COMPLETE DEDUCTIVE SYSTEMS FOR PROBABILITY LOGIC WITH APPLICATION TO HARSANYI TYPE SPACES

Chunlai Zhou

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Mathematics
Indiana University
August 2007

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Jon Michael Dunn, Ph.D.

Lawrence S. Moss, Ph.D.

Daniel Leivant, Ph.D.

Amr Sabry, Ph.D.

August 2007

To Professors Mike Dunn and Qingyu Zhang

Happy Life After Retirement

Acknowledgements

I have been very lucky to meet so many outstanding individuals along my five-year Ph.D. study in Bloomington. Without their help, I could not have finished my doctorate in mathematics. The following words are just a part of the gratitude that I owe them.

It was because of Professor Mike Dunn that I came to IU to pursue my Ph.D. degree. When I studied logic in China, I became fascinated by his algebraic methods on relevance logic. After I arrived at IU, he has always tried to find time from his busy university administration to teach it to me. I am more than grateful to him for his enormous trust of my ability in logic and for his generous support of my academic activities. He recommended me to many important researchers in my interested fields. When I attended summer schools and could not work in the Mathematics Department in the summer of 2005, he gave me a financial support for the whole summer. He also helped me several times when I really needed it. His kindness towards me makes me feel like family. I have never felt constrained in his Dean's office.

Without constant encouragement and guidance from Professor Larry Moss, I could not have completed my Ph.D. in mathematics. His recommendation helped me to get into the Ph.D. program in the Mathematics Department. In 2005, the most difficult year during my stay here, he gave me invaluable encouragement to build up my confidence so that I could bounce up and move on. His captivating teaching and insightful guidance definitely helped me take the first step in my dissertation research. He carefully read the first part of my thesis and made "countably infinite" detailed suggestions. I am lucky to learn from his expertise in so many forefront fields in logic.

I would like to thank my committee members Professors Leivant and Sabry. Their teaching and research has fostered my strong interest in logics for programming languages. The faculty of the Department of Mathematics enriched my intellectual life. I am particularly thankful to Bill Ziemer and Richard Bradley. My appreciation of the beauty of modern analysis is due to their patient teaching. My special gratitude is to the Mathematics Graduate Director Professor Zumbrun. I received from him a William B. Wilcox Mathematics Award, which is the most important honor during my Ph.D. life. I cherish this forever. Many thanks to the Department Graduate Secretaries Misty Cummings and Kate Bowman for their help. Also I want to thank my professors in Amsterdam Professors Dick de Jongh, Johan van Benthem and Yde Venema. My one-year focus study there laid a sound foundation for my future research in logic. I sincerely acknowledge a Graduate Fellowship (2001-2002) from the IU Graduate School and a Spinoza Fellowship (2002-2003) from Professor Johan van Benthem.

During my study here, I found a lot of stimulating study companions and friends. Jason Teutsch is my most helpful friend in the Mathematics Department. He and I have worked together on many topics in mathematics such as real analysis, probability theory and, of course, logic. I have learned from him a lot about Latex, mathematical logic and theory of computing. Joshua Sack is another important colleague and has helped me through regular study meetings with my preparation for Tier 3 and with my thesis research. He also gave suggestions to the introduction to this thesis. Jiho Kim helped me with programming languages and in proofreading this thesis to make it more readable. Thanks also to Cornelia, Jung-Miao, Jiun-Chau, Qingshan and Wahkwan at IU and Fenrong, Xinwen and Xiaoyu in China. Outside my study and research, my Chinese and Taiwanese friends made my life in Bloomington quite enjoyable: Jingyun, Sherry, Hongxin, Wangxin, Xingxing, Tongxin, Congyi, Liangyi, Duyu, Jiahu, Yuan-Shuo, Shenjie, Wuyin, Mengqi, Youwei, Yuanyuan, Zhanglan and Zhiping but most of all, Tongxin and Zhiping. My special thanks are to

Professor Tim O'Conner, Liequan, Daniel and Kevin, who helped me survive in my first year in Bloomington.

Finally I would like to acknowledge the support from my family members especially my two lovely twin nieces Meirong and Manrong, who make me realize that life is always so beautiful.

Abstract

These days, the study of probabilistic systems is very popular not only in theoretical computer science but also in economics. There is a surprising concurrence between game theory and probabilistic programming. J.C. Harsanyi introduced the notion of type spaces to give an implicit description of beliefs in games with incomplete information played by Bayesian players. Type functions on type spaces are the same as the stochastic kernels that are used to interpret probabilistic programs. In addition to this semantic approach to interactive epistemology, a syntactic approach was proposed by R.J. Aumann. It is of foundational importance to develop a deductive logic for his probabilistic belief logic.

In the first part of the dissertation, we develop a sound and complete probability logic Σ_+ for type spaces in a formal propositional language with operators L_r^i which means "the agent i's belief is at least r" where the index r is a rational number between 0 and 1. A crucial infinitary inference rule in the system Σ_+ captures the Archimedean property about indices. By the Fourier-Motzkin's elimination method in linear programming, we prove Professor Moss's conjecture that the infinitary rule can be replaced by a finitary one. More importantly, our proof of completeness is in keeping with the Henkin-Kripke style. Also we show through a probabilistic system with parameterized indices that it is decidable whether a formula ϕ is derived from the system Σ_+. The second part is on its strong completeness. It is well-known that Σ_+ is not strongly complete, i.e., a set of formulas in the language may be finitely satisfiable but not necessarily satisfiable. We show that even finitely satisfiable sets of formulas that are closed under the Archimedean rule are not satisfiable. From these results, we develop a theory about probability logic that is parallel to the relationship between explicit and implicit descriptions of belief types in game theory. Moreover, we use a linear system about probabilities over trees to prove that there is no strong completeness even for probability logic with finite indices. We conclude that the lack of strong completeness does not depend on the non-Archimedean property in indices but rather on the use of explicit probabilities in the syntax.

We show the completeness and some properties of the probability logic for Harsanyi type spaces. By adding knowledge operators to our languages, we devise a sound and complete axiomatization for Aumann's semantic knowledge-belief systems. Its applications in labeled Markovian processes and semantics for programs are also discussed.

Contents

CHAPTER 1

Introduction

"*I've always believed in numbers. In equations and logics that lead to reason. But after a lifetime's such pursuits, I ask, what truly is logic? Who decides reason?* "

–John Nash, *A Beautiful Mind*

Knowledge and Belief Knowledge and belief have been actively studied by researchers starting with ancient Greek philosophers. It was not until recently that they received a rigorous formal treatment in modal logic and in modern probability theory, which was necessitated by their substantive uses in artificial intelligence and game theory. Computer scientists and game theorists are more concerned about interactive epistemology, i.e., knowledge and belief among a group of agents such as the relationship among knowledge, belief, and action, while philosophers are more interested in questions about epistemology of a single agent, such as "what does it mean to say that someone knows something." Kripke structures, a standard semantics for modal logic, provide a mathematical framework for reasoning about knowledge. It is well-known that the deductive system $S5$ is a sound and complete axiomatization with respect to the class of Kripke structures for knowledge where accessibility relations are equivalence relations. Now we use a simple example to illustrate this kind of structures.

$$\begin{tikzcd}
(H,H) \arrow[r, leftrightarrow, "1"] \arrow[d, leftrightarrow, "2"'] & (H,T) \arrow[d, leftrightarrow, "2"] \\
(T,H) \arrow[r, leftrightarrow, "1"'] & (T,T)
\end{tikzcd}$$

Example 1.0.1. Consider a game between player 1 and player 2. Each of them has a coin. After tossing his coin, each of them knows the outcome of his own toss but does not know the outcome of the other. There are four possible states after tossing in the game: $(H, H), (H, T), (T, H)$ and (T, T). (H, T) means that player 1's outcome is heads and player 2 gets tails. The meanings of other states are similar. The Kripke structure for this simple game is given by the diagram above. For example, at state (H, H), player 1 considers only (H, H) and (H, T) possible because he knows his own outcome but does not know the outcome of player 2's toss. According to this analysis, at each state, each player knows that the other player knows his own outcome.

Reasoning about belief is more subtle than reasoning about knowledge. In this thesis, we are concerned only about *quantified beliefs* like the statement "I believe that the chance of it raining today is *ninety percent*." We reconsider the example above, now adding probabilities to it.

EXAMPLE 1.0.2. The game is played the same as in the previous example except that we further assume that player 1 has a fair coin, i.e. after tossing it, it is equally likely to land heads and tails, and player 2 has a biased coin which is twice as likely to land heads as to land tails. At state (H, H), player 1 believes that the chance for player 2's coin to land head and tail are 2/3 and 1/3, respectively. So he assigns 2/3 to the state (H, H) (or itself) and 1/3 to (H, T). Since, at that state, he never thinks the states (T, H) or (T, T) possible, he assigns 0 to both of them. Similarly, we can define both players' beliefs at other states. We can then consider beliefs of agents about other agents' beliefs. At the state (H, H), player 1 has one-third belief in player 2's belief that chance of both coins' landing tails is one half.

How do we represent and reason about quantified beliefs? Can we find a mathematical framework for belief analogous to Kripke structures for knowledge?

Modeling Beliefs and Belief Types Reconsider the statement "I believe that the chance of rain today is at least ninety percent." This belief is about another statement that it will rain today. Mathematically, statements are modeled as events. Since the first statement involves probabilities, it is natural to consider these statements in the context of a measurable space $S = \langle \Omega, \mathcal{A} \rangle$ where Ω is a state space and $\mathcal{A}$ is a σ-algebra on this space. So we interpret this quantified belief using an operator $L_{0.9}$ on $\mathcal{A}$, i.e. a mapping from $\mathcal{A}$ to $\mathcal{A}$. L here stands for "at least". If A stands for the event that it will rain today, then $L_{0.9}A$ denotes the belief that the chance of rain today is at least ninety percent. For any index $r \in \mathcal{Q} \cap [0, 1]$, the operator L_r is defined similarly. For such a family of belief operators L_r, it is easy to check that they satisfy the following properties: for events $A, A_n \in \mathcal{A}$,

$$L_0(A) = \Omega \tag{1.1}$$

$$L_1(\Omega) = \Omega \tag{1.2}$$

$$L_r A \subseteq \sim L_s(\sim A), r + s > 1 \tag{1.3}$$

$$r_n \uparrow r \Rightarrow L_{r_n} A \downarrow L_r A \tag{1.4}$$

$$L_r(A \cap B) \cap L_s(A \cap (\sim B)) \subseteq L_{r+s} \tag{1.5}$$

$$\sim L_r(A \cap B) \cap \sim L_s(A \cap (\sim B)) \subseteq \sim L_{r+s} A \tag{1.6}$$

$$A_n \downarrow A \Rightarrow L_r A_n \downarrow L_r A \tag{1.7}$$

where $\uparrow$ ($\downarrow$) means infinitely approaching by an (a) increasing (decreasing) sequence. The first three properties say that degrees of beliefs are always between 0 and 1. (1.5) and (1.6) state that belief operators are finitely additive. (1.7) is the *continuity from above property* from measure theory. These three properties ensure that belief operators are σ additive. (1.4) states that these operators are continuous in degrees (indices r in L_r). This kind of treatment of belief interpreted by operators is analogous to the treatment of knowledge in Kripke structures, interpreted by partition-induced operators. In order to make this parallel between knowledge and belief complete, we still need a quantified version of "Kripke structures" for quantified beliefs. Such quantified Kripke structures, which are called *belief types* in game theory, have played a major role in economic theory and game theory.

There are two approaches to belief types in game theory. The first represents beliefs explicitly, and is called the *explicit description* of beliefs. Such a description starts with a space of states of nature, which specify parameters of a game, such as payoff functions. Next it specifies the beliefs of the agents about the space of states of nature, and then the beliefs about the combination of the nature space with the beliefs about the nature space and so on. An explicit belief type consists of a hierarchy of beliefs which satisfy the *coherence* requirement that different levels of beliefs of every agent do not contradict one another. In the first layer, the beliefs are represented by probability measures over the nature space, and the beliefs in the second layer are represented by probability measures over the space of probability measures in the first layer, and so on. Such a straightforward description provides all possible belief types, which form an explicit model for beliefs. However, this model is hardly a workable model considering the complexity of the representations of the belief types in it.

The second is the *implicit* description of beliefs, which is what we want for giving formal semantics for beliefs. It was introduced by Harsanyi in 1960's [16] for games with incomplete information played by Bayesian players. This description is defined in a measure space $S = \langle \Omega, \mathcal{A} \rangle$. For each agent, each state of this space is associated with a state of nature and a probability measure on the space. His *implicit* belief type at the state is this probability measure. It provides a belief over the nature space. Since each state is associated with a belief type, it also defines beliefs of beliefs about the nature space, and so on. So we can extract a hierarchy of beliefs (or simply explicit description) from this implicit description. If we ignore the association with the nature space, namely the economic content, the above association of states to probability measures is called a *type function* from Ω to the measure space $\Delta(S)$ of probability measures on S. This type function corresponds to a Markovian kernel T_i for each agent i on $\Omega \times \mathcal{A}$, i.e., a function from $\Omega \times \mathcal{A}$ to $[0, 1]$ which satisfies the following two conditions:

- for each $w \in \Omega$, $T_i(w, \cdot)$ is a probability measure on S; and
- for each $A \in \mathcal{A}$, $T_i(\cdot, A)$ is an $\mathcal{A}$-measurable function.

The triplet $\langle \Omega, \mathcal{A}, T_i \rangle$ is called *a type space*. It was shown in Samet [31] that there is a natural one-to-one correspondence between such defined type spaces and the families above of belief operators. Type spaces are exactly the expected "quantified Kripke structures" for beliefs. Note that the introspectivity conditions would be captured by the standard Harsanyi type spaces within the multi-agent setting of game theory.

Syntactic Formalism An alternative to the above semantic approach to belief is Aumann's syntactic approach. The building blocks of the syntactic formalism are *formulas*. They are constructed from propositional letters (which are interpreted by Aumann as "natural occurrences", in contrast with the nature space in Harsanyi's type spaces) by the Boolean connectives and a family of belief operators L_r^i where $r \in \mathcal{Q} \cap [0,1]$. The characteristic feature of the syntax is this family of operators. The interpretation of $L_r^i \phi$ is that the agent i's belief in the event ϕ is at least r. L_r^i is the syntactic counterpart of the semantic belief operator L_r [1] on σ-algebras. In this thesis, we mainly focus on reasoning about beliefs of one agent. So, for simplicity, we omit the label i.

With this language, we devise a deductive system for quantified beliefs which we prove to be sound and complete with respect to the class of type spaces above. It is not surprising that our logical formulation Σ_+ contains many properties from semantic belief operators.

Probability Logic Σ_+

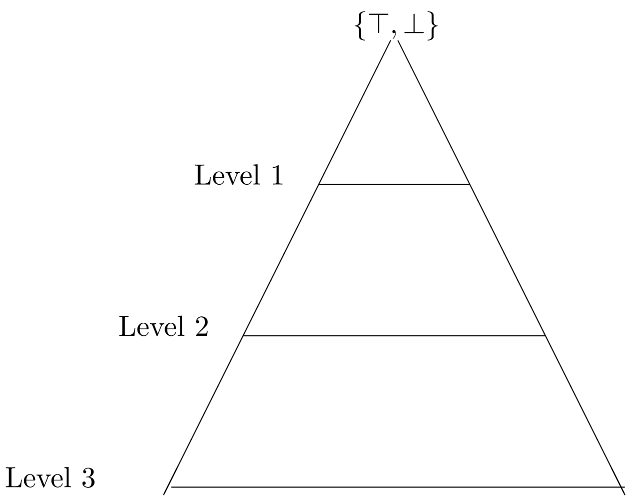
- (A0) Propositional logic
- (A1) $L_0 \phi$
- (A2) $L_r \top$
- (A3) $L_r(\phi \wedge \psi) \wedge L_t(\phi \wedge \neg\psi) \to L_{r+t}\phi$, $r + t \leq 1$
- (A4) $\neg L_r(\phi \wedge \psi) \wedge \neg L_s(\phi \wedge \neg\psi) \to \neg L_{r+s}\phi$, $r + s \leq 1$
- (A5) $L_r \phi \to \neg L_s \neg\phi$, $r + s > 1$

[1] We use the same symbol L_r^i for both the syntactic and semantic operators. Context should determine which we mean.

