation of EUM rule by DM rule)**.** *Let a BR G be an EUM rule in $\triangle^M$ with utility u.*

G is invariant under the same input-representation (IIR) and u is U-SP
iff
G is a DM rule in $\triangle^M$.

Proof. It is clear that u is U-SP iff $G(P) = b'$ for any $P \in \mathbb{P}(W)$ such that $p = b'$, which is equivalent to (ii) in Theorem 4.2. Thus, by Theorem 4.2 our claim holds. □

We remark that if G is an EUM rule in $\triangle^W$, IIR always holds. Thus, u is U-SP iff G is a DM rule. This theorem raises the question: what kind of divergence can we use for an EUM rule to be a DM rule. Let G be an EUM rule satisfying the conditions to be a DM rule. We can construct the following divergence.

$$d(p,b) := \begin{cases} 0 & p = b \\ \mathbb{E}_{w\sim P}[-u(w,b)] + \max_b \mathbb{E}_{w\sim P}[u(w,b)] + \epsilon & \text{otherwise} \end{cases}$$

where $\epsilon > 0$ is any positive real number. We need the small error ϵ to avoid the case where $d(p,b) = 0$ when $p \neq b$. Note that P can be any extension of p. Even though $d(p,b)$ might be different depending on which extension we choose, $G(P)$ produces, whatever the choice of extension, the same belief set because G satisfies IIR.

If an EUM rule can be a DM rule, this implies that the EUM rule also obeys the suspension principle that characterizes the DM rules. This raises a question: which kind of utility function should we use to satisfy the suspension principle? Let us give the most simple examples.

Example 4.5. *(1) In Example 4.4 (1), $R_1 < W_1$ iff u is U-SP iff G is a DM rule.*

(2) Compare (1) with the case where $|W| = 3$ in Figure 4.15. We can easily check the following:

(i) $R < W$ is neither necessary nor sufficient to be a DM rule. (d) is a DM rule although $R > W$ and (f) is not a DM rule although $R < W$.

(ii) If $2R' > W'$, then it cannot be a DM rule.(See (f).)

(iii) If $R > 2W$, then it cannot be a DM rule.(See (g).)

Thus what we know so far is just the necessary condition for an EUM rule to be a DM rule such that $2R' \leq W'$ and $R \leq 2W$. Further research is needed to investigate the necessary and sufficient condition on informativeness-sensitive utility for the EUM rule to be a DM rule in the case where $|W| \geq 3$.

(3) In Example 4.4 (2), u is not U-SP.

Now we shall generalize Theorem 4.4 and Corollary 4.9.2 (2) and find a necessary and sufficient condition on d for DM rules to be represented by EUM rules, which is a counterpart of Theorem 4.10. We note that the condition that d is a Bregman divergence is a sufficient condition for a DM rule to be an EUM rule. (See Theorem 4.4.) However, we will show that it is not a necessary condition. Like the case of the Euclidean distance and the squared Euclidean divergence, when two divergences give us the same result, we call them equivalent divergences. Let us give its formal definition.

Definition 4.12 (Equivalence of Divergences)**.** *Let d and d' be a divergence on $\triangle^M$. d is equivalent to d' iff*

$$\underset{b}{\operatorname{argmin}}\, d(p,b) = \underset{b}{\operatorname{argmin}}\, d'(p,b)$$

for all $p \in \triangle^M$

The following theorem shows the condition we are looking for. A DM rule can be an EUM rule iff there is an equivalent distance such that distance minimizing is the same as expected distance minimizing. We call such DM rules the *EDM* rules.

Theorem 4.11 (Representation of DM rule by EUM rule)**.** *Let a BR G be a DM rule in $\triangle^M$ with divergence d. There is a divergence d' equivalent to d such that*

$$\underset{b}{\operatorname{argmin}}\, d'(p,b) = \underset{b}{\operatorname{argmin}}\, \mathbb{E}_{w\sim P}[d'(v_w,b)]$$

for all $P \in \mathbb{P}(W)$

iff
G is a EUM rule in $\triangle^M$.

Proof. ($\rightarrow$) Let $u(w,b) := -d'(v_w,b)$. Then, $G(P) = \operatorname{argmin}_b d(p,b) = \operatorname{argmin}_b d'(p,b) = \operatorname{argmax}_b \mathbb{E}_{w\sim P}[u(w,b)]$.

($\leftarrow$) Let G be an EUM rule with a utility u and $p \in \triangle^M$. For any $b \in U^M$, define

$$d'(p,b) := \begin{cases} \mathbb{E}_{w\sim P}[-u(w,b)] + \max_b \mathbb{E}_{w\sim P}[u(w,b)] & p \in U^M \\ \mathbb{E}_{w\sim P}[-u(w,b)] + \max_b \mathbb{E}_{w\sim P}[u(w,b)] + \epsilon & \text{otherwise} \end{cases}$$

where $\epsilon > 0$ is any positive real number and we can choose P among the extensions of p. (d' can be different depending on which extension we choose.) We will show the following (i), (ii) and (iii). (i) d' is a divergence since G is both a EUM rule and a DM rule and thus u is U-SP. (ii) d' is equivalent to d, since $\operatorname{argmin}_b d(p,b) = \operatorname{argmax}_b \mathbb{E}_{w\sim P}[u(w,b)] = \operatorname{argmin}_b d'(p,b)$ whether $p \in U^M$ or not. (iii) From the definition of d', we have

$$d'(v_w,b) = -u(w,b) + \max_b[u(w,b)]$$

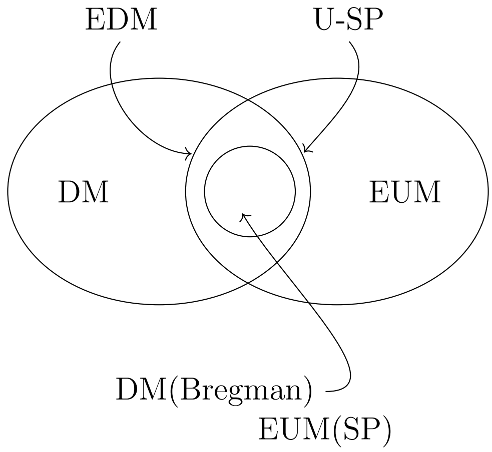


Figure 4.16: U-SP and EDM

Thus, for all $P' \in \mathbb{P}(W)$ such that $p' = p$,

$$\begin{aligned}\operatorname*{argmin}_b \mathbb{E}_{w\sim P'}[d'(v_w, b)] &= \operatorname*{argmin}_b \mathbb{E}_{w\sim P'}[-u(w,b)] \\ &= \operatorname*{argmin}_b \mathbb{E}_{w\sim P}[-u(w,b)] = \operatorname*{argmin}_b d'(p,b)\end{aligned}$$

Note that the second equality holds since G is a DM rule and thus $G(P) = G(P')$, which follows from the fact that DM rules satisfy IIR. (See Theorem 4.2.) □

To summarize this section, we proved that the EUM rules with U-SP utilities satisfying IIR are DM rules, and the DM rules that are also EDM rules are EUM rules.

4.3.4 Convexity of the EUM rules and EUM-rationalizability

Convexity of the EUM rules In Section 3.5, we proposed that the notion of *convexity* should be studied and discussed in the context of belief binarization. When a rational agent determines binary beliefs from probability functions, it is natural that if two probability functions yield the same result, then so does any probability function between them — any linear combination of them. However, explicating the convexity norm is not straightforward and there can be various types of convexity norms. Since the DM rules and EUM rules are defined to be binarization correspondences — i.e., multiple outputs are allowed — , C-convexity and R-covexity are relevant in this chapter. Considering that the ouput of the DM rules or EUM rules is the set of some points $b \in U^M$, we reformulate them as follows[23]:

(C-Convexity) A binarization correspondence G satisfies the convexity of binarization correspondence (*C-convexity*) iff for all $P, P' \in \mathbb{P}(W)$, if $G(P) = G(P')$, then

$$G(\alpha P + (1-\alpha)P') = G(P)(= G(P'))$$

[23]See Definition 3.21

for all $\alpha \in [0, 1]$.

(R-Convexity) A binarization correspondence G satisfies the convexity of binarization relation (*R-convexity*) iff for all $P, P' \in \mathbb{P}(W)$ and $b \in U^M$, if $b \in G(P)$ and $b \in G(P')$, then

$$b \in G(\alpha P + (1-\alpha)P')$$

for all $\alpha \in [0, 1]$.

We showed that the threshold-based rules introduced in Chapter 3 — $\mathrm{HT}^r(\mathrm{S})$, CS^S and CCT^g — satisfiy neither C-convexity nor R-convexity.[24] Moreover, we proved that HT^r satisfies R-convexity but not C-convexity and that DM(SE) satisfies both convexity norms.[25] What about other DM rules and the EUM rules? Notice that every preimage-region of the representation point of a belief core under the EUM(I) rule — see Figure 4.15 — is convex. In this section, we prove that it holds for every EUM rule. This is easy to prove because we can make use of the linearity of the expectation operator.