- (DIS) If $\vdash \phi \leftrightarrow \psi$, then $\vdash L_r\phi \leftrightarrow L_r\psi$.
- (ARCH): If $\vdash \gamma \rightarrow L_s\phi$ for all $s < r$, then $\vdash \gamma \rightarrow L_r\phi$

The most important principle is the rule (ARCH). It captures the Archimedean property of the index set $\mathcal{Q} \cap [0, 1]$ of our language. In the presence of this rule, we can give a natural proof of the completeness of the system Σ_+ with respect to the above semantics, which can also be regarded as a natural generalization of standard completeness proofs through canonical models in modal logics to probability logics. There is one more thing that we need to deal with. The rule (ARCH) is not finitary but *infinitary*. In order to replace it with a finitary principle, we need to consider its relation with reasoning about linear inequalities.

Reasoning about quantified beliefs can be stratified into different levels according to depths of formulas. At each level, probabilistic reasoning is reduced to reasoning about linear inequalities plus propositional reasoning with probabilistic constraints.



The rule (ARCH) makes sure that at each level each finite *consistent* set of linear inequalities is *satisfiable* in a type space. Another rule (DIS) says that L_r is a syntactic congruence operator. *Coherence* of probabilistic beliefs is already implicit in this syntactic representation. Indeed, our syntactic formulation is very similar to the *explicit* description of belief types in games with incomplete information.

With this interpretation in terms of linear programming, we can replace the infinitary rule (ARCH) with a finitary rule by the Fourier-Motkzin elimination method. The elimination method tells us that if a system of linear inequalities has a solution, we can use a simple algorithm to find it within a finite number of steps. In terms of syntax, it follows that, for any formulas γ and ϕ, there is a sufficiently small rational ϵ which depends only on the syntactic forms of these two formulas such that if $\gamma \to L_{r-\epsilon}\phi$ is derivable in Σ_+, so is $\gamma \to L_r\phi$. So we conclude that Σ_+ with this replacement provides a finitary deductive system to reason about beliefs.

Strong Completeness and Compactness It is well-known that strong completeness and compactness fail for probability logics. For example, the set $\{L_{\frac{1}{2}-\frac{1}{4^n}}p : n \in \mathcal{N}\} \cup \{\neg L_{\frac{1}{2}}p\}$ is consistent and hence finitely satisfiable, but it is not satisfiable. How about the sets Γ of formulas that avoid such obvious non-satisfaction, i.e. if $L_s\phi \in \Gamma$ for all $s < r$, then $L_r\phi \in \Gamma$? Such sets of formulas, if consistent, are called *admissible*. In other words, are admissible sets satisfiable? The question is parallel to the question whether coherent beliefs are always types in game theory. This kind of analogy between belief types and probability logic is summarized in the following table:

belief type	probability logic
implicit description	semantics
explicit description	syntax
coherence	weak completeness
coherent beliefs are types	strong completeness
universal type space	canonical model
certain topological properties	compactness

Heifetz and Samet [19] showed that coherent beliefs are not always types. In the thesis, we prove that admissible sets are not necessarily satisfiable. Just as the equivalence between implicit description and explicit description breaks down for belief type in the general measure-theoretic context, the strong completeness of the deductive system with respect to the semantics does not hold for admissible sets. It was shown by Heifetz and Samet [18] and by Moss and Viglizzo [29] that, in the general measure theoretic context, there is a universal type space which consists of all hierarchies of beliefs that are types. In this thesis, we prove that the canonical model whose elements are satisfiable admissible sets is a type space and actually the "biggest" type space. A similar result for the multi-agent case was proved in Moss and Viglizzo[29]. In other words, in the context of general measure theory, belief types in game theory correspond elegantly to our deductive system for reasoning about beliefs. Only in the last row of the above table is the analogy between belief type and probability logic less strong. For a belief type, certain topological properties such as compact Hausdorff and Polish [26] can be imposed on the nature space, and these properties are preserved along the hierarchies of beliefs. So, under these topological conditions, coherent beliefs are types. But, for probability logic, it is hard to imagine how to impose certain topological properties on the algebra of formulas of depth 0, which are Boolean combinations of propositional letters, i.e. purely propositional formulas.

Surprisingly, compared to the easy-going proofs in modal logic, it is difficult to generalize the completeness proof through the canonical model to Σ_+ plus some simple high-order probability formulas like modal logics. Here we take the logic $\Sigma_1 = \Sigma_+ + (L_1\phi \to L_1L_1\phi)$ as an illustration. For separable measurable spaces $S = \langle \Omega, \mathcal{A} \rangle$, the extra axiom $L_1\phi \to L_1L_1\phi$ characterizes the transitivity property of S: $T_i(w)(\{w' \in \Omega : \forall A \in \mathcal{A}(T_i(w')(A) = 0 \Rightarrow T_i(w)(A) = 0)\}) = 1, \forall w \in \Omega$. In order to show the completeness of Σ_1 with respect to the class of transitive type spaces, we should prove that the canonical type space is transitive. If we use such a finitary method as filtration, the syntactic characterization of transitivity requires formulas that are of depth one more than the maximal depth of formulas in the canonical frame. So finitary methods are not applicable here. Infinitary methods are the only candidate. To make sure that infinitary methods do not produce the non-Archimedean property, we can restrict the index set to a finite set of rationals between 0 and 1. If we could show that any consistent set of formulas with finite indices is satisfiable, then the completeness of Σ_1 would follow immediately. However, we show that there is no hope of such strong completeness for probability logics with finite indices. So, it seems that *the failure of strong completeness does not ultimately come from the non-Archimedean property in indices but rather from the use of explicit probabilities in the syntax.*

Two Important Systems Harsanyi type spaces are those type spaces where each agent is *introspective*, i.e., is certain of his type, at each state. They have played a major role in games with incomplete information played by Bayesian players. It was proved in [18] and [29] that there is a universal Harsanyi type space, which contains all possible types. So implicit descriptions do not lose the generality of explicit descriptions. The property of being introspective can be characterized by the two axiom schemas: $L_r\phi \to L_1 L_r\phi$ and $\neg L_r\phi \to L_1\neg L_r\phi$. Let Σ_H denote the logic Σ_+ plus these two axiom schemas. We show that this system provides a complete axiomatization of Harsanyi type spaces. Moreover, we support the idea that introspective beliefs of a single agent gain simplicity by showing that each Σ_H-consistent atom in a finite language has a unique maximal consistent extension if we only increase the depth.

In his seminal paper [2], Aumann introduced another important kind of structures to reason about both knowledge and belief. Aumann's knowledge-belief semantic systems are *conservative* combinations of Harsanyi type spaces and partition structures for knowledge. He proved that the canonical system consisting of all semantically-closed maximal consistent sets of formulas is an Aumann's knowledge-belief system. But his definition of consequence is not syntactic but semantic. He did not succeed in devising a deductive system for a syntactic definition of consistency. By adding some axiom schemata to the basic probability logic Σ_+ and S_5, we formulate a logical system which is sound and complete with respect to the class of Aumann's knowledge-belief systems. Because of the simplicity of both epistemology and the beliefs of a single agent, so is his *combination* of knowledge and belief. Common knowledge is one of the most important notions in *interactive epistemology.* By adding Halpern and Moses' axioms for common knowledge [15], we get a complete axiomatization of Aumann's knowledge-belief semantical systems with common knowledge.

In addition to their applications in belief types of game theory, probability logics are also used in theoretical computer science. *Probabilistic bisimulation* is a mathematical formulation of the equivalence of two probability models (on type spaces). Two models M and M' are probabilistic bisimilar if they satisfy the same propositional letters and they match in their type functions. Actually, our language of probability logics provides a logical characterization of probabilistic bisimulation in the sense that two probability models are probabilistic bisimilar if and only if they satisfy the same formulas. We can use this theory to finitely approximate labeled Markov processes. Type functions are also used as semantics for programs. They were employed by Kozen [22], [23] to provide a semantics for probabilistic programs. With this interpretation, a logical language can be developed to allow explicit reasoning about probabilistic programs. Moreover, programs as type functions seem to give insights to Peter Selinger's semantics for quantum programming language [33]. His slogan for this approach is "quantum data and classical control." He interpreted quantum programs as superoperators on density matrices, which is just the quantum counterpart of Kozen's interpretation of probabilistic programs as linear operators with norms ≤ 1 on measure spaces. These promising applications in theoretical computer science will guide us in our future research on probability logics.

Overview Chapter 2 introduces core concepts of probability logics such as semantics, syntax, axiomatization and decidability. In Chapter 3, we develop a theory about probability logic that parallels the theory of belief types in economics. Two important systems from interactive epistemology are analyzed in detail in Chapters 4 and 5. The last two chapters sketch the application of probability logics to labeled Markov processes and to semantics of programs.

CHAPTER 2

Basic Probability Logic

2.1. Introduction

Reasoning about knowledge and belief has been an active topic of investigation for researchers in such diverse fields as philosophy [20], economics [2] and artificial intelligence [14]. It is well-known that $S5$ provides a sound and complete axiomatization for the logic of knowledge with respect to the class of Kripke structures of knowledge. However, the logical system of belief is still a subject to be explored.

First we demonstrate how probabilities are involved in expressing beliefs. Here we cite one example from [31]. Jane says that the probability of putting a human being on Mars in her lifetime is eighty percent. This is a belief about the statement: “In her lifetime, human beings will land on Mars.” In probability theory, statements are modeled as events. But, if we want to make beliefs the subject of beliefs, then beliefs to some degree themselves must be events. So, beliefs transform statements to statements; mathematically, they map events to events. This gives an intuitive semantics for beliefs.

Ever since [16], *type spaces* have been the most important models for providing implicit description of belief types in economic theory and game theory. Each point of a type space is called *a state*, and is associated with a state of nature (which can be though of as a specification of the parameters of a game) as well as a probability measure for each agent. Each agent’s probability measure at each state is his *type*, which describes his beliefs about the events of the space. Type functions are also called Markovian kernels in probability theory.

In addition to this semantic approach to belief types, R. Aumann in [4] proposed the second syntactic one. Beliefs are expressed in formulas which are recursively constructed from atomic formulas according to some well-known rules. The logical formulation is of foundational importance because it enables the *explicit* construction of Kripke structures for beliefs from more primitive statements and, moreover, it would show that the above two approaches are *equivalent*. The canonical model in the proof of the completeness would provide a *universal* type space in [18]. It is essential to provide a deductive logic for belief types.

In this dissertation we give an axiomatization of probability logic with respect to the class of type spaces. Our system is based on the work in [14] by Fagin and Halpern and in [17] by Heifetz and Mongin. In [14], Fagin and Halpern define a rich langauge for their logic. The logical language includes not only formulas expressing probabilities but also linear combinations of probability formulas. In order to accommodate this rich syntax with "arithmetic" connectives, they have to formulate an independent system for linear inequalities. Their system is sound and complete with respect to the class of type spaces. In [17], Heifetz and Mongin use a much simpler syntax suggested in [4] by Aumman. Its characteristic feature is captured by the belief operators L_r^i for rationals $r \in [0, 1]$, which is interpreted as "the agent i's belief is at least r". In some sense, this syntax is a quantified syntax of modal logic. Their finitary system, based on Aumann's original system, is also shown to be sound and complete with respect to the class of type spaces.

However, our formulation is essentially different from the above work in that we don't use any arithmetic formulas like reasoning about linear inequalities in [14] or any arithmetic style rule like the rule (B) in [17]. The most important contribution of our thesis is our proof of completeness. In order to show completeness, both [14] and [17] use some theorems from linear programming and convex analysis, and their definition of probability measures on their canonical models is not canonical in the sense that it is not totally determined by the syntax. This method does not agree with the tradition from [24]. In our proof of completeness, however, the probability measure on the canonical model is totally determined by the syntax. In some sense, our completeness is a generalization of Kripke's completeness proof to the modal logic of probability.

2.2. Semantics and Syntax

The syntax of our logic is very similar to that of modal logic. We start with a fixed infinite set $P := \{p_1, p_2, \cdots\}$ of propositional letters. We also use $p, q, \cdots$ to denote propositional letters. The set of formulas Φ is built from propositional letters as usual by connectives $\neg$, $\wedge$ and a countably infinite modalities L_r for each $r \in \mathcal{Q} \cap [0, 1]$, where $\mathcal{Q}$ is the set of rational numbers. Equivalently, a formula ϕ is formed by the following syntax:

$$\phi := p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid L_r\phi \ (r \in \mathcal{Q} \cap [0, 1])$$

L_r is the primitive modality in our language. But we also use a derived modality M_r which means "at most" in our semantics through the following definition:

$$\text{(DEF M)} \quad M_r\phi := L_{1-r}\neg\phi.$$

Let $\mathcal{L}$ be the formal language consisting of the above components. We use $r, s, \alpha, \beta, \cdots$ (also with subscripts) to denote rationals. Next we describe the semantics of our system. A *probability model* is a tuple

$$M := \langle \Omega, \mathcal{A}, T, v \rangle$$

where

- Ω is a non-empty set, which is called *the universe or the carrier set* of M;
- $\mathcal{A}$ is a σ-field of subsets of Ω;
- T is a measurable mapping from Ω to the space $\Delta(\Omega, \mathcal{A})$ of probability measures on Ω, which is endowed with the σ-field generated by the sets:

$$\{\mu \in \Delta(\Omega, \mathcal{A}) : \mu(E) \geq \alpha\} \text{ for all } E \in \mathcal{A} \text{ and rational } \alpha \in [0, 1],$$

- v is a mapping from P to $\mathcal{A}$, i.e. $v(p) \in \mathcal{A}$.

T is called a *type function* on the space and $\langle \Omega, \mathcal{A}, T \rangle$ is called a *type space*.

Remark. Let T be a type function. Define $S(w, A) := T(w)(A)$ for any $w \in \Omega$ and $A \in \mathcal{A}$. It is easy to check that S is a Markovian kernel, i.e., it satisfies the following two conditions:

(1) $S(w, \cdot)$ is a probability measure on the sigma-algebra $\mathcal{A}$ for any $w \in \Omega$;

(2) $S(\cdot, A)$ is an $\mathcal{A}$-measurable function for any $A \in \mathcal{A}$.

Conversely, if S is a Markovian kernel on $(\Omega, \mathcal{A})$, then such a defined function $T : \Omega \to \Delta(\Omega, \mathcal{A})$ that $T(s)(A) := S(s, A)$ for any $s \in \Omega$ and $A \in \mathcal{A}$ is a type function. In the proof of the converse statement, we need the following standard theorem from real analysis [1] or [39]:

Theorem 2.2.1. *Let $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ be a function between measurable spaces, and let $\mathcal{C}$ generate Σ_Y. Then f is measurable if and only if $f^{-1}(C) \in \Sigma_X$ for each $C \in \mathcal{C}$.*

The forcing relation $\models$ between states and formulas is defined inductively as follows:

- $M, w \models p$ iff $w \in v(p)$ for propositional letters p;
- $M, w \models \phi_1 \wedge \phi_2$ iff $M, w \models \phi_1$ and $M, w \models \phi_2$;
- $M, w \models \neg\phi$ iff $M, w \not\models \phi$;
- $M, w \models L_r\phi$ iff $T(w)([[\phi]]) \geq r$, where $[[\phi]] := \{w \in \Omega : M, w \models \phi\}$.

ϕ is *valid* in the probability model M if $M \models \phi$, i.e. for all states $w \in M$, $M, w \models \phi$. ϕ is *valid in a class of probability models* $\mathcal{C}$ if, for each $M \in \mathcal{C}$, $M \models \phi$. ϕ is *valid in a class* $\mathcal{T}$ *of type spaces* if ϕ is valid in all the probability models defined on $\mathcal{T}$.

2.3. Completeness

In this section we will give a complete axiomatization of probability logic with respect to the class of type spaces. Our system is different from that by Heifetz and Mongin in that we don't need the rule (B)[1] but we need another rule which is similar to the induction rule

[1]We will define this rule formally in the next section. It is closely related to the semantic fact that, given a probability measure μ on a state space Ω, if a function f can be written as a sum of characteristic functions in two different ways, then the two ways of calculating the integral with respect to μ will give the same result.

in PDL.

Probability Logic Σ_+

- (A0) propositional calculus
- (A1) $L_0\phi$
- (A2) $L_r\top$
- (A3) $L_r(\phi \wedge \psi) \wedge L_t(\phi \wedge \neg\psi) \rightarrow L_{r+t}\phi$, $r + t \leq 1$
- (A4) $\neg L_r(\phi \wedge \psi) \wedge \neg L_s(\phi \wedge \neg\psi) \rightarrow \neg L_{r+s}\phi$, $r + s \leq 1$
- (A5) $L_r\phi \rightarrow \neg L_s\neg\phi$, $r + s > 1$
- (DIS) If $\vdash \phi \leftrightarrow \psi, \vdash L_r\phi \leftrightarrow L_r\psi$.
- (ARCH): If $\vdash \gamma \rightarrow \neg M_s\phi$ for all $s < r$, then $\vdash \gamma \rightarrow L_r\phi$[2].

Observe that the rule (ARCH) is the only rule that is really about the indexes of the modalities. Since the index set $\mathcal{Q} \cap [0, 1]$ has the *Archimedean property*, i.e., the property of having no infinitely small elements, the fact that the rule has infinitely many premises seems unavoidable. Despite its infinite flavor, it is very easy to *use* the rule (ARCH) to show other propositions in Σ_+. In the next section, we will show as an illustration the admissibility of the rule (B) in our Σ_+. From there, one may find that the rule is finite in nature. Before we show the completeness, we will prove some basic theorems of Σ_+. For simplicity, we will sometimes drop $\vdash$ in front of theorems.

LEMMA 2.3.1. *The following two principles are provable in Σ_+:*

(1) $\neg L_r\bot$ *if* $r > 0$*;*

(2) $\neg M_r\top$ *if* $r < 1$*, which is just dual to the first part.*

PROOF. Since $r > 0$, there is a t such that $t < 1$ such that $t + r > 1$. Reason inside Σ_+:

$L_t\top$ (A1)

$L_t\top \rightarrow \neg L_r\bot$ (A5)

[2] $ARCH$ here stands for the Archimedean property of rational numbers $\mathcal{Q}$.

$\neg L_r \bot$ (A0).

□

THEOREM 2.3.2. *The following principles are provable in Σ_+:*

(1) *If $\phi \to \psi$, then $L_r\phi \to L_r\psi$;*
(2) $L_r\phi \to L_s\phi$ *if $r \geq s$;*
(3) $\neg L_r\phi \to M_r\phi$*;*

PROOF. We reason inside Σ_+.

(1) $\phi \to \psi$ (Assumption)

$\psi \leftrightarrow \phi \wedge \psi$ (A0)

$L_r\psi \leftrightarrow L_r(\phi \wedge \psi)$ (DIS)

$L_r(\phi \wedge \psi) \wedge L_0(\phi \wedge \neg\psi) \to L_r(\phi)$ (A3)

$L_0(\phi \wedge \neg\psi)$ (A1)

$L_r(\phi \wedge \psi) \to L_r(\phi)$ (A0)

$L_r\phi \to L_r\psi$ (A0)

(2) If $r = t$, it is trivially true. Assume that $r > t$.

$\neg L_t(\phi \wedge \phi) \wedge \neg L_{r-t}(\phi \wedge \neg\phi) \to \neg L_r\phi$ (A4)

$\neg L_{r-t}(\phi \wedge \neg\phi)$ (above lemma)

$\neg L_t\phi \to \neg L_r\phi$ (DIS and A0)

$L_r\phi \to L_t\phi$ (A0)

(3) $\neg L_r(\top \wedge \phi) \wedge \neg L_{1-r}(\top \wedge \neg\phi) \to \neg L_1\top$ (A4)

$\neg L_r\phi \wedge \neg L_{1-r}(\neg\phi) \to \neg L_1\top$ (DIS and A0)

$\neg L_r\phi \to M_r\phi$ (A2 and A0)

□

COROLLARY 2.3.3. *The following principles follow immediately from the above theorem:*

(1) *If $\phi \to \psi$, then $M_r\psi \to M_r\phi$;*
(2) $L_r\phi \to L_s\phi$ *if $r \leq s$;*
(3) $\neg M_r\phi \to L_r\phi$*;*

LEMMA 2.3.4. *The following two propositions hold:*

(1) *If* $\vdash \neg(\phi \wedge \psi)$, *then* $\vdash L_r\phi \wedge L_s\psi \rightarrow L_{r+s}(\phi \vee \psi), r + s \leq 1$;

(2) *If* $\vdash \neg(\phi \wedge \psi)$, *then* $\vdash \neg L_r\phi \wedge \neg L_s\psi \rightarrow \neg L_{r+s}(\phi \vee \psi), r + s \leq 1$.

PROOF. Note that, from propositional reasoning, $(\phi \vee \psi) \wedge \phi \leftrightarrow \phi, (\phi \vee \psi) \wedge \psi \leftrightarrow \psi$ and $(\phi \vee \psi) \wedge \psi \rightarrow (\phi \vee \psi) \wedge \neg\phi$. It follows from (**A3**) that

$$L_r((\phi \vee \psi) \wedge \phi) \wedge L_s((\phi \vee \psi) \wedge \neg\phi) \rightarrow L_{r+s}(\phi \vee \psi).$$

From (**DIS**) and (A0), we know

$$L_r\phi \wedge L_s\psi \rightarrow L_r((\phi \vee \psi) \wedge \phi) \wedge L_s((\phi \vee \psi) \wedge \neg\phi)$$

It follows from the above two that:

$$\vdash L_r\phi \wedge L_s\psi \rightarrow L_{r+s}(\phi \vee \psi).$$

The proof of the second part is similar to that of the first part.

□

Now we are prepared to show the completeness. Fix ψ and assume that it is consistent in Σ_+. We need to show that it has a probability model.

DEFINITION 2.3.5. The *depth* $dp(\phi)$ of a formula ϕ is defined inductively:

- $dp(p) := 0$ for propositional letters p;
- $dp(\neg\phi) := dp(\phi)$;
- $dp(\phi_1 \wedge \phi_2) := max\{dp(\phi_1), dp(\phi_2)\}$;
- $dp(L_r\phi) := dp(\phi) + 1$.

Now we define a local language $\mathcal{L}[\psi]$ to be the language satisfying the following conditions:

- The propositional letters in $\mathcal{L}[\psi]$ are those occurring in ψ;
- The indexes of formulas in $\mathcal{L}[\psi]$ are in the finite set $I[\psi]$, which is the set of all rationals in the form of $p/q \in [0, 1]$ where q is the least common multiple of all denominators of the indices appearing in ψ;
- The formulas in $\mathcal{L}[\psi]$ are of depth $\leq dp(\psi)$.

The above q is called the *accuracy* of the language $\mathcal{L}[\psi]$ and $I[\psi]$ the *index set* of the language. Let $\mathcal{L}^+$ be the language obtained by only increasing the depth of $\mathcal{L}[\psi]$ by one[3]. Such a defined language $\mathcal{L}[\psi]$ gives rise to a set of maximal consistent subsets and it is denoted Ω, which will be the carrier set of our canonical model. For any formula $\phi \in \mathcal{L}[\psi]$, define $[\phi] := \{\Delta \in \Omega : \phi \in \Delta\}$. It is easy to see that for any $\Gamma \in \Omega$, since Γ is consistent in $\mathcal{L}^+$, there is a maximal consistent extension $\Gamma^+ \in \mathcal{L}^+$ such that $\Gamma \subseteq \Gamma^+$. For any $\Gamma \in \Omega$ and $\phi \in \mathcal{L}[\psi]$, define

$$\alpha^\Gamma_\phi = max\{\alpha \in I[\psi] : L_\alpha\phi \in \Gamma^+\} \text{ and } \beta^\Gamma_\phi = min\{\beta \in I[\psi] : M_\beta\phi \in \Gamma^+\}$$

Note that $L_{\alpha^\Gamma_\phi}\phi \in \Gamma^+$ and $M_{\beta^\Gamma_\phi}\phi \in \Gamma^+$.

Lemma 2.3.6. *For the above defined α^Γ_ϕ and β^Γ_ϕ, either $\alpha^\Gamma_\phi = \beta^\Gamma_\phi$ or $\beta^\Gamma_\phi = \alpha^\Gamma_\phi + 1/q$.*

Proof. First we show that $\alpha^\Gamma_\phi \leq \beta^\Gamma_\phi$. Suppose that $\alpha^\Gamma_\phi > \beta^\Gamma_\phi$. Since $L_{\alpha^\Gamma_\phi}\phi \in \Gamma^+$, $\neg\beta^\Gamma_\phi\phi \in \Gamma^+$ (by **(A5)**). But this contradicts the fact that $M_{\beta^\Gamma_\phi}\phi \in \Gamma^+$.

Now we show that $\alpha^\Gamma_\phi \leq \beta^\Gamma_\phi + 1/q$. Suppose that $\alpha^\Gamma_\phi > \beta^\Gamma_\phi + 1/q$. It follows that there is a rational $s \in I[\psi]$ such that $s := \alpha^\Gamma_\phi + 1/q$. Since $s > \alpha^\Gamma_\phi$, $L_s\phi \notin \Gamma^+$. This implies that $\neg L_s\phi \in \Gamma^+$. According to one lemma above, $M_s\phi \in \Gamma^+$. But this contradicts the fact $s < \beta^\Gamma_\phi$.

□

Given $\Gamma \in \Omega$ and $\phi \in \mathcal{L}[\psi]$ we define $\mathcal{T}^\Gamma_\phi$ to be either $\{\alpha^\Gamma_\phi\}$ if $\alpha^\Gamma_\phi = \beta^\Gamma_\phi$ or the open interval $(\alpha^\Gamma_\phi, \beta^\Gamma_\phi)$ if $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$. A language $\mathcal{L}_1$ is *more accurate* than the language $\mathcal{L}_2$ if the accuracy q_1 of $\mathcal{L}_1$ is a multiple of the accuracy q_2 of the language $\mathcal{L}_2$. $\mathcal{L}_1$ is *strictly more accurate* than the language $\mathcal{L}_2$ if $q_1 = m \cdot q_2$ for some natural number $m \geq 2$. For any maximal consistent set Δ in the finite language $\mathcal{L}'$ with accuracy q' (and the index set I') and any formula ϕ in $\mathcal{L}$, we define:

$$\alpha^\Delta_\phi = max\{\alpha \in I' : L_\alpha\phi \in \Delta\} \text{ and } \beta^\Delta_\phi = min\{\beta \in I' : M_\beta\phi \in \Delta\}.$$

[3]This increase is needed in the definition of α and β to come.

LEMMA 2.3.7. *Assume that $\mathcal{L}_1$ is more accurate than $\mathcal{L}_2$, Δ is a maximal consistent set of formulas in $\mathcal{L}_2$ and Δ' is a maximal consistent extension of Δ in $\mathcal{L}_1$. Then, for any $\phi \in \mathcal{L}_2$,*

$$\alpha_\phi^\Delta \leq \alpha_\phi^{\Delta'} \leq \beta_\phi^{\Delta'} \leq \beta_\phi^\Delta$$

To show the truth lemma, we want to define a probability measure $T(\Gamma)$ on the subsets of Ω with the property that

$$(P) : \forall \phi \in \mathcal{L}[\psi], T(\Gamma)([\phi]) \in \mathcal{T}_\phi^\Gamma$$

In order to achieve (P), Rockfellar's Lemma is used in [17]. In the following, we show that, with addition of the rule (ARCH), we don't need this lemma. This also implies that the rule (B) is unnecessary in the complete axiomatization. Fagin and Halpern's reasoning about linear inequalities are not necessary either. However, we don't know how to achieve the property (P) directly for the above Γ and Ω in our system. But instead we will show the property in a more accurate language than $\mathcal{L}[\psi]$. Our canonical model is constructed in the local language $\mathcal{L}[\psi]$ but the definition of the mapping T is obtained in the "infinitely accurate language" $\mathcal{L}$, which is also the formal langauge of the logic. Our proof of the Truth Lemma is quite unconventional. For any maximal consistent extension Γ^∞ in $\mathcal{L}$ of Γ^+ such that $\Gamma^+ \subseteq \Gamma^\infty$, we define, for any $\phi \in \mathcal{L}[\psi]$,

$$\alpha_\phi^\infty = max\{\alpha \in \mathcal{Q} : L_\alpha\phi \in \Gamma^\infty\} \text{ and } \beta_\phi^\infty = min\{\beta \in \mathcal{Q} : M_\beta\phi \in \Gamma^\infty\}$$

Both α_ϕ^∞ and β_ϕ^∞ might be irrational.

LEMMA 2.3.8. *For any maximal extension Γ_ϕ^∞, $\alpha_\phi^\infty = \beta_\phi^\infty$.*

PROOF. Suppose that $\alpha_\phi^\infty < \beta_\phi^\infty$. This implies that there is a rational r between: $\alpha_\phi^\infty < r < \beta_\phi^\infty$. Therefore, $L_r\phi \notin \Gamma^\infty$. Since Γ_ϕ^∞ is maximal consistent, $\neg L_r\phi \in \Gamma_\phi^\infty$. It follows that $M_r\phi \in \Gamma^\infty$. But this is impossible because β_ϕ^∞ is the greatest lower bound.

Suppose that $\alpha_\phi^\infty > \beta_\phi^\infty$. Then there are two rationals r_1 and r_2 such that $\alpha_\phi^\infty > r_1 > r_2 > \beta_\phi^\infty$. Since $\alpha_\phi^\infty > r_1$, $L_{r_1}\phi \in \Gamma^\infty$. But this implies that $\neg M_{r_2}\phi \in \Gamma^\infty$ by **(A5)**. On other hand, we have that $M_{r_2}\phi \in \Gamma^\infty$ because $r_2 > \beta^\infty_\phi$. This is a contradiction. So we have shown that $\alpha^\infty_\phi = \beta^\infty_\phi$.

$\square$

LEMMA 2.3.9. (1) $2^\Omega = \{[\phi] : \phi \in \Phi(\psi)\}$;

(2) *For any* $\phi_1, \phi_2 \in \mathcal{L}[\psi]$, $\vdash_{\Sigma_+} \phi_1 \to \phi_2$ *iff* $[\phi_1] \subseteq [\phi_2]$.

PROOF. We need only propositional reasoning to show the first part. The second is just a standard result in modal logic. Note that each formula is actually an equivalence class under the Σ_+-derived equivalence relation.

$\square$

We define $T(\Gamma) : 2^\Omega \to [0,1]$ as follows: $T(\Gamma)([\phi]) = \alpha^\infty_\phi$.

THEOREM 2.3.10. *Such a defined function* $T : 2^\Omega \to [0,1]$ *is well-defined.*

PROOF. Assume that $[\phi_1] = [\phi_2]$. According to the second part of the above lemma, $\vdash \phi_1 \leftrightarrow \phi_2$ and hence $\vdash (L_r\phi_1 \leftrightarrow L_r\phi_2) \wedge (M_s\phi_1 \leftrightarrow M_s\phi_2)$. So $\alpha^\infty_{\phi_1} = \beta^\infty_{\phi_2}$. In other words, $T(\Gamma)([\phi_1]) = T(\Gamma)([\phi_2])$.

$\square$

From the first part of Lemma 2.3.9, it follows that T is total. It is easy to see that $T(\Gamma)(\Omega) = T(\Gamma)([\top]) = 1$ since $L_1\top \in \Gamma^\infty$. Next we show the finite additivity.

LEMMA 2.3.11. *For* $A, B \in 2^\Omega$, *if* $A \cap B = \emptyset$, *then* $T(\Gamma)(A) + T(\Gamma)(B) = T(\Gamma)(A \cup B)$.