Theorem 4.12 (Convexity of the EUM rule)**.** *Every EUM rule satisfies C-convexity and R-convexity.*

Proof. Let G be an EUM rule with $G(P) := \mathrm{argmax}_b \, \mathbb{E}_{w \sim P}[u(w, b)]$. Since

$$\mathbb{E}_{w \sim \alpha P + (1-\alpha)P'}[u(w, b)] = \alpha \mathbb{E}_{w \sim P}[u(w, b)] + (1-\alpha)\mathbb{E}_{w \sim P'}[u(w, b)]$$

if

$$\underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim P}[u(w, b)] = \underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim P'}[u(w, b)]$$

then

$$\underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim \alpha P + (1-\alpha)P'}[u(w, b)] = G(P)$$

Thus G satisfies C-convexity.

Moreover, if

$$b' \in \underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim P}[u(w, b)] \text{ and } b' \in \underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim P'}[u(w, b)]$$

then

$$b' \in \underset{b}{\mathrm{argmax}} \, \mathbb{E}_{w \sim \alpha P + (1-\alpha)P'}[u(w, b)]$$

Thus G satisfies R-convexity.

□

Since we know that some DM rules can be considered as EUM rules, we obtain the following corollary.

[24] See Table 3.7.

[25] See Theorem 3.11.

Corollary 4.12.1. *(1) Every DM rule with a Bregman divergence satisfies C-convexity and R-convexity.*

(2) Every DM rule with a divergence d such that there is a divergence d' that is equivalent to d and $\text{argmin}_b\, d'(p,b) = \text{argmin}_b\, \mathbb{E}_{w\sim P}[d'(v_w, b)]$ for all $P \in \mathbb{P}(W)$ satisfies C-convexity and R-convexity.

Proof. From Theorem 4.4 and Theorem 4.11, we know that the above rules are the EUM rules. □

For example, not only DM(SE) but also DM(KL1), defined in $\triangle^M$ as follows, satisfies C- and R-convexity.

Example 4.6. *Consider a DM rule $DM(KL1)$ with KL1 divergence $D_{KL1}(p,b)$ in $\triangle^W$ where*

$$D_{KL1}(p,b) := \begin{cases} \sum_{w\in Supp(p)} p_w log \frac{p_w}{b_w} & \textit{if } Supp(p) \subseteq Supp(b) \\ \infty & \textit{otherwise} \end{cases}$$

KL1 divergence is a Bregman divergence, and thus $DM(KL1)$ satisfies C-convexity and R-convexity.

In addition, d_E is equivalent to D_{SE}, and therefore DM(E) is also an EUM rule. The above theorem and corollary show that the EUM rules and the DM rules that can also be considered as an EUM rules are more advantages than the threshold-based rules in Chapter 3 in terms of convexity.

Convexity and EUM-rationalizability From the above theorem, it follows that a binarization rule G is not an EUM rule if it does not satisfy C-convexity or R-convexity. Note that if there is a $b \in U^M$ whose preimage-region $G^{-1}(b) := \{p \in \triangle^M | G(P) = b\}$ is not convex, then G satisfies neither C-convexity nor R-convexity. Therefore, we can easily conclude that a BR is not an EUM rule when there exists a preimage-region that is not convex. In other words, unless all preimage-regions are convex, we can say that the BR is not an EUM rule. For instance, we can easily check from Figure 3.8, 3.9 and 3.10 that the threshold-based rules HT^r(S), CS^S and CCT^g, respectively, cannot be considered as an EUM rule. This property of not being an EUM rule can be understood as violating an epistemic norm that we call *EUM-rationalizable*. We say a binarization rule is EUM-rationalizable if there is a utility function with which the rule can be viewed as an EUM rule. Thus, the above rules are problematic not only because they do not satisfy convexity but also because they are not EUM-rationalizable, i.e., we cannot interpret them as an EUM rule with any utility function. Therefore, not only convexity but also EUM-rationalizability can be an epistemic norm to be respected in the binarization problem. Moreover, if we accept EUM-rationalizability, then we should also accept the convexity norm. In this sense, EUM-rationalizability supports the convexity norm and gives it one more

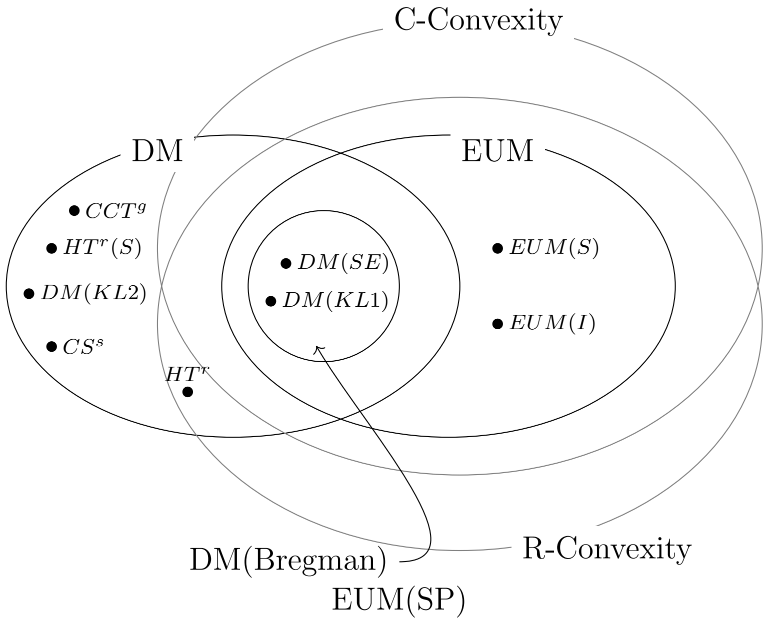


Figure 4.17: Classification of DM and EUM rules

justification. The following example shows the application of non-convexity to show that DM(KL2) is not EUM-rationalizable.

Example 4.7. *Consider a DM rule $DM(KL2)$ with KL2 divergence $D_{KL2}(p, b)$ in $\triangle^W$ where*

$$D_{KL2}(p,b) := \begin{cases} \sum_{w \in Supp(b)} b_w log \frac{b_w}{p_w} & \textit{if } Supp(b) \subseteq Supp(p) \\ \infty & \textit{otherwise} \end{cases}$$

where $Supp(p) := \{w \in W : p_w \neq 0\}$, which is called the support of p for any $p \in \triangle^W$. Let $|W| = 3$. $DM(KL2)$ has a non-convex preimage-region, thus it cannot be EUM-rationalizable. Note that KL2 divergence is not a Bregman divergence.

Proof. Consider the case where $p_1 = (0.7, 0.18, 0.12)$ and $p_2 = (0.18, 0.7, 0.12)$. Then $\mathrm{argmin}_b\, D_{KL2}(p_1, b) = \mathrm{argmin}_b\, D_{KL2}(p_2, b) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. However $\mathrm{argmin}_b\, D_{KL2}(p, b) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ where $p = \frac{p_1+p_2}{2} = (0.44, 0.44, 0.12)$.

$\square$

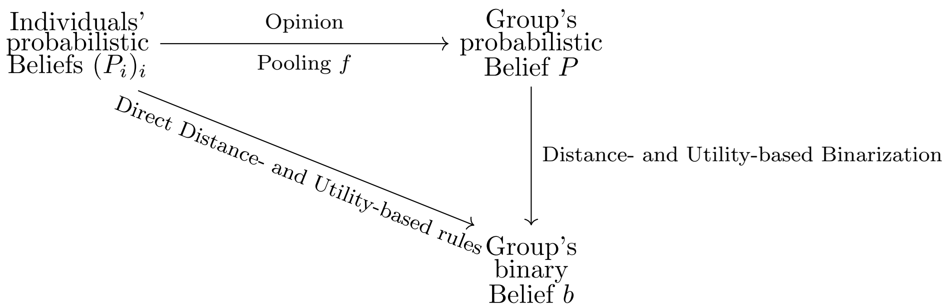
We can now collect the above results and classify the EUM rules and DM rules as illustrated in Figure 4.17. Except for DM(SE), DM(KL1) and DM(KL2), each point represents a type of multiple binarization rules with certain parameters, and belonging to a category — DM/EUM/DM(Bregman)/EUM(SP)/C-Convexity/R-Convexity — means that they always belong to the category regardless of the values of parameters. EUM(S) is introduced in Example 4.4 (1) in the previous section.

It will become clear in the next section what kind of role C-convexity of a binarization rule plays for heterogeneous belief aggregation. It is closely related to some sort of unanimity.

4.4 Distance- and Utility-based Heterogeneous Belief Aggregation and its Properties

This section is based on joint work with Chisu Kim. In this chapter, we deal with distance- and utility-based rules. In the first section, we introduce two groups of distance- and utility-based rules and various examples of them. In the second section, we propose various properties of distance- and utility-based rules centering on the interrelation between the two groups. Moreover, we examine properties of representative distance-based rules. It is worth noting that we also explore C-convexity, introduced in the belief binarization context, in the heterogeneous belief aggregation context. Indeed, we will show that the C-convexity norm is related to some unanimity norms. This permits us to view the C-convexity norm from a broader social perspective.