PROOF. It is easy to see that there are formulas $\phi_1, \phi_2 \in \mathcal{L}[\psi]$ such that $A = [\phi_1]$, $B = [\phi_2]$ and $\vdash \phi_1 \to \neg\phi_2$. Let α_1, α_2 and α_+ denote $T(\Gamma)([\phi_1])$, $T(\Gamma)([\phi_2])$ and $T(\Gamma)([\phi_1 \vee \phi_2])$, respectively. So we only need to show that $\alpha_1 + \alpha_2 = \alpha_+$.

Suppose that $\alpha_1+\alpha_2 < \alpha_+$. Then there are $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that $(\alpha_1+\epsilon_1)+(\alpha_2+\epsilon_2) < \alpha_+$, $\alpha_1 + \epsilon_1 \in \mathcal{Q}$ and $\alpha_2 + \epsilon_2 \in \mathcal{Q}$. Let $\alpha'_1 := \alpha_1 + \epsilon_1$ and $\alpha'_2 := \alpha_2 + \epsilon_2$. It follows that $L_{\alpha'_1}\phi_1 \notin \Gamma^\infty$ and hence $\neg L_{\alpha'_1}\phi_1 \in \Gamma^\infty$. Similarly, $\neg L_{\alpha'_2}\phi_2 \in \Gamma^\infty$. By **(A4)** (actually by one lemma following from it), we know that $\neg L_{\alpha'_1+\alpha'_2}(\phi_1 \vee \phi_2) \in \Gamma^\infty$. But this is impossible because $\alpha'_1+\alpha'_2 < \alpha_+$ and α_+ is the least upper bound such that $L_\beta(\phi_1 \vee \phi_2) \in \Gamma^\infty$.

The following argument is dual to the above one. Suppose that $\alpha_1 + \alpha_2 > \alpha_+$. Then there are two $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that $(\alpha_1 - \epsilon_1) + (\alpha_2 - \epsilon_2) > \alpha_+$, $\alpha_1 - \epsilon_1 \in \mathcal{Q}$ and $\alpha_2 - \epsilon_2 \in \mathcal{Q}$. Let $\alpha''_1 := \alpha_1 - \epsilon_1$ and $\alpha''_2 := \alpha_2 - \epsilon_2$. It follows that $M_{\alpha''_1}\phi_1 \notin \Gamma^\infty$ and hence $L_{\alpha''_1}\phi_1 \in \Gamma^\infty$. Similarly, $L_{\alpha''_2}\phi_2 \in \Gamma^\infty$. By **(A3)**, we know that $L_{\alpha''_1+\alpha''_2}(\phi_1 \vee \phi_2) \in \Gamma^\infty$. But this is impossible because $\alpha''_1 + \alpha''_2 > \alpha_+$ and α_+ is the smallest upper bound such that $L_\alpha(\phi_1 \vee \phi_2) \in \Gamma^\infty$.

□

Since $\Gamma^+ \subseteq \Gamma^\infty$, $\alpha^\Gamma_\phi \leq T(\Gamma)([\phi]) \leq \beta^\Gamma_\phi$ for any $\phi \in \mathcal{L}[\psi]$. But, as we will see in the following proof of the Truth Lemma, we have to eliminate the possibility that $T(\Gamma)([\phi]) = \beta^\Gamma_\phi$ when $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$. This is also the reason why we need the rule (ARCH).

LEMMA 2.3.12. *In $\mathcal{L}[\psi]$, given any $\Gamma \in \Omega$, there is a strictly more accurate language $\mathcal{L}_\Gamma$ and a maximal consistent extension Δ in $\mathcal{L}_\Gamma$ of Γ^+ satisfying the following condition:*

(E): for all $\phi \in \mathcal{L}[\psi]$ such that $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$, either $\alpha^\Delta_\phi = \beta^\Delta_\phi = \beta^\Gamma_\phi$ or $\beta^\Delta_\phi < \beta^\Gamma_\phi$.

PROOF. We claim that, for any ϕ such that $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$, there is a rational r strictly between α^Γ_ϕ and β^Γ_ϕ such that $\Gamma^+ \cup \{M_r\phi\}$ is consistent. Suppose not. Then, for all $r < \beta^\Gamma_\phi$, $\vdash \bigwedge \Gamma^+ \to \neg M_r\phi$. It follows from the rule **(ARCH)** that $\vdash \bigwedge \Gamma^+ \to L_{\beta^\Gamma_\phi}\phi$. So, $L_{\beta^\Gamma_\phi}\phi \in \Gamma^+$. But, this contradicts the fact that α^Γ_ϕ ($< \beta^\Gamma_\phi$) is the largest such number in the index set $I[\psi]$. So, there is a $r < \beta^\Gamma_\phi$ such that $\Gamma^+ \cup \{M_r\phi\}$ is consistent.

Now we prove the main lemma. Our proof strategy is like "sandwich-making". First we enumerate all the formulas in $\mathcal{L}[\psi]$ such that $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$: $\phi_1, \phi_2, \cdots, \phi_n$. For ϕ_1, we know that there is a rational r_1 between $\alpha^\Gamma_{\phi_1}$ and $\beta^\Gamma_{\phi_1}$ such that $\Gamma^+ \cup \{M_{r_1}\phi_1\}$ is consistent. Let $\mathcal{L}^1$ be the language obtained from $\mathcal{L}^+$ just by increasing the accuracy of the language in the following way: if q is the accuracy of $\mathcal{L}^+$, i.e. the least common multiple of the denominators of all the indexes in ψ, then the accuracy of the language $\mathcal{L}^1$ is the integer q_1 which is the least common multiple of q and the denominator of r_1. Let I^1 be the index set of the langauge $\mathcal{L}^1$. It is easy to see that there is a maximal consistent extension $\Gamma^1 \in \mathcal{L}^1$ such that $\Gamma^+ \cup \{M_{r_1}\phi_1\} \subseteq \Gamma^1$.

Next we consider ϕ_2. Define:

$\alpha_{\phi_i}^{\Gamma^1} := max\{r \in I^1 : L_r\phi_i \in \Gamma^1\}$ and $\beta_{\phi_i}^{\Gamma^1} := min\{r \in I^1 : M_r\phi_i \in \Gamma^1\}$ for each i: $1 \leq i \leq n$.

It is easy to see that $\beta_{\phi_1}^{\Gamma^1} \leq r_1 < \beta_{\phi_1}^{\Gamma}$ and $\beta_{\phi_2}^{\Gamma^1} \leq \beta_{\phi_2}^{\Gamma}$ because $\Gamma^+ \subseteq \Gamma^1$. If $\beta_{\phi_2}^{\Gamma^1} < \beta_{\phi_2}^{\Gamma}$, then we are done. If $\beta_{\phi_2}^{\Gamma^1} = \beta_{\phi_2}^{\Gamma}$ and $\alpha_{\phi_2}^{\Gamma^1} = \beta_{\phi_2}^{\Gamma^1}$, we don't do anything either. Now we consider the possibility that $\beta_{\phi_2}^{\Gamma^1} = \beta_{\phi_2}^{\Gamma}$ and $\alpha_{\phi_2}^{\Gamma^1} < \beta_{\phi_2}^{\Gamma^1}$. As we already said above, this is the case where we need the rule **(ARCH)**. By appealing to the above claim, we know that there is r_2 such that $\alpha_{\phi_2}^{\Gamma^1} < r_2 < \beta_{\phi_2}^{\Gamma^1}$ and $\Gamma^1 \cup \{M_{r_2}\phi_2\}$ are consistent. Similarly, we can define the expanded language $\mathcal{L}^2$. Therefore, there is a maximal consistent extension $\Gamma^2 \in \mathcal{L}^2$ such that $\Gamma^1 \cup \{M_{r_2}\phi_2\} \subseteq \Gamma^2$.

Similarly, we can define $\alpha_{\phi_i}^{\Gamma^2}$ and $\beta_{\phi_i}^{\Gamma^2}$ for all ϕ_i. Then $\beta_{\phi_2}^{\Gamma^2} \leq r_2 < \beta_{\phi_2}^{\Gamma^1} = \beta_{\phi_2}^{\Gamma}$. Now we move to deal with ϕ_3. If, either $\beta_{\phi_3}^{\Gamma^2} < \beta_{\phi_3}^{\Gamma}$, or $\beta_{\phi_3}^{\Gamma^2} = \beta_{\phi_3}^{\Gamma}$ and $\alpha_{\phi_3}^{\Gamma^2} = \beta_{\phi_3}^{\Gamma^2}$, then we don't do anything. If $\beta_{\phi_3}^{\Gamma^2} = \beta_{\phi_3}^{\Gamma}$ and $\alpha_{\phi_3}^{\Gamma^2} < \beta_{\phi_3}^{\Gamma^2}$, then there is a maximal consistent extension Γ^3 in the corresponding langauge $\mathcal{L}^3$ such that $\Gamma^2 \cup \{M_{r_3}\phi_3\} \subseteq \Gamma^3$ for some $r_3 \in (\alpha_{\phi_3}^{\Gamma^2}, \beta_{\phi_3}^{\Gamma^2})$. We repeat this process, then we will get a chain of Γ^i's in the expanded language L^i's:

$$\Gamma \subseteq \Gamma^+ \subseteq \Gamma^1 \subseteq \Gamma^2 \subseteq \cdots \subseteq \Gamma^n$$

Observe that, in Γ^n, for all ϕ_i, either $\alpha_{\phi_i}^{\Gamma^n} = \beta_{\phi_i}^{\Gamma^n} = \beta_{\phi_i}^{\Gamma}$ or $\beta_{\phi_i}^{\Gamma^n} < \beta_{\phi_i}^{\Gamma}$. Therefore, given any $\Gamma \in \mathcal{L}[\psi]$, there is a maximal consistent extension Γ^n in the language $\mathcal{L}^n$ such that $\Gamma^n \supseteq \Gamma^+$ and, for any $\phi \in \mathcal{L}[\psi]$, either $\alpha_{\phi}^{\Gamma^n} = \beta_{\phi}^{\Gamma^n} = \beta_{\phi}^{\Gamma}$ or $\alpha_{\phi}^{\Gamma^n} \leq \beta_{\phi}^{\Gamma^n} < \beta_{\phi}^{\Gamma}$. Note that α_{ϕ}^{Γ} and β_{ϕ}^{Γ} are defined in the language $\mathcal{L}^+$, $\alpha_{\phi}^{\Gamma^n}$, $\beta_{\phi}^{\Gamma^n}$ are defined in the language $\mathcal{L}^n$. So $\mathcal{L}^n$ and Γ^n are what we want to get.

□

Now we enumerate all the maximal consistent sets in Ω:

$$\Gamma_1, \Gamma_2, \cdots, \Gamma_N.$$

From the above lemma, it follows that, for each Γ_i, there is a strictly more accurate language $\mathcal{L}_i$ and a maximal consistent extension Δ_i in the language $\mathcal{L}_i$ that satisfy the property (E). Assume that $\mathcal{L}^c$ is a language that is strictly more accurate than all the languages $\mathcal{L}_i$. The language $\mathcal{L}^c$ gives rise to a set Ω^c of maximal consistent sets of formulas in $\mathcal{L}^c$. For each Δ_i, there is a maximal consistent extension $\Delta_i^c \in \mathcal{L}^c$ such that $\Delta_i \subseteq \Delta_i^c$. Since property (E) is preserved for more accurate languages, it holds in all these Δ_i^c.

Further, for each $\Delta_i^c \in \mathcal{L}^c$, there is a maximal consistent extension $\Gamma_i^\infty \in \mathcal{L}$ such that $\Gamma_i^\infty \supseteq \Delta_i^c$. For each $\phi \in \mathcal{L}[\psi]$, let $\alpha_\phi^{\Gamma_i^\infty}$ and $\beta_\phi^{\Gamma_i^\infty}$ be just defined as above. It follows that $\alpha_\phi^{\Gamma_i^\infty} = \beta_\phi^{\Gamma_i^\infty}$. From the above lemmas, we know that such a Γ_i^∞ defines a probability measure $T(\Gamma_i) : 2^\Omega \to [0,1]$: $T(\Gamma_i)([\phi]) = \alpha_\phi^{\Gamma_i^\infty}$.

LEMMA 2.3.13. *For such $T(\Gamma_i)$, we can show that, if $L_r\phi \in \mathcal{L}[\psi]$, then*

$$\Gamma_i \models L_r\phi \text{ iff } L_r\phi \in \Delta_i^c.$$

PROOF. Assume that $\Gamma_i \models L_r\phi$. Then, $T(\Gamma_i)([\phi]) \geq r$. There are several cases that we have to consider. If $\alpha_\phi^{\Gamma_i} = \beta_\phi^{\Gamma_i}$, it is easy to see that $T(\Gamma_i)([\phi]) = \alpha_\phi^{\Gamma_i}$ and hence $\alpha_\phi^{\Gamma_i} \geq r$. From the definition of $\alpha_\phi^{\Gamma_i}$, it follows that $L_r\phi \in \Gamma_i^+ \subseteq \Delta_i^c$. If $\alpha_\phi^{\Gamma_i} < \alpha_\phi^{\Gamma_i}$ and $\alpha_\phi^{\Delta_i^c} = \beta_\phi^{\Delta_i^c} = \beta_\phi^{\Gamma_i}$, then $r \leq T(\Gamma_i)([\phi]) = \alpha_\phi^{\Delta_i^c} = \beta_\phi^{\Gamma_i}$. Obviously $L_r\phi \in \Delta_i^c$. The last case that we have to consider is when $\alpha_\phi^{\Gamma_i} < \beta_\phi^{\Gamma_i}$ and $\beta_\phi^{\Delta_i^c} < \beta_\phi^{\Gamma_i}$. It means that $T(\Gamma_i)([\phi]) < \alpha_\phi^{\Gamma_i} + 1/q$ where q is the accuracy of the local language $\mathcal{L}[\psi]$. Since $r \in I[\psi]$, $\alpha_\phi^{\Gamma_i} \geq r$. So $L_r\phi \in \Gamma_i \subseteq \Delta_i^c$.

Assume that $L_r\phi \in \Delta_i^c$. It follows that $r \leq \alpha_\phi^{\Delta_i^c} \leq \alpha_\phi^{\Gamma_i^\infty} = T(\Gamma_i)([\phi])$. So $\Gamma_i \models L_r\phi$.

□

LEMMA 2.3.14. *(Truth Lemma) For any $\phi \in \mathcal{L}[\psi]$,*

$$\Gamma_i \models \phi \textit{ iff } \phi \in \Delta_i^c \textit{ iff } \phi \in \Gamma_i.$$

PROOF. Here we only note that every time when we increase the accuracy of the working languages, the set of propositional variables remains the same. We can prove this by induction on the complexity of the formula ϕ. The nontrivial case has been proved in the above lemma. The second equivalence follows from the fact that $\phi \in \mathcal{L}[\psi]$ and Γ_i is the $\mathcal{L}[\psi]$-fragment of Δ_i^c.

□

THEOREM 2.3.15. *(Completeness)* $\models \psi$ *iff* $\vdash_{\Sigma^+} \psi$.

PROOF. Assume that ψ is consistent. Then it is contained in some maximal consistent extension $\Gamma_i \in \mathcal{L}[\psi]$. By the Truth Lemma, we know that $\Gamma_i \models \psi$. So ψ is satisfiable. We have shown the completeness.

□

Note that this is a weak completeness result. Strong completeness does not hold for Σ^+ because of the lack of compactness.

THEOREM 2.3.16. *(noncompactness)* Σ_+ *is not compact. That is to say, there is some set* Λ *of formulas in* $\mathcal{L}$*, although* Λ *is finitely satisfiable, it is not satisfiable.*

PROOF. It is easy to see that the set of formulas:

$$\{\neg L_{1/2}p\} \cup \{L_{1/2-(1/2)^{n+2}}p : n \in \mathcal{N}\}$$

is finitely satisfiable but not satisfiable in the class of probability models.

□

2.4. Equivalence to Heifetz & Mongin's System

Semantically, it is easy to see that our system is equivalent to that in [17] since they are sound and complete with respect to the same class of type spaces. In this section, we will give a *constructive* proof of their equivalence. More importantly, we will give the proof as an illustration how to use the induction rule (ARCH).

THEOREM 2.4.1. *The following forms of (ARCH) are equivalent:*

- *If* $\vdash \gamma \to \neg M_s\phi$ *for all* $s < r$, *then* $\vdash \gamma \to L_r\phi$;
- *If* $\vdash \gamma \to \L_s\phi$ *for all* $s < r$, *then* $\vdash \gamma \to L_r\phi$;
- *If* $\vdash \gamma \to M_s\phi$ *for all* $s > r$, *then* $\vdash \gamma \to M_r\psi$;

Here we show that our system Σ_+ subsumes the system in [17]. We only need to show that the rule (B) there is admissible in our system. But, before we do that, we will prove that the following axiom can be obtained in our system. It is actually the so called ($A13$) in [17].

LEMMA 2.4.2. $\vdash_{\Sigma_+} M_r(\phi \wedge \neg\psi) \wedge M_s(\phi \wedge \psi) \to M_{r+s}\phi$, *if* $r + s \leq 1$.

PROOF. By the rule (ARCH)[4], we know that it suffices to show that

$$(*) :\vdash M_r(\phi \wedge \neg\psi) \wedge M_s(\phi \wedge \psi) \to M_t\phi \text{ for all } 1 \geq t > r + s.$$

Note that $\vdash M_r(\phi \wedge \neg\psi) \to \neg L_{r+\varepsilon/4}(\phi \wedge \neg\psi)$ and $\vdash M_r(\phi \wedge \psi) \to \neg L_{s+\varepsilon/4}(\phi \wedge \psi)$ where ε is a sufficiently small rational number such that $t - r - s > \epsilon/2$. Moreover, since $\vdash \neg L_{r+\varepsilon/4}(\phi \wedge \neg\psi) \wedge \neg L_{s+\varepsilon/4}(\phi \wedge \psi) \to \neg L_{r+s+\varepsilon/2}\phi$ and $r + s + \epsilon < t$, $(*)$ follows immediately. □

The rule (B) in [17] is the defining rule in their system. This rule conveys the intuition that "if a function f can be written as a sum of characteristic functions in two different ways, then the two ways of calculating the integral with respect to μ will give the same result." Now we formalize this idea. Let $(\phi_1, \cdots, \phi_m)$ be a finite sequence of formulas and $\phi^{(k)}$ denote

$$\bigvee_{1 \leq i_1 < \cdots < i_k \leq n} (\phi_{i_1} \wedge \cdots \wedge \phi_{i_k}).$$

Then the above intuition that the sum of characteristic functions of the finite sequence $(\phi_1, \cdots, \phi_m)$ and that of $(\psi_1, \cdots, \psi_m)$ are equal to each other can be formalized as

[4]The principle that we use here is not the same as that in the system Σ_+ but the dual form, which follows immediately from the one in the system by (DEF M).

$$\bigwedge_{k=1}^{m}(\phi^{(k)} \leftrightarrow \psi^{(k)})$$

This formula is denoted as $(\phi_1, \cdots, \phi_m) \leftrightarrow (\psi_1, \cdots, \phi_m)$. The inference rule (B) can be stated as follows:

$$\text{If } ((\phi_1, \cdots, \phi_m) \leftrightarrow (\psi_1, \cdots, \psi_m)), \text{ then}$$
$$((\bigwedge_{i=1}^{m} L_{r_i}\phi_i) \wedge (\bigwedge_{j=2}^{m} M_{s_j}\psi_j)) \to L_{(r_1+\cdots+r_m)-(s_2+\cdots+s_m)}\psi_1$$

for $(r_1 + \cdots + r_m) - (s_2 + \cdots + s_m) \in [0, 1]$. Now we show that the rule (B) is admissible in our system. Before we give a formal proof, we will develop an intuition how to apply the rule (ARCH). We will not show the general form of (B). But instead we prove as an illustration the case $m = 2$:

$$\text{If } (\phi_1, \phi_2) \leftrightarrow (\psi_1, \psi_2), \text{ then } L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \to L_{r_1+r_2-s_2}\psi_1$$

for $r_1 + r_2 - s_2 \in [0, 1]$. Assume that $(\phi_1, \phi_2) \leftrightarrow (\psi_1, \psi_2)$. According to the rule (ARCH), it suffices to show:

$$L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \to L_{r_1+r_2-s_2-\epsilon}\psi_1 \text{ for any sufficiently small rational } \epsilon.$$

In the following we will use a backtracking deduction from the conclusion to the beginning to see what axioms we need in the expected proof. Since the ϵ gives us a lot of flexibility, we will just use the strict inequality symbols like $<$ and $>$. For a clear presentation, we will write, say, $\phi_1 \wedge \phi_2 > r$ instead of the formal $\neg M_r(\phi_1 \wedge \phi_2)$.

(1) In order to show that $\psi_1 > r_1 + r_2 - s_2 - \epsilon$, we need to show that $\phi_1 \wedge \psi_2 > t$ and $\psi_1 \wedge \neg\psi_2 > r_1 + r_2 - s_2 - t - \varepsilon$ where t is a parameter;

(2) In order to show that $\psi_1 \wedge \neg\psi_2 > r_1 + r_2 - s_2 - t - \varepsilon$, we need to show that $\psi_1 \vee \psi_2 > r_1 + r_2 - t - \varepsilon$ and $\psi_2 < s_2$;

(3) Note that $\vdash \psi_1 \vee \psi_2 \leftrightarrow \phi_1 \vee \phi_2$. In order to show that $\phi_1 \vee \phi_2 > r_1 + r_2 - t - \varepsilon$, we need to show that $\phi_1 > r_1$ and $\phi_2 \wedge \neg\phi_1 > r_2 - t - \varepsilon$;

(4) Note that $\vdash \phi_1 \wedge \phi_2 \leftrightarrow \psi_1 \wedge \psi_2$. In order to show that $\phi_2 \wedge \neg\phi_1 > r_2 - t - \varepsilon$, we need to show that $\phi_2 > r_2$ and $\phi_1 \wedge \phi_2 < t + \varepsilon$.

As you can see in the above reasoning, $\psi_1 \wedge \psi_2 (\leftrightarrow \phi_1 \wedge \phi_2)$ is the parameter.

LEMMA 2.4.3. *If* $\vdash \phi_1 \to \neg\phi_2$ *and* $r + s \leq 1$*, then*

(1) $\vdash M_r(\phi_1) \wedge M_s(\phi_2) \to M_{r+s}(\phi_1 \vee \phi_2)$;

(2) $\vdash \neg M_r(\phi_1) \wedge \neg M_s(\phi_2) \to \neg M_{r+s}(\phi_1 \vee \phi_2)$;

(3) $\vdash M_r(\phi_1) \wedge \neg L_s(\phi_2) \to \neg\neg L_{r+s}(\phi_1 \vee \phi_2)$;

(4) $\vdash L_r(\phi_1) \wedge \neg M_s(\phi_2) \to \neg M_{r+s}(\phi_1 \vee \phi_2)$;

DEFINITION 2.4.4. A proof $\mathcal{E}$ of $L_r\phi$ ($M_r\phi$) in Σ_+ is *stable* if $\mathcal{E}(r/(r-\varepsilon))$ ($\mathcal{E}(r/(r+\varepsilon))$) is a proof of $L_{r-\varepsilon}$ ($M_{r+\varepsilon}$) in Σ_+ for any sufficiently small ε, where $\mathcal{E}(r/(r-\varepsilon))$ ($\mathcal{E}(r/(r+\varepsilon))$) is obtained by substituting $r-\varepsilon$ ($r+\varepsilon$) for r occurring in the proof sequence of $L_r\phi$ ($M_r\phi$).

The rule (ARCH) just tells us that all of our proofs in Σ_+ are stable.

THEOREM 2.4.5. *If* $\vdash (\phi_1, \phi_2) \leftrightarrow (\psi_1, \psi_2)$, *then* $\vdash L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \to L_{r_1+r_2-s_2-\epsilon}\psi_1$ *for any arbitrarily small* ϵ.

PROOF. Observe that the assumption says that we can use $\phi_1 \wedge \phi_2$ ($\phi_1 \vee \phi_2$) and $\psi_1 \wedge \psi_2(\psi_1 \vee \psi_2)$ interchangeably. First we just formalize the above intuition. We reason in Σ_+: given any ε,

(1) $L_{r_2}\phi_2 \wedge (M_t(\phi_1 \wedge \phi_2) \wedge L_{t-\delta}(\phi_1 \wedge \phi_2)) \to L_{r_2-t}(\phi_2 \wedge \neg\phi_1)$;

(2) $L_{r_1}\phi_1 \wedge \neg M_{r_2-t}(\phi_2 \wedge \neg\phi_1) \to \neg M_{r_1+r_2-t}(\phi_1 \vee \phi_2)$;

(3) $L_{r_1+r_2-t}(\psi_1 \vee \psi_2) \wedge M_{s_2}\psi_2 \to L_{r_1+r_2-s_2-t}(\psi_1 \wedge \neg\psi_2)$;

(4) $(L_{t-\delta}(\psi_1 \wedge \psi_2) \wedge M_t(\phi_1 \wedge \phi_2)) \wedge L_{r_1+r_2-s_2-t-\varepsilon}(\psi_1 \wedge \neg\psi_2) \to L_{r_1+r_2-s_2-\varepsilon}\psi_1$ for any $\delta < \epsilon$.

If we sum up all these theorems and then apply propositional calculus, we will get the following:

$$L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge ((M_t(\phi_1 \wedge \phi_2) \wedge L_{t-\delta}(\phi_1 \wedge \phi_2))) \to L_{r_1+r_2-s_2-\varepsilon}\psi_1.$$

But this is equivalent to:

$$(R) : L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \to (M_t(\phi_1 \wedge \phi_2) \to M_{t-\delta}(\phi_1 \wedge \phi_2)).$$

Since $L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \to M_{r_1+r_2-s_2-\varepsilon}(\phi_1 \wedge \phi_2)$,

$$L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \to M_{r_1+r_2-s_2-\varepsilon-\delta}(\phi_1 \wedge \phi_2)$$

If we apply the induction axiom (R) $n+1$ times, we will get:

$$L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \rightarrow M_{r_1+r_2-s_2-\varepsilon-n\cdot\delta}(\phi_1 \wedge \phi_2)$$

Eventually, we will get closer and closer to zero by the Archimedean property of $\mathcal{Q}$, since δ is an arbitrary small rational $< \varepsilon$. By the rule (ARCH), we know that the above proof is stable. This means that we reach a formula that will lead to a contradiction:

$$L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \rightarrow M_0(\phi_1 \wedge \phi_2).$$

Now we see the intuition how we can get a contradiction. The premise of the above formula tells us that:

(1) the probability of ϕ_1 is bigger than or equal to r_1;
(2) the probability of ϕ_2 is bigger than or equal to r_2;
(3) the probability of ψ_1 is smaller than $r_1 + r_2 - s_2 - \varepsilon$;
(4) the probability of ψ_2 is smaller than or equal to s_2.

Since the probability of $\phi_1 \wedge \phi_2$ (and $\psi_1 \wedge \psi_2$) is zero, the first two premises will get that the probability of $\phi_1 \vee \phi_2$ is bigger than or equal to $r_1 + r_2$ while the last two premises will tell us that the probability is smaller than $r_1 + r_2 - s_2 - \varepsilon$, which is contradiction. Next we formalize the above intuition. ϕ_0 denote $L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1$.

Since $\vdash \phi_0 \rightarrow M_0(\phi_1 \wedge \phi_2)$,

$$(C) :\vdash \phi_0 \rightarrow L_{r_1+r_2}(\phi_1 \vee \phi_2) \text{ and } \vdash \phi_0 \rightarrow M_{r_1+r_2-\varepsilon}(\psi_1 \vee \psi_2)$$

by the first two conjuncts and the last two conjuncts in ϕ_0, respectively. According to the assumption, $\vdash (\phi_1 \vee \phi_2) \leftrightarrow (\psi_1 \vee \psi_2)$. So (C) implies that

$$\vdash \phi_0 \rightarrow L_{r_1+r_2}(\phi_1 \vee \phi_2) \wedge M_{r_1+r_2-\varepsilon}(\psi_1 \vee \psi_2).$$

But, since $r_1 + r_2 > r_1 + r_2 - \varepsilon$, $\vdash L_{r_1+r_2}(\phi_1 \vee \phi_2) \rightarrow \neg M_{r_1+r_2-\varepsilon}(\psi_1 \vee \psi_2)$. Obviously, this leads to a contradiction. So we conclude that $\vdash \phi_0 \rightarrow \perp$, i.e. $\vdash L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \wedge \neg L_{r_1+r_2-s_2-\varepsilon}\psi_1 \rightarrow \perp$. Obviously, this is equivalent to:

$$\vdash L_{r_1}\phi_1 \wedge L_{r_2}\phi_2 \wedge M_{s_2}\psi_2 \rightarrow L_{r_1+r_2-s_2-\varepsilon}\psi_1.$$

□

2.5. $L_r(M_r)$ is *the* Modality for Inner (Outer) Measures

In this section, we will give a complete axiomatization of the general case: $[[\phi]](=\{w \in X : w \models \phi\}$ is not necessarily a measurable set. For the completeness of the presentation, we will give the definition of the semantics although it is very similar to the measurable case. We call the logic for the general case Σ_L. The only difference in syntax from Σ_+ is that we only need one modality L_r for rational $r \in [0,1]$. This implies that $L_{1-r}\neg\phi$ no longer means "at most". *A probability model with an inner measure* is defined as follows:

$$M = \langle \Omega, \mathcal{A}, T, \nu \rangle$$

where Ω is a non-empty set; $\mathcal{A}$ is a σ-algebra of subsets of Ω; T is a mapping from Ω to the space $\Delta(\Omega, \mathcal{A})$ of inner probability measures on Ω, which is endowed with the σ-field generated by the sets $\{\mu_* : \mu_*(E) \geq \alpha\}$ for all $E \in \mathcal{A}$ and rational $r \in [0,1]$, and ν is a mapping from the set P of propositional letters to 2^{Ω}. This also means that $\nu(p)$ is not necessarily measurable. The semantical clause for the nontrivial modality L_r is correspondingly defined as:

$$M, w \models L_r\phi \text{ iff } T(w)([\phi]) \geq r.$$

The axiomatization is quite expected after our axiomatization for measurable case. We only need the *almost positive*[5] half of the L_r half system of Σ_+.

Probability Logic Σ_L

- (A0) propositional calculus
- (A1) $L_0\phi$
- (A2) $L_r\top$
- (A3) $L_r(\phi \wedge \psi) \wedge L_s(\phi \wedge \neg\psi) \to L_{r+s}\phi$, $r + s \leq 1$
- (A5) $L_r\phi \to \neg L_s\neg\phi$ if $r + s > 1$;
- (A7) $L_r\phi \to L_s\phi$, if $r \geq s$;
- (DIS) If $\vdash \phi \to \psi, \vdash L_r\phi \to L_r\psi$.

---

[5]It is in the sense that we don't use any explicit negation in our axiomatization.

- (ARCH): If $\vdash \gamma \rightarrow L_s\phi$ for all $s < r$, then $\vdash \gamma \rightarrow L_r\psi$.

The validity of all theorems of Σ_L is obvious. Note that the axiom (A4) is not valid here. Now we show the completeness. Assume that ψ is consistent. We need to show that ψ is satisfiable in the above inner probability models. Here we will still use the notations from Section 3. Similarly, we can define the index set $I[\psi]$ and the local language $\mathcal{L}[\psi]$. This language gives rise to a set Ω of maximal consistent sets in $\mathcal{L}[\psi]$. For any Γ and ϕ, α^Γ_ϕ (defined through a maximal extension Γ^+ in the expanded language $\mathcal{L}^+$ and uses (**A7**) implicitly) can be defined similarly but we cannot define β^Γ_ϕ because the modality M_r is not in our formal language. In order to prove truth lemma, we have to eliminate the possibility that the probability measure of $[\phi]$ at Γ $T(\Gamma)([\phi]) = \alpha^\Gamma_\phi + 1/q$, where q is the accuracy of the language $\mathcal{L}[\psi]$. This is the place where we need the rule (ARCH).