4.4.1 Distance- and Utility-based Heterogeneous Belief Aggregation



We introduce distance- and utility-based heterogeneous belief aggregation rules. There are two categories of distance- and utility-based HA rules that have the following form:

(1) Pooling + distance- and utility-based BR

$$F((P_i)_i) = \operatorname*{argmin}_b g(f((P_i)_i), b)$$

(2) Direct distance- and utility-based HA

$$F((P_i)_i) = \operatorname*{argmin}_b h((g(P_i, b)_i))$$

where f is an opinion pooling function, g is a distance/(expected) utility function, and h is a function aggregating individual distances/(expected) utilities. Notice that we seek the representation point b of the uniform distribution $U(B)$, instead of the belief core B, which causes no harm because of Lemma 4.1 in Section 4.2.1. We can construct various kinds of distance- and utility-based rules depending on which opinion pooling method is combined with which distance- and utility-based binarization rule, and how individual distances/utilities are aggregated.

Let us consider two kinds of pooling method: linear pooling (LP) and geometric pooling (GP). Here is the formal definition of LP.

Definition 4.13 (Linear Pooling (LP)). *(1) A pooling f is a global LP iff there is a weighting vector $(\alpha_i)_i$ such that for all $(P_i)_i$*

$$f((P_i)_i) = \sum_i \alpha_i P_i$$

(2) A pooling f is a local LP iff for all $(P_i)_i$, there is a weighting vector $(\alpha_i)_i$ such that $f((P_i)_i) = \sum_i \alpha_i P_i$.

Here weighting vector $(\alpha_i)_i$ — $\alpha_i \in [0, 1]$ and $\sum_i \alpha_i = 1$ for all i — is differently used. In global LP, $(\alpha_i)_i$ is fixed for all probability profiles while in local LP, $(\alpha_i)_i$ depends on probability profiles.[26]

Now let us move to the case where f is GP. To make GP well defined, we give up universal domain (UD) and assume that the intersection of the support of each P_i is not empty, i.e., $\bigcap_i Supp(P_i) \neq \emptyset$, which guarantees that $\sum_w \prod_i P_i(w)^{\alpha_i} \neq 0$. Here is the formal definition of GP.

Definition 4.14 (Geometric Pooling (GP)). *(1) An opinion pooling function f is a global GP iff there is a weighting vector $(\alpha_i)_i$ such that for all $(P_i)_i$ and all w*

$$f((P_i)_i)(w) = \frac{\prod_i P_i(w)^{\alpha_i}}{\sum_w \prod_i P_i(w)^{\alpha_i}} (=: GP((P_i)_i)(w))$$

(2) An opinion pooling function f is a local GP iff for all $(P_i)_i$, there is a weighting vector $(\alpha_i)_i$ such that $f((P_i)_i)(w) = \frac{\prod_i P_i(w)^{\alpha_i}}{\sum_w \prod_i P_i(w)^{\alpha_i}}$ for all w.

We can build various kinds of Pooling+distance- and utility-based BR by combining LP or GP with EUM, DM(SE), DM(KL1) or DM(KL2).

Another kind of distance- and utility-based rules are direct distance- and utility-based HA rules. Here are some examples of them.

Definition 4.15 (Linear direct distance- and utility-based HA). *A HA F is a linear direct distance- and utility-based HA iff there are a function $g : \mathbb{P}(W) \times U^M \to \overline{\mathbb{R}}$ and a weighting vector $(\alpha_i)_i$ such that*

$$F((P_i)_i) = \underset{b}{\operatorname{argmin}} \sum_{i=1}^{n} \alpha_i g(P_i, b)$$

for all probability profile $(P_i)_i$.

[26]We distinguish local LP from global LP. It helps us classify various arguments for LP. For example, Pettigrew's arguments for LP (Pettigrew (2020a) can be divided into two types: (1) if the pooled credence lies outside the convex hull generated by the given probability profile, there is a credence, actually the projection onto the convex hull, which (such-and-such)-dominates the former credence; (2) if the pooled credence lies in the convex hull generated by the given probability profile , there is no credence (such-and-such)-dominates the pooled credence. The first type of argument supports only local LP while the second type of argument can be used to support global LP too.

Definition 4.16 (Geometric direct distance- and utility-based HA). *A HA F is a geometric direct distance- and utility-based HA iff there are a function $g : \mathbb{P}(W) \times U^M \to \overline{\mathbb{R}}$ and a weighting vector $(\alpha_i)_i$ such that*

$$F((P_i)_i) = \operatorname*{argmin}_b \prod_{i=1}^n g(P_i, b)^{\alpha_i}$$

for all probability profile $(P_i)_i$.

4.4.2 Commutativity, Unanimity and C-convexity

In this section we investigate properties of distance- and utility-based heterogeneous belief aggregation rules. Our main concern here is properties of Pooling + distance- and utility-based BR, but we address it focusing on the relation with direct distance- and utility-based HA rules.

Let F be a Pooling + distance- and utility-based BR that has the form of $F((P_i)_i) := G(f((P_i)_i)) := \operatorname{argmin}_b g(f((P_i)_i), b)$. First, consider the commutativity with the minimization of a linear mean of individual values $g(P_i, b)$ as follows.

$$\begin{tikzcd}
(P_i)_i \arrow[r, "f \text{ (e.g., LP or GP)}"] \arrow[dr, "\operatorname*{argmin}_b \sum_i \alpha_i g(P_i, b)"'] & P = f((P_i)_i) \arrow[d, "G \text{ (e.g., EUM or DM)}"] \\
& \operatorname*{argmin}_b g(P, b)
\end{tikzcd}$$

Definition 4.17 (Commutativity). *Let $F((P_i)_i) := \operatorname{argmin}_b g(f((P_i)_i), b)$.*

(1) F satisfies global commutativity iff there is a weighting vector $(\alpha_i)_i$ such that for all $(P_i)_i$

$$F((P_i)_i) = \operatorname*{argmin}_b \sum_i \alpha_i g(P_i, b)$$

(2) F satisfies local commutativity iff for all $(P_i)_i$, there is a weighting vector (α_i) such that

$$F((P_i)_i) = \operatorname*{argmin}_b \sum_i \alpha_i g(P_i, b)$$

We note that F satisfies global commutativity iff F can be considered as a linear direct distance-based HA rule with g in Definition 4.15. In this regard, if F satisfies global commutativity, then we can say that F is commute with some linear direct distance-based HA rule with g.

Apart from the minimization of the linear mean of g, consider other direct rules that satisfy some kind of unanimity about g. To commute with those rules, the necessary condition would be that the combination of f and G, denoted by $f + G$, also satisfies the same kind of unanimity.

$$\begin{tikzcd}
(P_i)_i \arrow[r, "f \text{ (e.g., LP or GP)}"] \arrow[dr, "\text{Unanimity}"'] & P = f((P_i)_i) \arrow[d, "G \text{ (e.g., EUM or DM)}"] \\
& \operatorname*{argmin}_b g(P, b)
\end{tikzcd}$$

Definition 4.18 (Unanimity). *Let* $F((P_i)_i) = \operatorname{argmin}_b h(g(P_i, b)_i)$ *or* $F((P_i)_i) = \operatorname{argmin}_b g(f((P_i)_i), b)$.

(1) F satisfies strong unanimity iff if there is a b^* such that for all i $g(P_i, b^*) < g(P_i, b')$, then $b' \notin F((P_i)_i)$, for all b' and for all $(P_i)_i$.

(2) F satisfies weak unanimity iff if for all i and for all b $g(P_i, b^*) < g(P_i, b)$, then $F((P_i)_i) = b^*$, for all b^* and for all $(P_i)_i$.

We note that linear and geometric direct distance- and utility-based HA rules in Definition 4.15 and 4.16 satisfy strong unanimity, therefore weak unanimity as well. Unanimity is a weaker notion than commutativity as the following lemma shows.

Lemma 4.13 (Commutativity and Unanimity). *Let* $F((P_i)_i) := \operatorname{argmin}_b g(f((P_i)_i), b)$.
F satisfies global commutativity
$\Rightarrow$ F satisfies local commutativity
$\Rightarrow$ F satisfies strong unanimity
$\Rightarrow$ F satisfies weak unanimity

Proof. It is clear that from global commutativity follows local commutativity. Assume that given any $(P_i)_i$ we have $F((P_i)_i) = \operatorname{argmin}_b \sum_i \alpha_i g(P_i, b)$. If for all i, $g(P_i, b^*) < g(P_i, b')$, then $\sum_i \alpha_i g(P_i, b^*) < \sum_i \alpha_i g(P_i, b')$. Therefore b' cannot be $F((P_i)_i)$, thus F satisfies strong unanimity. Now assume that F satisfies strong unanimity and for all i, $\operatorname{argmin}_b g(P_i, b) = b^*$. Then all $b'(\neq b^*)$ cannot be $F((P_i)_i)$ by strong unanimity because $g(P_i, b^*) < g(P_i, b')$ for all i. Thus, we must have $F((P_i)_i) = b^*$. □

The above defined 'unanimity' has a close relationship with 'C-convexity' discussed in Section 3.5 and Section 4.3.4 when f is global or local LP.

Theorem 4.14 (C-Convexity and Weak Unanimity). *Let* $F((P_i)_i) := G(f((P_i)_i))$ *— i.e.,* $F = f + G$ *— where* $G(P) := \operatorname{argmin}_b g(P, b)$.

(1) The followings are equivalent:

(i) G satisfies C-convexity

(ii) $F(:= f+G)$ satisfies weak unanimity for all local LP f

(iii) $F(:= f+G)$ satisfies weak unanimity for all global LP f

(2) Let f be global or local LP. If G satisfies C-convexity, then F satisfies weak unanimity.

Proof. Since (2) follows from (1), it is enough to show (1).
(i) $\Rightarrow$ (ii): Let f be a local LP, i.e., for all $(P_i)_i$ there is $(\alpha_i)_i$ such that $f(P_i)_i = \sum_i \alpha_i P_i$. Given $(P_i)_i$, assume that $\text{argmin}_b\, g(P_i, b)(= G(P_i))= b^*$ for all i. If G satisfies C-convexity, which is equivalent to the following: if $G(P_i) = b^*$ for all i, then for all weighting vectors $(\alpha'_i)_i$, $G(\sum_i \alpha'_i P_i) = b^*$, we have $G(\sum_i \alpha_i P_i) = \text{argmin}_b\, g(\sum_i \alpha_i P_i, b) = b^*$.
(ii) $\Rightarrow$ (iii): Every global LP is a local LP.
(iii) $\Rightarrow$ (i): Assume that (iii) holds and $G(P_i) = b^*$ for all i and let $(\alpha_i)_i$ be a weighting vector. Then $G(\sum_i \alpha_i P_i) = G(f((P_i)_i))$ where f is a global LP such that $f((P'_i)_i) = \sum_i \alpha_i P'_i$ for all $(P'_i)_i$. By (iii), $G(f((P_i)_i)) = b^*$. □

In Section 4.3.4, we showed that every EUM rule satisfies C-convexity. Thus all $LP + EUM$ rules satisfy weak unanimity. Note that the converse of (2) does not hold. For a given global or local LP f, C-convexity of G does not follow from weak unanimity of $f+G$. We need to check all other global or all local LP as well in order to obtain C-convexity of G.

The following theorem shows further relations between the above properties and some EUM or DM rules, when f is LP.

Theorem 4.15 (LP+EUM/DM, commutativity and unanimity). *(1) All global $LP+ EUM$ rules satisfy global commutativity.*

(2) All local $LP + EUM$ rules satisfy local commutativity.

(3) Some $LP + DM(KL2)$ rules do not satisfy weak unanimity.

Proof. (1)Since the pooling is a global LP, there is $(\alpha_i)_i$ such that for all $(P_i)_i$, $F((P_i)_i) = \text{argmin}_b\, \mathbb{E}_{w\sim\sum_i \alpha_i P_i}[u(w, b)]$, which is the same as $\text{argmin}_b \sum_i \alpha_i \mathbb{E}_{w\sim P_i}[u(w, b)]$ by the linearity of the expectation operator.
(2)Since the pooling function is a local LP, for all $(P_i)_i$, there is $(\alpha_i)_i$ such that $F((P_i)_i) = \text{argmin}_b\, \mathbb{E}_{w\sim\sum_i \alpha_i P_i}[u(w, b)] = \text{argmin}_b \sum_i \alpha_i \mathbb{E}_{w\sim P_i}[u(w, b)]$.
(3) It follows from Theorem 4.14(1) and Example 4.7(2). □

The following remark shows what happens if $f+G$ does not satisfy weak unanimity when f is a global or local LP.

Remark 4.3. *We have the following if-then chain, when f is a global or local LP.*

(1) $f + G$ does not satisfy weak unanimity $\Rightarrow$ $f + G$ does not satisfy strong unanimity $\Rightarrow$ $f + G$ does not satisfy global/local commutativity $\Rightarrow$ G is not EUM-rationalizable.

(2) $f + G$ *does not satisfy weak unanimity* $\Rightarrow$ G *does not satisfy C-convexity* $\Rightarrow$ G *is not EUM-rationalizable.*

Now we consider the case where f is not LP. Our first question is whether $f + EUM$ rules can satisfy global or local commutativity. Notice that

$$\operatorname*{argmin}_b \sum_i \alpha_i \mathbb{E}_{w\sim P_i}[u(w,b)] = \operatorname*{argmin}_b \mathbb{E}_{w\sim \sum_i \alpha_i P_i}[u(w,b)]$$

Thus, we can say that in order to satisfy global/local commutativity, f needs to be equivalent to a global/local LP with regard to EUM rule with utility u, in the sense that $f + EUM$ rule with u gives us the same result as some $LP + EUM$ rule with u.

Now consider $GP + EUM$ rules. The first part of the following theorem shows that in contrast to $LP+EUM$ rules, not all $GP+EUM$ rules satisfy weak unanimity. Next, our question would be whether there is a distance-based binarization combined with GP that satisfies strong or weak unanimity. The second part of the theorem shows Kullback-Leibler measure goes well with GP.

Theorem 4.16 (GP+DM, commutativity and unanimity)**.** *(1) Some* GP*+*$DM(SE)$ *rules do not satisfy weak unanimity.*

(2) All GP*+*$DM(KL1)$ *rules(Recall* $D_{KL1}(p,b) := \sum_{w\in Supp(p)} p_w log\frac{p_w}{b_w}$ *if* $Supp(p) \subseteq Supp(b)$*. Otherwise,* $:= \infty$*) in* $\triangle^W$ *satisfy weak unanimity and some* $GP + DM(KL1)$ *rules do not satisfy strong unanimity.*

(3) All GP*+*$DM(KL2)$ *rules(Recall* $D_{KL2}(p,b) := \sum_{w\in Supp(b)} b_w log\frac{b_w}{p_w}$ *if* $Supp(b) \subseteq Supp(p)$*. Otherwise,* $:= \infty$*) in* $\triangle^W$ *satisfy global or local commutativity.*

Proof. (1) Let $N = \{1,2\}$ and $|W| = 3$. In $\triangle^3$, consider the case where $p_1 = (\frac{1}{2}, \frac{67}{200}, \frac{33}{200})$ and $p_2 = (\frac{67}{200}, \frac{1}{2}, \frac{33}{200})$. Then $\operatorname{argmin}_b D_{SE}(p_1, b) = \operatorname{argmin}_b D_{SE}(p_2, b) = (\frac{1}{2}, \frac{1}{2}, 0)$. However $\operatorname{argmin}_b D_{SE}(p, b) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ where p is approximately $(0.416, 0.416, 0.168)$, which is GP of p_1 and p_2 with $(\alpha_i)_i = (0.5, 0.5)$.

(2) First of all, we will prove that for all $p \in \triangle^W$,

$$\operatorname*{argmin}_b D_{KL1}(p,b) = uni(Supp(p)) \tag{4.7}$$

(Recall that $uni(Supp(p))(\in \triangle^W)$ is a point that represents the uniform distribution on $Supp(p)$.) Note that $D_{KL1}(p,b) < \infty$, i.e., $Supp(p) \subseteq Supp(b)$, for b that minimizes the divergence. We will compare $D_{KL1}(p, uni(Supp(p)))$ with $D_{KL1}(p,b)$ for b such that $Supp(p) \subsetneq Supp(b)$.

$$\begin{aligned} D_{KL1}(p,b) - D_{KL1}(p, uni(Supp(p))) &= \sum_{w\in Supp(p)} p_w log\frac{p_w}{b_w} - \sum_{w\in Supp(p)} p_w log\frac{p_w}{uni(Supp(p))_w} \\ &= \sum_{w\in Supp(p)} p_w log\frac{uni(Supp(p))_w}{b_w} \\ &> 0 \end{aligned}$$

The inequality follows from the fact that $uni(Supp(p))_w > b_w$, which holds because

$$uni(Supp(p))_w = \frac{1}{|Supp(uni(Supp(p)))|} = \frac{1}{|Supp(p)|} > \frac{1}{|Supp(b)|} = b_w$$

Now to prove weak unanimity, assume that for all i, $\operatorname{argmin}_b D_{KL_1}(P_i, b) = b^*$. It means that $Supp(P_i)$s are all the same and $b^* = uni(Supp(P_i))$ by equation(4.7). Since

$$Supp(GP((P_i)_i)) = \bigcap_i Supp(P_i) = Supp(P_i)$$

we have $\operatorname{argmin}_b D_{KL_1}(GP((P_i)_i), b) = b^*$ by equation(4.7).

Now let us suggest a counterexample to strong unanimity. Let $N = \{1, 2\}$ and $|W| = 3$. In $\triangle^3$, consider the case where $p_1 = (\frac{1}{2}, \frac{1}{2}, 0)$ and $p_2 = (\frac{1}{2}, 0, \frac{1}{2})$. Then $D_{KL1}(p_i, (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))(< \infty) < D_{KL1}(p_i, (1, 0, 0))(= \infty)$ for $i = 1, 2$. However, since $GP((P_i)_i) = (1, 0, 0)$ and $\operatorname{argmin}_b D_{KL1}(p, b) = (1, 0, 0)$, strong unanimity does not hold.