LEMMA 2.5.1. *For any $\Gamma \in \Omega$ and $\phi \in \mathcal{L}[\psi]$, there is a maximal consistent extension Γ^∞ in the language $\mathcal{L}$, which is the formal language of the logic, such that $\Gamma^+ \subseteq \Gamma^\infty$ and $\alpha^\Gamma_\phi \leq \alpha^\infty_\phi < \alpha^\Gamma_\phi + 1/q$ where*

$$\alpha^\infty_\phi = max\{r \in \mathcal{Q} : L_r\phi \in \Gamma^\infty\}.$$

PROOF. We claim that, for any $\Gamma \in \Omega$ and ϕ, $\Gamma \cup \{\neg L_r\phi\}$ is consistent for some rational $r < \alpha^\Gamma_\phi + 1/q$. Suppose not. Then $\vdash \bigwedge \Gamma^+ \rightarrow L_r\phi$ for all $r < \alpha^\Gamma_\phi + 1/q$. By the rule (**ARCH**), we know that $\vdash \bigwedge \Gamma \rightarrow L_{\alpha^\Gamma_\phi + 1/q}\phi$. It follows that $L_{\alpha^\Gamma_\phi + 1/q}\phi \in \Gamma^+$ since $L_{\alpha^\Gamma_\phi + 1/q}\phi \in \mathcal{L}[\psi]$. But this is against the definition of α^Γ_ϕ.

We enumerate all the formulas in $\mathcal{L}[\psi]$:

$$\phi_1, \phi_2, \cdots, \phi_n.$$

Suppose that $\Gamma^+ \cup \{\neg L_{r_1}\phi_1\}$ is consistent for some $\alpha^\Gamma_{\phi_1} \leq r_1 < \alpha^\Gamma_{\phi_1} + 1/q$. Since our reasoning here is exactly parallel to that in Lemma 2.3.12., we will not repeat here. We just sketch the proof. By propositional reasoning, we know that there is a maximal consistent extension Γ^1 in the language $\mathcal{L}^1$ such that $\Gamma^1 \supseteq \Gamma^+ \cup \{\neg L_{r_1}\phi_1\}$ where $\mathcal{L}^1$ is a strictly more accurate langauge than $\mathcal{L}[\psi]$. Then it is easy to see that the above define $\alpha^{\Gamma^1}_{\phi_1}$ for Γ^1 is strictly smaller than $\alpha^{\Gamma}_{\phi_1} + 1/q$.

Now we consider ϕ_2 and Γ^1. Observe that $\alpha^{\Gamma}_{\phi_2} \leq \alpha^{\Gamma^1}_{\phi_2} < \alpha^{\Gamma^1}_{\phi_2} + \frac{1}{q_1} \leq \alpha^{\Gamma}_{\phi_2} + 1/q$ where q_1 is the accuracy of the langauge $\mathcal{L}^1$ (since $\neg L_{\alpha^{\Gamma}_{\phi_2}+1/q}\phi_2 \in \Gamma^+ \subseteq \Gamma^1$). By the same argument as for ϕ_1, we know that there is a maximal consistent extension Γ^2 in the strictly more accurate langauge $\mathcal{L}^2$ such that $\Gamma^2 \supseteq \Gamma^1 \cup \{\neg L_{r_2}\phi_2\}$ for some r_2: $\alpha^{\Gamma^1}_{\phi_2} \leq r_2 < \alpha^{\Gamma^1}_{\phi_2} + \frac{1}{q_1}$. So $\alpha^{\Gamma}_{\phi_2} \leq \alpha^{\Gamma^1}_{\phi_2} \leq \alpha^{\Gamma^2}_{\phi_2} \leq r_2 < \alpha^{\Gamma^1}_{\phi_2} + \frac{1}{q_1} \leq \alpha^{\Gamma}_{\phi_2} + \frac{1}{q}$.

If we repeat this process for all ϕ_i for $1 \leq i \leq n$, then we will get a chain of maximal consistent extensions Γ_i's such that

$$\Gamma \subseteq \Gamma^+ \subseteq \Gamma^1 \subseteq \cdots \subseteq \Gamma^n$$

and, for each ϕ_i,

$$\alpha^{\Gamma}_{\phi_i} \leq \alpha^{\Gamma^1}_{\phi_i} \leq \cdots \leq \alpha^{\Gamma^{i-1}}_{\phi_i} \leq \alpha^{\Gamma^i}_{\phi_i} \leq r_i < \alpha^{\Gamma^i}_{\phi_i} + 1/q_i \leq \alpha^{\Gamma^{i-}}_{\phi_i} + 1/q_{i-1} \leq \cdots \leq \alpha^{\Gamma}_{\phi_i} + 1/q$$

Note that there is one place where the inequality is strict. For such a Γ^n there is a maximal consistent extension Γ^∞ in the langauge $\mathcal{L}$ such that $\Gamma^n \subseteq \Gamma^\infty$. It is easy to see that any maximal extension of Γ^n will preserve the above inequality. This implies that, for any $\phi_i \in \mathcal{L}[\psi]$ and $\alpha^{\infty}_{\phi_i} := \{r \in \mathcal{Q} : L_r\phi_i \in \Gamma^\infty\}$,

$$\alpha^{\Gamma}_{\phi_i} \leq \alpha^{\Gamma^1}_{\phi_i} \leq \cdots \leq \alpha^{\Gamma^{n-1}}_{\phi_i} \leq \alpha^{\Gamma^n}_{\phi_i} \leq \alpha^{\infty}_{\phi_i} \leq \alpha^{\Gamma^n}_{\phi_i} + 1/q_n \leq \cdots < \alpha^{\Gamma^i}_{\phi_i} + 1/q_i \leq \cdots \leq \alpha^{\Gamma}_{\phi_i} + 1/q$$

$\square$

Lemma 2.5.2. *The following two are needed in the following definition of inner probability measure on the canonical model:*

(1) *if* $\vdash \phi_1 \to \neg\phi_2$ *and* $r + s \leq 1$, *then* $\vdash L_r\phi_1 \wedge L_s\phi_2 \to L_{r+s}(\phi_1 \vee \phi_2)$;
(2) $\vdash \neg L_r\bot$ *for any* $r > 0$.

Now we define the inner probability measures on the canonical model based on such Γ^∞. Define

$$T(\Gamma) : 2^\Omega \to [0,1] \text{ as } T(\Gamma)([\phi]) = \alpha^{\infty}_{\phi}.$$

The following part is parallel to that in the measurable case. It is easy to see that $T(\Gamma)(\Omega) = T(\Gamma)([\top]) = 1$ and $T(\Gamma)(\emptyset) = T(\Gamma)([\bot]) = 0$ since $L_1\top \in \Gamma^\infty$ (**A2**) and $\neg L_r\bot \in \Gamma^\infty$ for any $r > 0$. Next we show the finite superadditivity.

LEMMA 2.5.3. *This is the very place where we need the axiom (A5):*

(1) $T(\Gamma)([\phi]) + T(\Gamma)([\neg\phi]) \leq 1$.

(2) *If* $\vdash \phi_1 \to \neg\phi_2$, *then* $T(\Gamma)([\phi_1]) + T(\Gamma)([\phi_2]) \leq 1$.

PROOF. We just prove the first part and leave the second part to the reader. Suppose that $T(\Gamma)([\phi]) + T(\Gamma)([\neg\phi]) > 1$. Then there are two sufficiently small rationals ε and ε' such that $r := T(\Gamma)([\phi]) - \varepsilon$, $s := T(\Gamma)([\neg\phi]) - \varepsilon'$ and $r + s > 1$. According to the above definition, we know that $L_r\phi \in \Gamma^\infty$ and $L_s(\neg\phi) \in \Gamma^\infty$. Since $r + s > 1$, $\neg L_s\neg\phi \in \Gamma^\infty$ (by **(A5)**), which is a contradiction. So, $T(\Gamma)([\phi]) + T(\Gamma)([\neg\phi]) \leq 1$.

□

LEMMA 2.5.4. *For* $A, B \in 2^\Omega$, *if* $A \cap B = \emptyset$, *then* $T(\Gamma)(A) + T(\Gamma)(B) \leq T(\Gamma)(A \cup B)$.

PROOF. We know that there are formulas $\phi_1, \phi_2 \in \mathcal{L}[\psi]$ such that $A = [\phi_1]$ and $B = [\phi_2]$. Let α_1, α_2 and α_+ denote $T(\Gamma)([\phi_1])$, $T(\Gamma)([\phi_2])$ and $T(\Gamma)([\phi_1 \vee \phi_2])$, respectively. By the above lemma, we know that $\alpha_1 + \alpha_2 \leq 1$. So we only need to show that $\alpha_1 + \alpha_2 \leq \alpha_+$.

Suppose that $\alpha_1 + \alpha_2 > \alpha_+$. Then there are two $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that $(\alpha_1 - \epsilon_1) + (\alpha_2 - \epsilon_2) > \alpha_+$, $\alpha_1 - \epsilon_1 \in \mathcal{Q}$ and $\alpha_2 - \epsilon_2 \in \mathcal{Q}$. Let $\alpha_1' := \alpha_1 - \epsilon_1$ and $\alpha_2' := \alpha_2 - \epsilon_2$. It follows that $L_{\alpha_1'}\phi_1 \in \Gamma^\infty$ Similarly, $L_{\alpha_2'}\phi_2 \in \Gamma^\infty$. By **(A3)** (one pervious lemma), we know that $L_{\alpha_1'+\alpha_2'}(\phi_1 \vee \phi_2) \in \Gamma^\infty$. But this is impossible because $\alpha_1' + \alpha_2' > \alpha_+$ and α_+ is the smallest upper bound such that $L_\beta(\phi_1 \vee \phi_2) \in \Gamma^\infty$.

□

Indeed such defined $T(\Gamma)$ is an inner probability measure for all Γ.

LEMMA 2.5.5. *(Truth Lemma) Let* $T(\Gamma)$ *defined as above. Then, for any* $L_r\phi \in \mathcal{L}[\psi]$,

$$L_r\phi \in \Gamma \text{ iff } \Gamma \models L_r\phi.$$

PROOF. Assume that $\Gamma \models L_r\phi$. This implies that $\alpha_\phi^\infty \geq r$. Since $\alpha_\phi^\infty < \alpha_\phi^\Gamma + 1/q$, $r \leq \alpha_\phi^\Gamma$. It follows that $L_r\phi \in \Gamma$.

Assume that $L_r\phi \in \Gamma$. It follows that $r \leq \alpha_\phi^\Gamma \leq \alpha_\phi^\infty$. Since $T(\Gamma)([\phi]) = \alpha_\phi^\infty$, $\Gamma \models L_r\phi$.

□

THEOREM 2.5.6. *(Completeness) Let $\mathcal{M}_o$ be the class of probability models with inner measures. Then*

$$\models_{\mathcal{M}_o} \psi \textit{ iff } \vdash_{\Sigma_L} \psi.$$

PROOF. Since ψ is consistent as we assume at the beginning of this section, ψ is contained in a maximal consistent extension Γ in the language $\mathcal{L}[\psi]$. It follows from the above lemma that $\Gamma \models \psi$.

□

The noncompactness can be proved similarly.

COROLLARY 2.5.7. *Σ_L is not compact.*

2.6. Decidability

In this section, we give a new logic Σ_v and show that its relation to the system Σ_+ through which we can get the decidability of Σ_+ easily. Also this logic will justify both our claim that Σ_+ is a finite system in nature and our meta-reasoning with index parameters in Σ_+. We need the following index language with equality $\mathcal{L}_I$ as an auxiliary language.

2.6.1. Term Calculus. The index language with equality $\mathcal{L}_I$ consists of

(1) countably infinite index variables $x, y, \cdots$ (and with subscripts) and two constants 0 and 1;
(2) two predicates: $\leq$ and $<$[6];
(3) function symbols: $+, -$.

[6]We can mae do with only either one of them and define the other using equality.

An *index term* t is defined by the following syntax:

$$t := x \mid 0 \mid 1 \mid t_1 + t_2 \mid t_1 - t_2$$

We use t (with subscripts) to denote index terms. A *literal index formula* in the language $\mathcal{L}_I$ is defined as follows:

$$g =: t_1 = t_2 \mid t_1 \leq t_2 \mid t_1 < t_2 \mid \neg g$$

Now we develop some machinery that is needed for our main theorem. The following abbreviations are very handy:

- n for $1 + \cdots + 1$ for n times;
- $n \cdot t$ is short for $t + \cdots + t$ for n times;
- $\frac{n}{m} + t = u$ is short for $n + mt = mu$;
- $\frac{n}{m} + t \leq u$ is short for $n + mt \leq mu$;
- $\frac{n}{m} + t < u$ is short for $n + mt < mu$;

So we can also assume that the language includes rationals as constants. Equivalently, a literal index formula is an atomic formula or the negation of an atomic formula in the langauge $\mathcal{L}_I$. *A guard index set* G is a set of literal index formulas. An *evaluation* v is a mapping from $X := \{x_1, \cdots, x_n, \cdots\}$ to $\mathcal{Q}$, the number system of rational numbers. An index formula g is *valid* if g is true under all evaluations. A guard index set G is *valid* if all $g's$ in G are valid. g is a *logical consequence* of g' (denoted as $g' \models g$) if every evaluation that satisfies g' also satisfies g.

Our goal is to find a calculus $\vdash$ that is sound and complete with respect to the above semantics:

$$g' \vdash g \text{ iff } g' \models g.$$

2.6.2. Axiomatization in Term Calculus. The system AX_I, *in addition to those for equality*, has the following axioms and rules:

Index Calculus AX_I[7]

- $t \leq t$, $0 < 1$,
- $t_1 + 0 = t_1$, $t_1 + t_2 = t_2 + t_1$, $t_1 + (t_2 + t_3) = (t_1 + t_2) + t_3$;
- $(t_1 + t_2) - t_2 = t_1$, $(t_1 - t_2) + t_2 = t_1$;
- $\dfrac{\neg(t_1 < t_2), \neg(t_1 = t_2)}{t_2 < t_1}$, $\dfrac{t_1 \leq t_2, \neg(t_1 = t_2)}{t_1 < t_2}$;
- $\dfrac{t_1 < t_2}{t_1 \leq t_2}$; $\dfrac{t_1 < t_2}{\neg(t_1 = t_2)}$; $\dfrac{\neg(t_1 < t_2)}{t_2 \leq t_1}$; $\dfrac{t_2 \leq t_1}{\neg(t_1 < t_2)}$;
- $\dfrac{t_1 \leq t_2, t_2 \leq t_3}{t_1 \leq t_3}$; $\dfrac{t_1 \leq t_2, t_2 \leq t_1}{t_1 = t_2}$;
- $\dfrac{t_1 \leq t_2}{t_1 + t_3 \leq t_2 + t_3}$; $\dfrac{t_1 \leq t_2}{t_3 - t_2 \leq t_3 - t_1}$;
- $\dfrac{nt_1 = nt_2}{t_1 = t_2}$; $\dfrac{nt_1 \leq nt_2}{t_1 \leq t_2}$; $\dfrac{nt_1 < nt_2}{t_1 < t_2}$.
- $\dfrac{g}{t \neq t'} \Leftrightarrow \dfrac{g, \neg(t > t')}{t < t'}$; $\dfrac{g, \neg g'}{0 \neq 0} \Leftrightarrow \dfrac{g}{g'}$; $\dfrac{g, g_1}{g_2} \Leftrightarrow \dfrac{g, \neg g_2}{\neg g_1}$.
- $\dfrac{t \neq t', g}{g'} \Leftrightarrow \dfrac{t > t', g}{g'} \& \dfrac{t < t', g}{g'}$.

A literal index formula g is *derivable* from a guard index set G (denoted as $G \vdash g$) if there is a sequence s each element of which is either:

(1) an element of G, or

(2) an instance of the axioms in AX_I, or

(3) a result by applying the rules in AX_I to other elements located before in the sequence.

A guard index set G_1 is *derivable from the guard set* G_2 if $G_2 \vdash g$ for all $g \in G_1$.