(3) We need to show that

$$\operatorname*{argmin}_b D_{KL2}(GP((P_i)_i), b) = \operatorname*{argmin}_b \sum_i \alpha_i D_{KL2}(p_i, b)$$

First, we will show that for all $(P_i)_i$, there is b such that $D_{KL2}(GP((P_i)_i), b) < \infty$, which is equivalent to $\sum_i \alpha_i D_{KL2}(p_i, b) < \infty$. The equivalence holds because

$$\begin{aligned} D_{KL2}(GP((P_i)_i), b) < \infty \text{ iff } & Supp(GP((P_i)_i)) \supseteq Supp(b) \\ \text{iff } & \bigcap_i Supp(P_i) \supseteq Supp(b) \\ \text{iff } & Supp(p_i) \supseteq Supp(b) \text{ for all i} \\ \text{iff } & \sum_i \alpha_i D_{KL2}(p_i, b) < \infty \end{aligned}$$

Furthermore for all $(P_i)_i$ there is always b such that $\bigcap_i Supp(P_i) \supseteq Supp(b)$, since $uni(\bigcap_i Supp(P_i))$ satisfies it.

Next, because of the first part, it is enough to show that

$$\operatorname*{argmin}_{b:\bigcap_i Supp(P_i) \supseteq Supp(b)} D_{KL2}(GP((P_i)_i), b) = \operatorname*{argmin}_{b:\bigcap_i Supp(P_i) \supseteq Supp(b)} \sum_i \alpha_i D_{KL2}(p_i, b)$$

		Commutativity	Strong Unanimity	Weak Unanimity
LP(Theorem 4.15)	DM(SE)	O	O	O
	DM(KL1)	O	O	O
	DM(KL2)	X	X	X
GP(Theorem 4.16)	DM(SE)	X	X	X
	DM(KL1)	X	X	O
	DM(KL2)	O	O	O

Table 4.2

Let us compute.

$$\begin{aligned}
LHS &= \underset{b}{\operatorname{argmin}} \sum_{w \in Supp(b)} b_w \Big(log b_w - \sum_i \alpha_i log P_i(w) + log \sum_w \prod_i P_i(w)^{\alpha_i} \Big) \\
&= \underset{b}{\operatorname{argmin}} \sum_{w \in Supp(b)} b_w \Big(log b_w - \sum_i \alpha_i log P_i(w) \Big) \\
&= \underset{b}{\operatorname{argmin}} \sum_{w \in Supp(b)} \sum_i \alpha_i b_w \Big(log b_w - log P_i(w) \Big) = RHS
\end{aligned}$$

The second equality follows from $\sum_{w \in Supp(b)} b_w = 1$ □

To summarize this section, we introduced some properties of heterogeneous belief aggregation and their relations: From commutativity follows strong unanimity, from which follows weak unanimity. Moreover, we showed that for all global or local LP, $LP + G$ satisfies weak unanimity iff G satisfies C-convexity. Table 4.2 shows whether LP+ or GP+DM rules with D_{SE}, D_{KL1} or D_{KL2} satisfy commutativity or unanimity, which we proved in this section. The combinations of global/local LP and EUM rules, like DM rules with Bregman divergence, for example D_{SE} or D_{KL1}, satisfy global/local commutativity, hence strong and weak unanimity. As for D_{KL2}, which is not a Bregman divergence, some $LP + DM(KL2)$ violate weak unanimity because the binarization does not satisfy C-convexity. Thus we conclude that strong unanimity and commutativity do not always hold as well. As far as GP concerned, we suggested a counterexamples in which some $GP + DM(SE)$ violate weak unanimity or some $GP + DM(KL1)$ violate strong unanimity. Furthermore, we proved that all $GP + DM(KL1)$ satisfy weak unanimity and all $GP + DM(KL2)$ satisfy commutativity, thus strong and weak unanimity.

4.5 Conclusion

This chapter presented two rational belief binarization rules to determine a belief core, which can be identified with the uniform distribution on it, given a probability function. We introduced a set of focused events $\mathcal{F}(\subseteq \mathcal{P}(W))$ and represented probability functions in De Finetti's coherent polytope $\triangle^{\mathcal{F}}(\subseteq \mathbb{R}^{|\mathcal{F}|})$ as well as in a probability simplex $\triangle^{W}(\subseteq \mathbb{R}^{|W|})$. Using each representation, we formulated the DM rules in $\triangle^{W}/\triangle^{\mathcal{F}}$. Moreover, we modified the definition of Bregman divergence D so that $D(p,q)$ is finite and continuous as long as $q \in \mathbb{F}_p$, i.e., as long as q does not exclude any world that p does not exclude. With this refined definition, we proved that the DM(Bregman) rules with a Bregman divergence D can be represented by the EUM(SP) rules with a strictly proper score I_D such that $I_D(w,q) = D(v_w, q)$. We also proved, the other way around, that the EUM(SP) rules with a strictly proper score I can be represented by the DM(Bregman) rules with a Bregman divergence D_I such that $D_I(p,q) = \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$.

Interestingly, we proved these theorems in both settings without assuming the additivity of scoring rules. We extended this discussion and proved that the EUM rules with U-SP utilities satisfying IIR are DM rules, and the DM rules that are also EDM rules are EUM rules. In addition, we proved that every EUM rule is convex and criticized the threshold-based binarization rules in Chapter 3 because they are not convex and therefore not EUM-rationalizable. Last but not least, we combined EUM rules and DM rules with opinion pooling LP or GP for heterogeneous belief aggregation problems, and compared these combined rules with direct heterogeneous belief aggregation rules aggregating individual distances or utilities. We discussed properties such as commutativity, strong and weak unanimity, and their interrelation: from commutativity follows strong unanimity, from which follows weak unanimity. Moreover, we showed that for all global or local LP, $LP+G$ satisfies weak unanimity iff G satisfies C-convexity. In addition, we examined whether LP+ and GP+DM rules with D_{SE}, D_{KL1} or D_{KL2} satisfy commutativity, strong unanimity or weak unanimity.

Chapter 5

Conclusion

In this thesis, we proposed the heterogeneous belief aggregation problem, and systematically investigated it. In Chapter 2, we first established three impossibility results when the agenda is a non-trivial algebra, and then provided three characterizations of impossibility agendas. In Chapter 3, we classified threshold-based heterogeneous belief aggregation rules and showed that various forms of monotonicity and conjunctiveness characterize them. Moving on to collective belief binarization, we focused on local threshold rules, reviewed the rational belief binarization rules, and identified which rational rules satisfy which properties. Moreover, we proposed various kinds of convexity norms and examined which of them are satisfied by which binarization methods including not only functions but also correspondences, relations and ordinalizations. In Chapter 4, we introduced two novel rational belief binarization rules, namely the DM rules and EUM rules. We investigated their relationship and studied them in the heterogeneous belief aggregation context as well. Among the various DM and EUM rules, we focused on DM(Bregman) and EUM(SP) rules, and proved that they can represent each other.

We would like to close this thesis by commenting on the topics that cross the chapters and adding future research topics. First, we attempted to connect research fields that have been separately studied.

- We proposed a general framework for different types of beliefs.
- We discussed heterogeneous belief aggregation in relation to judgment aggregation in Chapter 2, opinion pooling in Chapters 2 and 4, and belief binarization literature in Chapters 3 and 4.
- We tried to fill the gap between individual and social epistemology addressing e.g., the collective belief binarization problem in Chapters 3 and 4.

We believe that attempts to link different problem areas basically allow generalizations and also raise new problems.

- Unanimity norms can be weakened into super-majority norms. We can study these kinds of norms especially in relation to Chapter 3.

- We can compare collective belief binarization with individual belief binarization.

- Heterogeneous belief aggregation can be generalized to deal with new belief aggregation problems where different types of input data are allowed.

- Informativeness of belief and the veritistic value have not been successfully integrated and explored so far. We think that informativeness-sensitive scoring rules in Chapter 4 will open up new research directions.

- We can investigate diachronic norms imposed on heterogeneous belief aggregation.

- Heterogeneous belief aggregation can be generalized into a group decision problem e.g., aggregating individuals' quantitative beliefs and utilities into a group's qualitative belief and preference.

Second, we emphasized looking at the belief binarization problem and the heterogeneous belief aggregation problem from a geometric point of view throughout this thesis, especially in Chapters 3 and 4. The geometric point of view provided us with different kinds of new concepts, methods, and ideas.

- We utilized the Voronoi diagram to characterize the coherent belief binarization rules in Chapter 3.

- We employed distance measures to devise new belief binarization rules in Chapters 3 and 4.

- We introduced the convexity norms to evaluate belief binarization rules in Chapters 3 and 4.

We believe that this geometric approach is promising and opens up many new research challenges for us.

- We can investigate geometrical characterizations of epistemic norms imposed on belief binarization such as being stable and being r-likely.

- We can further address the relationship between the Voronoi diagrams and the belief bianrization problem, which will give us a bridge between the conceptual space program and the belief binarization problem.[1]

- The relationship between the linear pooling methods and the convexity norm can be examined. More generally, our study can be extended to distance- and utility-based opinion pooling.[2]

[1]For studies related to this question, see Decock et al. (2014) where the concept of *knowledge* is modeled using conceptual space.

[2]For studies related to this question, see Abbas (2009), Pettigrew (2019), Adamcik (2014a), Feldbacher-Escamill & Schurz (2020), Neyman & Roughgarden (2021).