THEOREM 2.6.1. *(Completeness of term calculus) For any literal index formulas g and* finite *guard index set G,*

$$G \models g \text{ iff } G \vdash_{AX_I} g.$$

In order to show this theorem, we have to apply some techniques from linear programming [32]. However, we don't use them directly but adapt them into our setting of term calculus. First we state the problem of linear programming. A term inequality system is a group

[7] Probably some of them below are redundant. But we are more concerned about the completeness of the axiomatization.

of linear inequalities with integral coefficients (but with rational constant terms) in the following form:

$$S = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n \geq c_1 \\ \cdots, \cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n \geq c_r \\ -a'_{11}x_1 - \cdots - a'_{1n}x_n > -c'_1 \\ \cdots \\ -a'_{s1}x_1 - \cdots - a'_{sn}x_n > -c'_s \end{cases}$$

Note that all the a_{ij} and a'_{ij} are integers and c_i and c'_i are rationals. First we prove the integral version of Farkas's lemma and then we show the integral version of Kuhn's transposition theorem. There are many adaptations that we can use to show these. But, for the following proof of the decidability of the term calculus, we choose to use *Fourier-Motzkin elimination method* in [32]. First we assume that $s = 0$. Then the term inequality system can be written as a more condensed matrix presentation: $Ax \geq c$ where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix}.$$

and $x = (x_1, \cdots, x_n)^T$ and $c = (c_1, \cdots, c_r)^T$.

LEMMA 2.6.2. *The following two statements are equivalent:*

(1) $Ax \geq c$ *is unsatisfiable in* $\mathcal{Q}$*;*

(2) *there exists an integral row vector* α *such that*

- *each coordinates of* α *is nonnegative;*
- $\alpha A = 0$*;*
- $\alpha c > 0$.

PROOF. We will use the Fourier-Motzkin elimination method. First we eliminate x_1. Without loss of generality, we assume that $a_{1,1}, \cdots, a_{r_+,1} > 0$, $a_{(r_+ +1),1}, \cdots, a_{r_-,1} < 0$ and

$a_{r_-+1,1}, \cdots, a_{r,1} = 0$. Let a_1 be the least (positive) common multiple of all $a_{1,1}, \cdots, a_{r_+,1}$, $a_{(r_++1),1}, \cdots, a_{r_-,1}$. Then by multiplying each of the first r_- rows by a corresponding *positive integer*, we can get the following term inequality system:

$$S_n = \begin{cases} a_1x_1 + \cdots + a^1_{1n}x_n \geq c^1_1 \\ \cdots, \cdots \\ a_1x_1 + \cdots + a^1_{r_+,n}x_n \geq c^1_{r_+} \\ -a_1x_1 + \cdots + a^1_{r_++1,n}x_n \geq c^1_{r_++1} \\ \cdots \\ -a_1x_1 + \cdots + a^1_{r_-,n}x_n \geq c^1_{r_-} \\ 0 \cdot x_1 + \cdots + a_{r_-+1,n}x_n \geq c_{r_-+1} \\ \cdots \\ 0 \cdot x_1 + \cdots + a_{r,n}x_n \geq c_r \end{cases}$$

where $a^1_{ij}(1 \leq i \leq r_-, 2 \leq j \leq n)$ are integers but $c^1_k(1 \leq k \leq r_-)$ are rationals. Then it is easy to see that the first r_- rows are equivalent to the following inequalities:

$\sum_{j=2}^n a^1_{k,j}x_j - c^1_k \geq a_1x_1 \geq c^1_i - (\sum_{j=2}^n a^1_{i,j}x_j)$ for all $i : 1 \leq i \leq r_+$ and $k : r_+ + 1 \leq k \leq r_-$.

The solvability of the system S_n with n unknown variables can be reduced to the following system S_{n-1} with $n-1$ unknown variables:

$$S_{n-1} = \begin{cases} \sum_{j=2}^n a^1_{k,j}x_j - c^1_k \geq c^1_i - (\sum_{j=2}^n a^1_{i,j}x_j) \\ a_{r_-+1,2}x_2 + \cdots + a_{r_-+1,n}x_n \geq c_{r_-+1} \\ \cdots \\ a_{r,2}x_2 + \cdots + a_{r,n}x_n \geq c_r \end{cases}$$

for all $i : 1 \leq i \leq r_+$ and $k : r_+ + 1 \leq k \leq r_-$. Note the following three facts about the system S_{n-1}:

(1) S_{n-1} has only $n-1$ unknown variables;

(2) it has $r_+ \cdot r_- + (r - r_-)$ inequalities;

(3) S_n has a rational solution iff S_{n-1} has a rational solution.

Now we show the lemma. The direction from (2) to (1) is trivial true. Now we show the other direction. We use the induction on the dimension of the term inequality system.

(1) Assume that the number of the unknown variables is 1. Then the term inequality system is just equivalent to the following inequlaity:

$$max\{c_1, \cdots, c_{r_+}\} \leq a_1 x_1 \leq min\{-c_{r_++1}, \cdots, -c_{r_-}\}.$$

Since $Ax \geq c$ is unsatisfiable, there are k and l such that $1 \leq k \leq r_=$, $r_+ + 1 \leq l \leq r_-$ and $c_k > -c_l$ (otherwise, the system would have a rational solution by the density of rationals). Therefore we can choose the row vector α for this case to be:

$$(0_1, \cdots, 0_{k-1}, 1_k, 0_{k+1}, \cdots, 1_l, 0_{l+1}, \cdots, 0_r)$$

where the subscripts indicate the location of the element in the vector. In other words, all coordinates are zeros except the k-th and l-th coordinates. It is easy to see that $\alpha A = 0$ and $\alpha c > 0$. So we finish the proof for the base case.

(2) Assume that $S_n : Ax \geq c$ have n unknown variables and does not have a rational solution. Equivalently, S_n does not have a rational solution. Then, from the above observations, we know that $S_{n-1} : A'x' = c'$, where A' is just the above matrix in the Fourier-Motzkin elimination method and $x' = (x_2, \cdots, x_n)$, does not have a rational solution, either. By induction hypothesis, we know that there is an integral vector α_{n-1} such that

(a) $\alpha_{n-1} \geq 0$;

(b) $\alpha_{n-1} A' = 0$;

(c) $\alpha_{n-1} c' > 0$.

To put it in another word, $(0, \cdots, 0, \alpha_{n-1} c')$ is a nonnegative integral combination of the row vectors of the matrix $[A', c']$. Moreover, each of the first $r_+ \cdot (r_- - r_+)$ rows of S_{n-1} is a sum of two row vectors of the first r_- rows of the matrix $[A, c]$. Therefore, $(0, 0, \cdots, 0, \alpha_{n-1} c')$ is also a nonnegative integral combination of the row vectors of the matrix $[A, c]$. That is to say, there is a row vector α_n such that

- $\alpha_n \geq 0$;
- $\alpha_n A = 0$;

- $\alpha_n c > 0$.

So we finish the inductive step and hence the whole lemma.

□

The above proof can also be adapted for the proof for the general case: $s > 0$. The system S can be represented as a matrix presentation: $Ax \geq c$ and $A'x > c'$.

LEMMA 2.6.3. *The following two statements are equivalent:*

(1) *The system $S : Ax \geq c$ and $A'x > c'$ does not have a rational solution;*

(2) *there is an integral row vector (α, α') such that:*
   - $\alpha \geq 0$ *and* $\alpha' \geq 0$*;*
   - $\alpha A + \alpha A' = 0$*;*
   - *Either* $\alpha c + \alpha c' > 0$ *or some entry of* α' *is positive, and* $\alpha c + \alpha c' \geq 0$*.*

PROOF. The proof of this lemma is similar to that for the special case $s = 0$.

□

As an illustration, we show the following lemma.

LEMMA 2.6.4. *Show that* $\{a_1t_1 + a_2t_2 \geq c_1, b_1t_1 + b_2t_2 \geq c_2\} \vdash (2a_1 + 3b_1)t_1 + (2a_2 + 3b_2)t_2 \geq 2c + 3d$*.*

PROOF. Reason inside AX_I:

$b_1t_1 + b_2t_2 \geq d, b_1t_1 + b_2t_2 \geq d$

$(b_1t_1 + b_2t_2) + (b_1t_1 + b_2t_2) \geq d + d$

$(2b_1)t_1 + (2b_2)t_2 \geq 2d$

$[(2b_1)t_1 + (2b_2)t_2] + (b_1t_1 + b_2t_2) \geq 2d + d$

$(3b_1)t_1 + (3b_2)t_2 \geq 3d$

$(2a_1)t_1 + (2a_2)t_2 \geq 2c$ (Similarly)

$(2a_1 + 3b_1)t_1 + (2a_2 + 3b_2) \geq (2c + 3d)$.

□

LEMMA 2.6.5. *The following propositions are provable in AX_I:*

(1) $0x_1 + 0x_2 + \cdots 0x_n = 0$

(2) $\neg(0 > n)$ *for any positive integer* n.

PROOF. First note that $0x = 0$. According to our convention, $0x = (1-1)x = x - x$. On the other hand, $x - x = 0$ because $0 + x = x$. Now we can see that $0 = 0x_1 = 0x_1 + 0 = 0x_1 + 0x_2 = \cdots = 0x_1 + \cdots + 0x_n$. For the second proposition, we prove by contraposition. Suppose that $0 > n$. but, according to the axiom $0 < 1$, we have that $0 < n$. Then, by one axiom in AX_I, we know that it contradicts the fact that $0 > n$.

□

Now we go back to prove Theorem 2.6.1.

PROOF. (Completeness) It is easy to check the soundness of the term calculus. Now we are showing the completeness of the term calculus. Given a finite guard index set G and an literal index formula g, assume that $G \neg \vdash g$. If g is of the form $t = t'$, then $G \not\vdash t \geq t'$ or $G \not\vdash t \leq t'$. Now we need to show that there is an evaluation v such that both G and $\neg g$ are verified under v. Let $G' = G \cup \{\neg g\}$. We can transform G' to a term inequality system. If some element of G' is an equality $t_1 = t_2$, then we can just replace it by $t_1 \leq t_2$ and $t_1 \geq t_2$. According to the last rule, we can assume that all the elements in G' are inequalities.

There are two cases that we need to consider.

(1) Assume that all of the elements of G' are of the form $t \geq t'$. By the axioms of AX_I, we can transform it into the following term inequality system:

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq c_1 \\ \cdots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rn}x_n \geq c_r \end{cases}$$

Usually it is written in matrix form: $Ax \geq c$. Suppose that it is not satisfiable. According to one previous lemma, there is an integral row vector $\alpha = (\alpha_1, \cdots, \alpha_r)$ such that:

- $\alpha \geq 0$;

- $\alpha A = 0$;
- $\alpha c > 0$.

As we have shown in the above lemma, inside AX_I we can multiply the i-th row by α_i and then add all of them together. So we will get on the left side $0x_1 + 0x_2 + \cdots + 0x_n$. On the right hand side, we will get αc, which is a positive rational. But, according to one lemma above, we will get that $0 > \alpha c$, or equivalently, $0 > n$ for some positive integer n. From this we can easily deduce that $0 \neq 0$. Sum up all the above arguments, we get that: $G' \vdash 0 \neq 0$. It follows that $G \vdash g$. But this contradicts to our assumption that $G \nvdash g$. This implies that $G \models g$.

(2) Assume that some elements of G' are of the form $t > t'$. We can transform it into the following term inequality system:

$$S = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n \geq c_1 \\ \cdots, \cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n \geq c_r \\ -a'_{11}x_1 - \cdots - a'_{1n}x_n > -c'_1 \\ \cdots \\ -a'_{s1}x_1 - \cdots - a'_{sn}x_n > -c'_s \end{cases}$$

Correspondingly it can be written as a matrix presentation: $Ax \geq c$ and $A'x > c'$. We need to show that $G \not\models g$. It suffices to show that G' is satisfiable. Suppose that it is not satisfiable. By the main lemma above, we have that there are two integral row vectors $\alpha = (\alpha_1, \cdots, \alpha_r)$ and $\alpha' = (\alpha'_1, \cdots, \alpha'_s)$ such that

- $\alpha \geq 0$ and $\alpha' \geq 0$;
- $\alpha A + \alpha A' = 0$;
- Either $\alpha c + \alpha' c' > 0$ or some entry of α' is positive and $\alpha c + \alpha' c' \geq 0$.

Here we only discuss the second case: some entry of α' is positive and $\alpha c + \alpha' c' \geq 0$. Assume that $\alpha'_s > 0$. We multiply each row by the corresponding integers:

$$\begin{cases} \alpha_1 a_{11} x_1 + \cdots + \alpha_1 a_{1n} x_n \geq \alpha_1 c_1 \\ \cdots, \cdots \\ \alpha_r a_{r1} x_1 + \cdots + \alpha_r a_{rn} x_n \geq \alpha_r c_r \\ -\alpha'_1 a'_{11} x_1 - \cdots - \alpha'_1 a'_{1n} x_n > -\alpha'_1 c'_1 \\ \cdots \\ -\alpha'_s a'_{s1} x_1 - \cdots - \alpha'_s a'_{sn} x_n > -\alpha'_s c'_s \end{cases}$$

Add the first $(r+s-1)$ rows together we will get the following inequality: $a_1x_1 + \cdots + a_nx_n \geq d$ for some integers a_i and rational d. Since $\alpha A + \alpha' A' = 0$, we have that: $a_1 = \alpha'_s a'_{s1}, \cdots, a_n = \alpha'_s a'_{sn}$. It follows that $-(a_1x_1 + \cdots + a_nx_n) > -\alpha'_s c's$. Since $\alpha c + \alpha' c' \geq 0$, $-\alpha'_s c's \geq d$. It follows that $a_1x_1 + \cdots + a_nx_n > -d$. So $(a_1x_1 + \cdots + a_nx_n) < d$. But this contradicts the fact that $a_1x_1 + \cdots + a_nx_n \geq d$. This implies that $G' \vdash 0 \neq 0$. Equivalently, $G \vdash g$, which contradicts our assumption that $G \nvdash g$. We conclude that $G \not\models g$. So we finish the proof of the completeness.

□

THEOREM 2.6.6. *The term calculus is decidable, i.e. given a finite guard index set G and a literal index formula g, it is decidable inside AX_I whether $G \vdash g$ or not.*

PROOF. The theorem follows directly from Fourier-Motzkin's elimination method. We only need to see whether $G' := G \cup \{\neg g\}$ has a rational solution or not by the above completeness. But this can be achieved by the elimination method in finite many steps. If G' has a solution, then we can conclude that $G \nvdash g$ by the completeness result. If G' does not have a solution, then we can conclude that $G \vdash g$ because otherwise G' would have a solution.

□

2.6.3. Decidability of Σ_+. After setting up the auxiliary system for our logic, we are now defining the syntax for the deductive system Σ_v. A formula $\phi[G]$ in this logic consists of two parts: the propositional part ϕ and the index part G. G is a guard index set of literal index formulas containing all index terms occurring in ϕ. The index part specifies the condition which the indexes in the propositional part must satisfy. For example, the formula $(L_{t_1}\phi \to \neg L_{t_2}\neg\phi)[t_1 + t_2 > 1]$ means that the indexes in the formula part $L_{t_1}\phi \to \neg L_{t_2}\neg\phi$ satisfies the condition: $t_1 + t_2 > 1$. If the propositional part does not contain the modalities L, we can regard the corresponding index part as $(0 = 0)$, which we just omit. $G_1 \preceq G_2$ if all the index terms occurring in G_1 occur also in G_2. Propositional part ϕ are defined as follows:

$$\phi := p \mid \phi_1 \wedge \phi_2 \mid \neg\phi \mid L_t\phi$$

When the context is clear, we also call the propositional part ϕ a formula. Given any formula ϕ, the guard index set $G_0(\phi)$ specifying that, for all index terms t occurring in ϕ, $0 \leq t \leq 1$ is called *the basic guard index set*. Without further notice, any guard index set for any formula ϕ is a superset of its basic guard index set. In order to "evaluate" the indexes variables, we also write ϕ as $\phi(x_1, \cdots, x_n)$, where $x_1, \cdots, x_n$ are all the index variables occurring in ϕ.