Mathematical Background

Definition .1 (Convex set and Convex function)**.** *Let $A \subseteq \mathbb{R}^m$.*

(1) A is called convex if for all $x, y \in A$ it holds that $\alpha x + (1-\alpha)y \in A$ for all $\alpha \in [0,1]$.

(2) The convex hull of A: $Conv(A) = \{z \in \mathbb{R}^m |\ z = \alpha x + (1-\alpha)y$ for some $x, y \in A, \alpha \in [0,1]\}$

(3) Let A be convex. A function $f : A \to \mathbb{R}$ is convex if for all $x, y \in A$ it holds that $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ for all $\alpha \in [0,1]$.

Definition .2 (Basic topological notions)**.** *Let $x \in \mathbb{R}^n$, $\epsilon > 0$, and $A \subseteq \mathbb{R}^n$.*

(1) The open ball of the centre x and radius ϵ: $B_\epsilon(x) = \{y \in \mathbb{R}^n |\ ||x-y|| < \epsilon\}$

(2) An element $x \in A$ is called an interior point of A if $\exists \epsilon > 0 B_\epsilon(x) \subseteq A$

(3) The interior of A: $int(A) = \{x \in A |\ \exists \epsilon > 0 B_\epsilon(x) \subseteq A\}$

(4) A is called open if x is an interior point of A for all $x \in A$

(5) A is called closed if $\overline{A}$ is open

(6) the closure of A: $cl(A) = \{x \in \mathbb{R}^n |\ \forall \epsilon > 0\ B_\epsilon(x) \cap A \neq \emptyset\}$

(7) The boundary of A: $bd(A) = cl(A) \setminus int(A)$

(8) A is said to be affine if $l[x,y] \subset A$ for all $x, y \in A$ where $l[x,y] = \{\lambda x + (1-\lambda)y |\ \lambda \in \mathbb{R}\}$

(9) The affine hull of A: $aff(A) = \bigcap \{C \subseteq \mathbb{R}^m |\ C$ is affine and $A \subseteq C\}$

(10) The relative interior of A: $ri(A) = \{x \in A |\ \exists \epsilon > 0 (B_\epsilon(x) \cap aff(A)) \subset A\}$

(11) The relative boundary of A: $rb(A) = cl(A) \setminus ri(A)$.

Deutsche Zusammenfassung

Diese Dissertation schlägt ein neues Forschungsthema dahingehend vor, wie wir mehrere individuelle probabilistische Überzeugungen in Hinblick auf logisch zusammenhängende Propositionen zu einer kollektiven binären Überzeugung aggregieren sollten: heterogene Überzeugungsaggregation. Wir argumentieren, dass heterogene Überzeugungsaggregation eine Untersuchung wert ist, da es viele Situationen gibt, in denen probabilistische Überzeugungen und binäre Überzeugungen plausible und naheliegende Inputs bzw. Outputs von Aggregationsverfahren darstellen. Das Hauptproblem besteht darin, dass heterogene Überzeugungsaggregation anfällig für ein Dilemma wie das diskursive Dilemma oder das Lotterieparadox ist: Propositionsbezogene unabhängige Verfahren können möglicherweise keine deduktive Abgeschlossenheit und Konsistenz gewährleisten. Angesichts dieser Situation haben wir zwei Hauptfragen: Wie das Dilemma präzisiert und verallgemeinert werden könnte und welche Arten von Aggregationsverfahren das Dilemma vermeiden und rationale kollektive Überzeugungen erhalten können.

Um die erste Frage zu beantworten, wenden wir den axiomatischen Ansatz an, um allgemeine Aggregationsverfahren wie in der Urteilsaggregation und der Theorie der sozialen Wahl behandeln zu können. Wir untersuchen, welche individuellen und kollektiven Rationalitätsanforderungen und welche Eigenschaften von Aggregationsverfahren der heterogenen Überzeugungsaggregation auferlegt werden sollten und welche ihrer Kombinationen unmöglich sind. Wir gehen hauptsächlich von deduktiver Abgeschlossenheit und nicht von Vollständigkeit aus, anders als in der meisten Literatur zur Urteilsaggregation. Darüber hinaus adressieren wir Unmöglichkeitsergebnisse ohne Anonymitätsbedingungen, die bei der Überzeugungsbinarisierung nicht berücksichtigt werden können. Dies führt zu drei Arten von Unmöglichkeitsergebnissen, und wir bestimmen die notwendige und hinreichende Agenda-Bedingung für jedes der Ergebnisse. Darüber hinaus analysieren wir Ähnlichkeiten und Unterschiede zwischen unseren Beweisen und anderen verwandten Beweisen und kommen zu dem Schluss, dass das Problem der heterogenen Überzeugungsaggregation nicht auf die anderen verwandten Probleme reduziert werden kann. Schließlich zeigen wir dass unsere Methoden auf andere ähnliche Unmöglichkeiten angewendet werden können.

Für die zweite Frage untersuchen wir spezifische heterogene Aggregationsverfahren und deren Eigenschaften. Es gibt dabei zwei Arten von heterogenen Aggregationsverfahren: die kollektive Überzeugungsbinarisierung kombiniert mit einer probabilistischen Meinungspooling-Methode und die direkten Regeln.

Was die kollektive Überzeugungsbinarisierung betrifft, so sind Theorien der Überzeugungsbinarisierung anwendbar. Dazu analysieren wir zunächst die bestehenden schwellenwertbasierten Verfahren, insbesondere solche, die die Lockesche These lockern und die Rationalität bewahren. Wir kategorisieren sie als lokale Schwellenwertregeln —- wobei Schwellenwerte von Wahrscheinlichkeitsmaßen abhängen —- und Weltschwellenwertregeln – wobei Schwellenwerte nicht auf eine Proposition, sondern auf eine mögliche Welt angewendet werden. Die nämlichen Regeln können mittels der Eigenschaft der lokalen Monotonie bzw. der Weltmonotonie charakterisiert werden. Wir vergleichen und setzen diese Eigenschaften mit anderen bestehenden Eigenschaften wie Stabilität in der Stabilitätstheorie von Überzeugungen und mit neuen — noch einzuführenden — Eigenschaften in Beziehung. Ob einige bestehende rationale Verfahren, wie die Kamera-Shutter-Regel, diese Eigenschaften erfüllen, ist eine interessante und philosophisch wichtige Frage. Wir geben geometrische Charakterisierungen einiger der Eigenschaften an, um diese Frage zu beantworten. Darüber hinaus schlagen wir vor, Konvexitätsnormen im Kontext der Überzeugungsbinarisierung zu diskutieren. Wir führen verschiedene Arten von Konvexitätsnormen ein und untersuchen, ob diese nämlichen Verfahren diese erfüllen.

Weiters schlagen wir zwei neue Arten von Überzeugungsbinarisierungsmethoden vor, die Rationalität bewahren, aber nicht auf Schwellenwerten basieren: Distanzbasierte Binarisierung und Epistemischer-Nutzen-basierte Binarisierung. Die erste ist eine holistische Methode, die den Abstand von einem gegebenen Wahrscheinlichkeitsmaß zu der resultierenden binären Überzeugung minimiert. Der zweite basiert auf einer Genauigkeitsnorm, die die erwartete Unrichtigkeit (inaccuracy) minimiert. Wir entwickeln neue Wege, um Distanz und Unrichtigkeit zu messen. Wir entwickeln neue Wege, um Distanz und Ungenauigkeit zu messen. Darüber hinaus untersuchen wir Distanzminimierung mit Bregman-Divergenz, Nutzenmaximierung mit strikten "proper score" und deren Beziehung.

Direkte heterogene Überzeugungsaggregationsregeln werden ebenfalls vorgeschlagen und hinsichtlich Schwellenwert, Distanz und epistemischer Nützlichkeit untersucht. Wir erstellen eine Klassifizierung und Charakterisierung dieser Regeln. Darüber hinaus untersuchen wir verschiedene, besonders in sozialen Kontexten relevante Normen, wie verschiedene Einstimmigkeits- und Konvexitätsnormen, die in sozialen Kontexten interpretiert werden, und Kommutativitätsnormen, die den Zusammenhang zwischen direkten Regeln und Kombinationen von probabilistischem Meinungspooling und kollektiver Überzeugungsbinarisierung aufzeigen.

Zusammenfassend kommen wir zu dem Schluss, dass heterogene Überzeugungsaggregation ein philosophisch fruchtbares Thema darstellt, das Aufmerksamkeit verdient. Heterogene Überzeugungsaggregation kann als ein allgemeiner Rahmen angesehen werden, in dem nicht nur heterogene Überzeugungsaggregation, sondern auch pobabilistisches Meinungspooling, Urteilsaggregation und Überzeugungsbinarisierung gemeinsam untersucht werden können. Erstens ist die Untersuchung der heterogenen Überzeugungssaggregation an sich interessant und lässt sich nicht auf andere Forschungsfelder reduzieren: Wir können uns mit unterschiedlichen Rationalitätsnormen in sozialen Kontexten auseinandersetzen und Eigenschaften der heterogenen Überzeugungsaggregation charakterisieren. Darüber hinaus sind es nicht nur die direkten Regeln, sondern auch die verschiedenen möglichen Kombinationen von Methoden aus unterschiedlichen Forschungsgebieten, die diese Unternehmung zu mehr als bloß der Summe ihrer Teile werden lassen. In der Tat verbindet zweitens dieser Rahmen unabhängig entwickelte Forschungsbereiche: Einerseits können wir gut entwickelte formale Theorien der formalen Erkenntnistheorie wie Überzeugungsbinarisierungstheorien und epistemische Entscheidungstheorien auf das Überzeugungsaggregationsproblem anwenden. Andererseits ermöglicht uns dieser Rahmen, soziale Kontexte zu Überzeugungsbinarisierungsproblemen und epistemischen Entscheidungstheorien hinzuzufügen, die somit auf die Behandlung sozialer Überzeugungen ausgedehnt werden können. Unsere Theorie der heterogenen Überzeugungsaggregation kann auf das (kollektive) Überzeugungsbinarisierungsproblem und die epistemische (kollektive) Entscheidungstheorie angewendet werden. Auf diese Weise schließt diese Arbeit die Lücke zwischen individueller Erkenntnistheorie und kollektiver Erkenntnistheorie oder verkleinert dieselbe zumindest.

Bibliography

[1] Abbas, Ali E. (2009). A Kullback-Leibler View of Linear and Log-Linear Pools. *Decsion Analysis* **6(1)**: 25-37.

[2] Adamcik, Martin (2014a). *Collective reasoning under uncertainty and inconsistency*. PhD thesis. University of Manchester.

[3] Adamcik, Martin (2014b). The information Geometry of Bregman Divergences and Some Applications in Multi-Expert Reasoning. *Entropy* **16**: 6338-6381.

[4] Arlo-Costa, Horacio, Vincent F. Hendricks, Johan Van Benthem (eds.) (2016). *Readings in Formal Epistemology : Sourcebook*. Springer.

[5] Banerjee, Arindam, Guo, Xin, and Wang, Hui (2005). On the optimality of conditional expectation as a bregman predictor. *IEEE Transactions of Information Theory* **51**: 2664–2669.

[6] Bauschke, Heinz H., and Patrick L. Combettes (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer.

[7] Blackwell, David, and Lester Dubins (1962). Merging of opinions with increasing information. *The Annals of Mathematical Statistics* **33**: 882–886.

[8] Brier, Glenn W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review.* **8**: 1–3.

[9] Briggs, Rachael, Fabrizio Cariani, Kenny Easwaran, and Branden Fitelson (2014). Individual Coherence and Group Coherence. in: J. Lackey (ed.) *Essays in Collective Epistemology*, pp. 215–239. Oxford: Oxford University Press.

[10] Cantwell, John, and Hans Rott (2019). Probability, coherent belief and coherent belief changes. *Annals of Mathematics and Artificial Intelligence* **87**: 259–291.

[11] Cariani, Fabrizio (2016). Local supermajorities. *Erkenntnis* **81**: 391–406.

[12] Chandler, Jake (2013). Acceptance, aggregation and scoring rules. *Erkenntnis* **8**: 201–217.

[13] Carvalho, Arthur (2016). An Overview of Applications of Proper Scoring Rules. *Decision Analysis* **13(4)**: 223–242.

[14] Decock, Lieven, Igor Douven, Chrisoph Kelp, and Sylvia Wenmackers (2014). Knowledge and Approximate Knowledge. *Erkenntnis* **79**: 1129–1150.

[15] DeGroot, Morris H. (1974). Reaching a Consensus. *Journal of the American Statistical Association* **69(345)**: 118–121.

[16] Deza, Michel Marie, and Elena Deza (2016). *Encyclopedia of Distances (4th.)*. Springer.

[17] Dietrich, Franz, and Christian List (2007). A generalised model of judgment aggregation. *Social Choice and Welfare* **28**: 529-565.

[18] Dietrich, Franz, and Christian List (2008). Judgment aggregation without full rationality. *Social Choice and Welfare* **31**: 15–39.

[19] Dietrich, Franz, and Christian List (2013). Propositionwise judgment aggregation: the general Case. *Social Choice and Welfare* **40**: 1067–1095.

[20] Dietrich, Franz, and Christian List (2017a). Probabilistic opinion pooling generalized. Part one: general agendas. *Social Choice and Welfare* **48**: 747–786.

[21] Dietrich, Franz, and Christian List (2017b). Probabilistic opinion pooling generalized. Part two: the premise-based approach. *Social Choice and Welfare* **48**: 787–814.

[22] Dietrich, Franz, and Christian List (2018). From degrees of belief to binary beliefs: Lessons from judgement-aggregation theory. *Journal of Philosophy* **115(5)**: 787-814.

[23] Dietrich, Franz, and Christian List (2021). The relation between degrees of belief and binary beliefs: a general impossibility theorem. In: I. Douven (ed.) *Lotteries, Knowledge, and Rational Belief. Essays on the Lottery Paradox*, pp. 223-254. CUP.

[24] Dietrich, Franz, and Philippe Mongin (2010). The premise-based approach to judgment aggregation. *Journal of Economic Theory* **145(2)**: 562–582.

[25] Dietrich, Franz, and Kei Spiekermann (2021). Social Epistemology. In: M. Knauff & W. Spohn (eds.) *The Handbook of Rationality*, pp. 579-590. MIT Press.

[26] Dietrich, Franz, and Kei Spiekermann (2021b). Jury Theorems. In Miranda, F., P. J. Graham, D. Henderson, N. Pedersen (eds.) *The Routledge Companion to Social Epistemology*. Routledge .

[27] Dokow, Elad, and Ron Holzman (2010a). Aggregation of binary evaluations. *Journal of Economic Theory* **145**: 495–511.

[28] Dokow, Elad, and Ron Holzman (2010b). Aggregation of binary evaluations with abstentions. *Journal of Economic Theory* **145**: 544–561.

[29] Douven, Igor, and Peter Gärdenfors (2019). What are natural concepts? A Design Perspective. *Mind and Language* **35(3)**: 313-334.

[30] Duddy, Conal, and Ashley Piggins (2012). A Measure of Distance Between Judgment Sets. *Social Choice and Welfare* **39**: 855–867.

[31] Easwaran, Kenny (2016). Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities. *NOUS* **50(4)**: 816–853.

[32] Easwaren, Kenny, and Branden Fitelson (2015). Accuracy, Coherence and Evidence. in *Oxford Studies in Epistemology* **5**:61-96.

[33] Engelking, Ryszard, and Karol Siekluci (1992). *Topology: A Geometric Approach*. Heldermann, Berlin.

[34] Feldbacher-Escamilla, Christian, and Gerhard Schurz (2020). Optimal Probability Aggregation Based on Generalized Brier Scoring. *Annals of Mathematics and Artificial Intelligence* **88**:717–734.

[35] Gärdenfors, Peter (1988). *Knowledge in Flux. Modeling the Dynamics of Epistemic States*. MIT Press.

[36] Gärdenfors, Peter (2000). *Conceptual Spaces*. Cambridge MA: Bradford.

[37] Gärdenfors, Peter (2006). A representation theorem for voting with logical consequences. *Economics and Philosophy* **22(2)**: 181–190.

[38] Geiting, Tilmann, and Adrian E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* **102(477)**: 359-378.

[39] Genest, Christian (1984). Pooling operators with the marginalization property. *Canadian Journal of Statistics* **12(2)**: 74-86.

[40] Genest, Christian (1984). A characterization theorem for externally Bayesian groups. *Annals of Statistics* **12(3)**: 1100-1105.

[41] Genest, Christian, K. J. McConway, and M. J. Schervish (1986). externally Bayesian pooling operators. *Annals of Statistics* **14(2)**: 487-501.

[42] Genest, Christian, and Carl Wagner (1987). Further evidence against independence preservation in expert judgment synthesis. *aequationes Mathematicae* **32(1)**: 74–86.

[43] Genest, Christian, and James V. Zidek (1986). Combining probability distributions: a critique and annotated bibliography. *Statistical Science* **1(1)**: 114–135.

[44] Goldman, Alvin I. (1999). *Knowledge in a Social World*. Oxford: Clarendon.

[45] Goodin, Robert E., and Kei Spiekermann (2018). *An epistemic theory of democracy.* Oxford: Oxford University Press.

[46] Greaves, Hilary, and David Wallace (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind* **115(459)**: 607–632.

[47] Grossi, Davide, and Gabriella Pigozzi (2014). *Judgment Aggregation: A Primer.* CA: Morgan and Claypool.

[48] Hájek, Alan (2009). Arguments For—or Against—Probabilism?. in: Huber and Schmidt-Petri (2009), 229–251.

[49] Halpern, Joseph Y. (2017). *Reasoning about uncertainty.* MIT press.

[50] Hansson, Sven Ove (1999). *A Textbook of Belief Dynamics. Theory Change and Database Updating*, Dordrecht: Kluwer.

[51] Hernandez-Conde, J. V. (2017). A case against convexity in conceptual spaces. *Synthese* **194(10)** : 4011–4037.

[52] Herzberg, Frederik (2015). Aggregating infinitely many probability measures. *Theory and Decision* **78, no. 2**: 319–337.

[53] Herzberg, Frederik (2017). Respect for experts vs. respect for unanimity: The liberal paradox in probabilistic opinion pooling. *Economic Letters* **151**: 44-47.

[54] Hintikka, Jaakko (1962) [2005]. *Knowledge and Belief: An Introduction to the Logic of the Two Notions, second edition*, Vincent F. Hendriks and John Symons (eds.), (Texts in Philosophy, 1). London: College Publications.

[55] Huber, Franz, and C. Schmidt-Petri (eds.) (2009). *Degrees of Belief.* Dordrecht: Springer.

[56] Hug, Daniel, and Wolfgang Weil (2020). *Lectures on Convex Geometry.* Springer.

[57] Ivanovska, Magdalena, and Marija Slavkovik (2019). Aggregating Probabilistic Judgments. *Proceedings TARK 2019.*

[58] Jackson, Elizabeth G (2020). The relationship between belief and credence. *Philosophy Compass.*

[59] Joyce, James M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, **65**: 575–603.

[60] Kornhauser, L. A., and L. G. Sager (1986). Unpacking the Court. *Yale Law Journal* **96**: 82-117.

[61] Kyburg, Henry E. Jr., and Michael Pittarelli (1991). Set-Based Bayesianism. *IEEE Transcations Systems, Man and Cybernetics* 26(3):324-339.

[62] Landes, Jürgen (2015). Probabilism, entropies and strictly proper scoring rules. *International Journal of Approximate Reasoning* **63**: 1-12.

[63] Leitgeb, Hannes (2013) Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic* **164**: 1338-1389.

[64] Leitgeb, Hannes (2014a). The Stability Theory of Belief. *Philosophical Review* **123(2)**: 131–171.

[65] Leitgeb, Hannes (2014b). Belief as a Simplication of Probability, and What This Entails. In: Baltag, Alexandru, and Sonja Smets (eds) *Johan van Benthem on Logic and Information Dynamics*, pp. 405-417. Springer.

[66] Leitgeb, Hannes (2015) The Humean Thesis on Belief. *Proceedings of the Aristotelian Society* **89/1**: 143-185.

[67] Leitgeb, Hannes (2017a). *The Stability of Belief: How Rational Belief Coheres with Probability*. Oxford: Oxford University Press.

[68] Leitgeb, Hannes (2017b). Imaging All the People. *Episteme* **14(4)**:463-479.

[69] Leitgeb, Hannes (2020). A Structural Justification of Probabilism: From Partition Invariance to Subjective Probability. *Philosophy of Science*.

[70] Leitgeb, Hannes (2021). Stability and the Lottery Paradox. In: I. Douven (ed.)*Lotteries, Knowledge, and Rational Belief: Essays on the Lottery Paradox*, pp. 147-170. CUP.

[71] Leitgeb, Hannes, and Richard Pettigrew (2010a). An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science* **77(2)**:201-235.

[72] Leitgeb, Hannes, and Richard Pettigrew (2010b). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science* **77(2)**:236-272.

[73] Lehrer, Keith, and Carl Wagner (1981). *Rational Consensus in Science and Society*. Dordrecht/Boston: Reidel.

[74] Lehrer, Keith, and Carl Wagner (1983). Amalgamation and the Independence Issue: A Reply to Laddaga. *Synthese* **55(3)**: 339-346.

[75] Levi, Issac (1967). *Gambling with the Truth: An Essay on Induction and the Aims of Science*. Cambridge, Mass.: MIT Press.

[76] Levi, Issac (1980). *The Enterprise of Knowledge*. MIT Press.

[77] Lin, Hanti, and Kevin T. Kelly (2012a). Propositional Reasoning that Tracks Probabilistic Reasoning. *Journal of Philosophical Logic* **41(6)**: 957–981.

[78] Lin, Hanti, and Kevin T. Kelly (2012b). A Geo-logical Solution to the Lottery Paradox. *Synthese* **186(2)**: 531–575.

[79] List, Christian (2004). A Model of Path-Dependence in Decisions over Multiple Propositions. *American Political Science Review*, **98(3)**: 495–513.

[80] List, Christian (2014). Three Kinds of Collective Attitudes. *Erkenntnis*, **79(S9)**: 1601-1622.

[81] List, Christian, and Philip Pettit (2002). Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy* **18(1)**: 89-110.

[82] List, Christian, and Philip Pettit (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.

[83] List, Christian, and Clemens Puppe (2009). Judgment aggregation: A survey. In C. List & C. Puppe (eds.), *Handbook of Rational and Social Choice*. Oxford: Oxford University Press.

[84] McCarthy, John (1981). Measures of the value of information. *Proc. Nat. Acad. Sci.* **42**: 654–655.

[85] McConway, K. J. (1981). Marginalization and linear opinion pools. *Journal of the American Statistical Association* 76: 410–414.

[86] Midolo, L.F., and G. De Marco (2009). Right inverses of linear maps on convex sets. *Topology and its Applicatipons* **156**: 1186–1191

[87] Mordukhovich, Boris S., and Nguyen Mau Nam (2013). *An Easy Path to Convex Analysis and Applications* (Synthesis Lectures on Mathematics and Statistics). Morgan & Claypool Publishers.

[88] Mongin, Philippe (2012). The doctrinal paradox, the discursive dilemma, and logical aggregation theory. *Theory and Decision* **73**: 315–345.

[89] Mongin, Philippe, and Marcus Pivato (2016). Social Preference and Social Welfare Under Risk and Uncertainty. In: M. D. Adler and M. Fleurbaey (eds.) *Handbook of Well-being and Public Policy*, pp. 711-745. Oxford: Oxford University Press.

[90] Mormann, Thomas (1993). Natural Predicates and Topological Structures of Conceptual Spaces. *Synthese* 95: 219–240.

[91] Morris, Peter A. (1974). Decision analysis expert use. *Management Science* **20**: 1233–1241.

[92] Nehring, Klaus, and Clemens Puppe (2005). The structure of strategy-proof social choice - Part II: Non-dictatorship, anonymity and neutrality. Mimeographed.

[93] Nehring, Klaus, and Clemens Puppe (2007). The structure of strategy-proof social choice - Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory* **135**: 269-305.

[94] Nehring, Klaus, and Clemens Puppe (2010). Abstract Arrovian aggregation. *Journal of Economic Theory* **145**: 467–494.

[95] Neyman, Eric, and Tim Roughgarden (2021). From Proper Scoring Rules to Max-Min Optimal Forecast Aggregation. *EC '21: Proceedings of the 22nd ACM Conference on Economics and Computation.*

[96] Pettit, Philip (2001). Deliberative Democracy and the Discursive Dilemma. *Philosophical Issues* 11(1): 268–299.

[97] Pettigrew, Richard (2016). *Accuracy and the laws of credence*, Oxford: Oxford University Press.

[98] Pettigrew, Richard (2017). Epistemic utility and the normativity of logic. *Logos and Episteme 8(4)*:455-492.

[99] Pettigrew, Richard (2019). Aggregating incoherent agents who disagree. *Synthese* **196**: 2737–2776.

[100] Pettigrew, Richard (2020a). On the Accuracy of Group Credences. In: T. S. Gendler and J. Hawthorne (ed.) *Oxford Studies in Epistemology Volume 6*. Oxford: Oxford University Press.

[101] Pettigrew, Richard (2020b). *The Dutch Book Arguments*, Cambridge, UK: Cambridge University Press.

[102] Predd, J., Seiringer, R., Lieb, E. H., Osherson, D., Poor, V., and Kulkarni, S. (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions of Information Theory* **55(10)**: 4786–4792.

[103] Raidl, Eric (2021). Epistemic-State Parallelism: Translating Between Probabilities and Ranks. *Erkenntnis* **86(1)**: 209-236.

[104] Raidl, Eric, and Niels Skovgaard-Olsen (2017). Bridging Ranking Theory and the Stability Theory of Belief. *Journal of Philosophical Logic*, **46(6)**: 577–609.

[105] Rockafellar, Ralph Tyrell (1970). *Convex Analysis*. Princeton University Press.

[106] Savage, Leonard J. (1971). Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association* **66(336)**: 783-801.

[107] Schlag, Karl H., Tremewan, James, van der Weele, Joël J.(2015). A penny for your thoughts: a survey of methods for eliciting beliefs. *Experimental Economics* **18(3)**: 457–490.

[108] Schurz, Gerhard (2019). Impossibility Results for Rational Belief. *Noûs* **53(1)**: 134–159.

[109] Spohn, Wolfgang (2012). *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford: Oxford University Press.

[110] Stirling, W. C., and D. R. Morrell (1991). Convex Bayes Decision Theory. *IEEE Transactions Systems, Man and Cybernetics* 21(1):173-183.

[111] Thorn, Paul (2018). The joint aggregation of beliefs and degrees of belief. *Synthese* **197(12)**:5389-5409.

[112] Wolfers, J. and Zitzewitz, E. (2004). Prediction Markets. *Journal of Economic Perspectives* **18(2)**: 107–126.

[113] Wright, Jonathan (2018). Rigged lotteries: a diachronic problem for reducing belief to credence. *Synthese* **195**: 1355-1373.