Probability Logic Σ_v

- (A1) $L_0\phi$
- (A2) $L_t\top$;
- (A3) $L_{t_1}(\phi \wedge \psi) \wedge L_{t_2}(\phi \wedge \neg\psi) \to L_{t_1+t_2}\phi$
- (A4) $\neg L_{t_1}(\phi \wedge \psi) \wedge \neg L_{t_2}(\phi \wedge \neg\psi) \to \neg L_{t_1+t_2}\phi$
- (A5) $L_{t_1}\phi \to \neg L_{t_2}\neg\phi\ [t_1 + t_2 > 1]$
- (ARC) $\dfrac{(\gamma \to (M_t\phi \to M_{t-t_\epsilon}\phi))[G]}{(\gamma \to M_0\phi)[G]}$, where all of index variables in t and t_ε are new;
- (DIS) $\dfrac{(\phi \leftrightarrow \psi)[G]}{(L_t\phi \leftrightarrow L_t\psi)[G, 0 \leq t \leq 1]}$, where all the index variables in t are new;
- (ARCH): $\dfrac{(\gamma \to L_{t-t_\epsilon}\phi)[G]}{(\gamma \to L_t\phi)[G]}$, where all index variables appearing in t_ϵ are new;
- (CL) propositional calculus with the rule MP replaced by the following rule:

$$MP_v\text{: } \frac{(\phi_1 \to \phi_2)[G], \phi_1[G']}{\phi_2[G, G']}, ;$$

Note that in this axiomatization, we omit all the basic guard index formulas to simplify the presentation. Here we just give a brief intuition of the rules in Σ_v. (ARC) and (ARCH) actually characterize the Archimedean property of $\mathcal{Q}$, which means that, if the pace is constant, it will hit the zero after finite step backwards. (DIS) says that the probability measure is monotonic. And MP_v is just the detachment rule in index variables. If a formula $\phi[G]$ is provable in Σ_v, it can be interpreted as follows:

if any realization *of* $\phi[G]$ [8] *satisfies the guard condition, then* ϕ *is valid in any probability structure.*

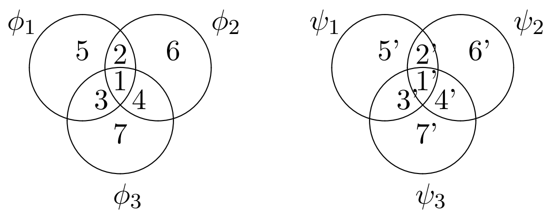
Note that according to our definition of formulas, for any formula $\phi[G]$, all the index terms in ϕ also occur in G. In the following, for any formula $\phi[G]$, $x_1, \cdots, x_n$ are all the index variables occurring in ϕ and $x_1, \cdots, x_n, \cdots, x_N$ are all the index variables occurring in G.

LEMMA 2.6.7. *The following rule (B) with index variables is admissible in* Σ_v*:*

$$\frac{((\phi_1, \cdots, \phi_n) \leftrightarrow (\psi_1, \cdots, \psi_n))[G]}{(\bigwedge_{i=1}^{n} L_{t_i}\phi_i \wedge \bigwedge_{j=2}^{n} M_{t'_j}\psi_j \to L_{\sum_{i=1}^{n} t_i - \sum_{j=2}^{n} t'_j}\psi_1)[G']},$$

where $t_1, \cdots, t_n, t'_2, \cdots, t'_n$ *are new and* $[G'] = [G, 0 \leq t_i, t_j, \sum_{i=1}^{n} t_i - \sum_{j=2}^{n} t'_j \leq 1]$*;*

PROOF. This is just an abstraction of the proof of Lemma 2.4.5. And it is needed in the connection of Σ_+ to Σ_v. In order to convince the reader, we give a proof for the case that $n = 3$.



Note that the numbers denote the areas for the corresponding formulas. Our proof is informal but informative enough. Assume that

[8]We will formalize this definition below.

$$\vdash (\phi_1, \phi_2, \phi_3) \leftrightarrow (\psi_1, \psi_2, \psi_3).$$

And assume that

$$\Xi = \begin{cases} a_1 \leq 1 \leq a_1 + \delta \\ a_2 \leq 2 \leq a_2 + \delta \\ a_3 \leq 3 \leq a_3 + \delta \\ a_4 \leq 4 \leq a_4 + \delta \end{cases}$$

Since $1 + 2 + 3 + 4 = 1' + 2' + 3' + 4'$, we can similarly assume that

$$\Xi' = \begin{cases} a_1 \leq 1' \leq a_1 + \delta \\ a_2 \leq 2' \leq a_2 + \delta \\ a_3 \leq 3' \leq a_3 + \delta \\ a_4 \leq 4' \leq a_4 + \delta \end{cases}$$

By a similar argument in Section 2.4, we have the following inequality system:

$$\begin{cases} 5 \geq r_1 - (a_1 + a_2 + a_3 + 3\delta) \\ 6 \geq r_2 - (a_1 + a_2 + a_4 + 3\delta) \\ 7 \geq r_3 - (a_1 + a_4 + a_3 + 3\delta) \end{cases}$$

Since all the numbered areas in the diagram are disjoint, we add them together and have the following inequality:

$$\phi_1 \vee \phi_2 \vee \phi_3 \geq (r_1 + r_2 + r_3) - (a_1 + a_2 + a_3 + a_4) - 9\delta - a_1.$$

which is equivalent to saying that

$$\psi_1 \vee \psi_2 \vee \psi_3 \geq (r_1 + r_2 + r_3) - (a_1 + a_2 + a_3 + a_4) - 9\delta - a_1.$$

Since $\psi_2 \leq s_2$, $6' \leq s_2 - (a_1 + a_2 + a_4)$. Similarly, we have $7' \leq s_3 - (a_1 + a_3 + a_4)$. Denote $A = a_1 + a_2 + a_3 + a_4$. Note that all the numbered areas in the second diagram are disjoint, we can use the inequality axioms in the system and get:

$$5' \geq (r_1+r_2+r_3) - (s_2+s_3) - A - 9\delta - a_1 - [s_2 - (a_1+a_2+a_4)] - [s_3 - (a_1+a_3+a_4)] - (A+4\delta)$$

By simplifying the above inequality, we have:

$$5' \geq (r_1 + r_2 + r_3) - (s_2 + s_3) - 13\delta - A + a_4$$

It follows immediately from this inequality that:

$$\psi_1 \geq (r_1 + r_2 + r_3) - (s_2 + s_3) - 13\delta.$$

By summing up the above argument, we have that: given any sufficiently small ε and δ such that $100\delta \leq \varepsilon$, from the assumptions Ξ, Ξ' and the following assumptions:

$$\phi_1 \geq r_1.\phi_2 \geq r_2, \phi_3 \geq r_3, \psi_2 \leq s_2, \psi_3 \leq s_3,$$

it follows that

$$\psi_1 \geq (r_1 + r_2 + r_3) - (s_2 + s_3) - 13\delta.$$

Suppose that $\psi_1 < (r_1 + r_2 + r_3) - (s_2 + s_3) - \varepsilon$. From the rule (ARC), we can deduce that

$$1 = 2 = 3 = 4 = 1' = 2' = 3' = 4' = 0.$$

It follows (as what we have shown in Section 3) that ϕ_1, ϕ_2 and ϕ_3 are disjoint and ψ_1, ψ_2, ψ_3 are disjoint. From this result, we have that

$$\phi_1 \vee \phi_2 \vee \phi_3 \geq r_1 + r_2 + r_3 \text{ while } \psi_1 \vee \psi_2 \vee \psi_3 < r_1 + r_2 + r_3 - \varepsilon.$$

But this contradicts our assumption that $\vdash (\psi_1 \vee \psi_2 \vee \psi_3) \leftrightarrow (\phi_1 \vee \phi_2 \vee \phi_3)$. So $\psi_1 \geq (r_1 + r_2 + r_3) - (s_2 + s_3) - \varepsilon$. By the rule (ARCH), we have that $\psi_1 \geq (r_1 + r_2 + r_3) - (s_2 + s_3)$, which is exactly what we want to show.

$\square$

DEFINITION 2.6.8. A *realization* τ of the formula $\phi[G]$ is a mapping from $X := \{x_1, \cdots, x_n\}$ to $\mathcal{Q} \cap [0, 1]$ such that

$$AX_I \vdash G(x_1/\tau(x_1), \cdots, x_n/\tau(x_n)).$$

Now we are well-equipped to state the main theorem of this section about the relationship between Σ_v and Σ_+:

THEOREM 2.6.9. *Given any formula $\phi(x_1, \cdots, x_n)[G(x_1, \cdots, x_N)]$, if $\Sigma_v \vdash \phi[G]$, then,* *(S): for any realization τ of G, $\Sigma_+ \vdash \phi(x_1/\tau(x_1), \cdots, x_n/\tau(x_n))$.*

Proof. Prove by induction on the proof sequence of $\phi[G] \in \Sigma_v$.

(1) Basis case: ϕ is an axiom in Σ_v. The it is easy to see that the proposition (S) holds. Here we take $L_t(\phi \wedge \psi) \wedge L_{t'}(\phi \wedge \neg\psi) \rightarrow L_{t+t'}\phi$ as an illustration. The implicit guard index set is the basic guard index set. For any realization τ, it is easy to see that

$$(L_t(\phi \wedge \psi) \wedge L_{t'}(\phi \wedge \neg\psi) \rightarrow L_{t+t'}\phi)[x_1/r_1, \cdots, x_n/r_n]$$

is an axiom in Σ_+.

(2) Assume that $\phi := (\gamma \rightarrow M_0\psi)[G]$ and it is a result by applying the rule (ARC): $\dfrac{(\gamma \rightarrow (M_t\phi \rightarrow M_{t-t_\epsilon}\psi))[G]}{(\gamma \rightarrow M_0\psi)[G]}$, where all of index variables in t and t_ε are new

Given any realization τ of ϕ, then it can be extended to a realization $\tau[y_1/r_1, \cdots, y_k/r_k]$ of $(\gamma \rightarrow (M_t\phi \rightarrow M_{t-t_\epsilon}\phi))[G]$ (where $y_1, \cdots, y_k$ are the index variables occurring in t or t_ϵ) because of the following two obvious facts:

(a) all of index variables in t and t_ε are new;

(b) $G \vdash_{AX_I} G[x_1/r_1, \cdots, x_k/r_k]$ for all $r_i : 0 \leq r_i \leq 1$, where x_i are those appearing in t or t_ε.

The Induction Hypothesis implies that, under the realization τ, for any $r > 0$, $\gamma \rightarrow M_r\phi$ is provable in Σ_+. By appealing to the rule (ARCH) in Σ_+, we know that $\phi(x_1/\tau(x_1), \cdots, x_n/\tau(x_n))$ is provable in Σ_+.

(3) Assume that $\phi := (L_t\gamma \leftrightarrow L_t\psi)[G, 0 \leq t \leq 1]$ is a result by applying the rule (DIS) in Σ_v:

$$\frac{(\gamma \leftrightarrow \psi)[G]}{(L_t\gamma \leftrightarrow L_t\psi)[G, 0 \leq t \leq 1]}, \text{ where all the index variables in } t \text{ are new;}$$

Let τ be a realization of $(L_t\gamma \leftrightarrow L_t\psi)[G, 0 \leq t \leq 1]$. If we take a corresponding restriction to the index variables in $(\gamma \rightarrow \psi)$, then we will get a realization τ' of $(\gamma \leftrightarrow \psi)[G]$ because G is a subset of $[G, 0 \leq t \leq 1]$ and t is new. Since τ' is a realization of $(\gamma \leftrightarrow \psi)[G]$, by the induction hypothesis, we know that $(\gamma \leftrightarrow \psi)[x_1/\tau'(x_1), \cdots, x_k/\tau'(x_k)]]$ is provable in Σ_+ where $x_1, \cdots, x_k$ are all the index variables occurring in $\gamma \leftrightarrow \psi$. By the rule (DIS) in Σ_+, we know that $(L_t\phi \leftrightarrow L_t\psi)[x_1/\tau(x_1), \cdots, x_k/\tau(x_k), \cdots x_n/\tau(x_n)]$ is provable in Σ_+, where $x_{k+1}, \cdots, x_n$ are the index variables occurring in t.

(4) The proof of the case for (ARCH) is similar to that for the case for (ARC);

(5) The case for the rule CL is also similar with a little more complication. Assume that $\phi := \phi_2[G, G']$ is a result by applying the rule MP_v:

$$\frac{(\phi_1 \to \phi_2)[G], \phi_1[G']}{\phi_2[G, G']}.$$

Assume that τ is a realization of $\phi_2[G, G']$. That is to say, $AX_I \vdash G[x_1/\tau(x_1), \cdots, x_n/\tau(x_n)]$ and $AX_I \vdash G'[x_1/\tau(x_1), \cdots, x_n/\tau(x_n)]$. Without loss of generality, we can assume that the index variable set of ϕ_1 is the same as that of ϕ_2. For example, if x is an index variable in ϕ_1 but not in ϕ_2. Then we can add the index formula $x = x$ to ϕ_2 to make up for this "shortcoming". Assume that τ is a realization of $\phi_2[G, G']$. It is easy to see that it is also a realization of both $(\phi_1 \to \phi_2)[G]$ and $\phi_1[G']$ because both $[G]$ and $[G']$ are subsets of $[G, G']$. According to the induction hypothesis,

$$\Sigma_+ \vdash (\phi_1 \to \phi_2)(x_1/\tau(x_1), \cdots, x_n/\tau(x_n)) \text{ and } \Sigma_+ \vdash \phi_1(x_1/\tau(x_1), \cdots, x_n/\tau(x_n)).$$

It follows immediately that

$$\Sigma_+ \vdash \phi_2(x_1/\tau(x_1), \cdots, x_n/\tau(x_n))$$

□

Corollary 2.6.10. *If $\Sigma_v \vdash \phi(x_1, \cdots, x_n)[G(x_1, \cdots, x_N])$ and, for any realization τ, $AX_I \vdash G(x_1/\tau(x_1), \cdots, x_n/\tau(x_n), \cdots, x_{n+j}/r_1, \cdots, x_{n+j+i_j}/r_{i_j}, \cdots, x_N)$, then $\Sigma_v \vdash \phi(x_1, \cdots, x_n)[G(x_1, \cdots, x_n, \cdots, x_{n+j}/r_1, \cdots, x_{n+j+i_j}/r_{i_j}, \cdots, x_N)]$*

Note that the realizations here evaluate only the index variables in ϕ but not completely those in G. The extra condition just says that any expansion of realizations must validate the guard set.

In the proof of the following theorem, we choose the axiomatization in [17] because their deductive system has a very nice tree property. Of course, we have shown that their system is equivalent to ours.

THEOREM 2.6.11. *If $\phi(x_1/r_1, \cdots, x_n/r_n)$ is provable in Σ_+, then there is a guard index set G such that $\Sigma_v \vdash \phi(\overrightarrow{x})[G]$ and $\tau : x_i \mapsto r_i (1 \leq i \leq n)$ is a realization of G.*

PROOF. We prove on the proof sequence of the formula $\phi(x_1/r_1, \cdots, x_n/r_n)$ in Σ_+.

(1) Basis case: $\phi(x_1/r_1, \cdots, x_n/r_n)$ is an axiom in Σ_+. We can just choose the basic guard index set G_0 for the corresponding axiom in Σ_v.

(2) Assume that $\phi := (L_r\gamma \leftrightarrow L_r\psi)$ is a result by applying the rule:
$$\frac{\gamma \leftrightarrow \psi}{L_r\gamma \leftrightarrow L_r\psi}.$$
According to the induction hypothesis, there is a guard set G such that $\Sigma_v \vdash \phi(\overrightarrow{x})[G]$ and $\tau : x_i \mapsto r_i$ is a realization. Define $[G'] := [G, 0 \leq x \leq 1]$ and $\tau' := \tau \cup \{(x, r\}$. It is easy to see that $\Sigma_v \vdash (L_x\gamma \leftrightarrow L_x\psi)[G']$ and τ' is a realization.

(3) Assume that $\phi := \bigwedge_{i=1}^n L_{r_i}\phi_i \wedge \bigwedge_{j=2}^n M_{r'_j}\psi_j \rightarrow L_{\sum_{i=1}^n r_i - \sum_{j=2}^n r'_j}\psi_1$ is a result by applying the rule:
$$\frac{(\phi_1, \cdots, \phi_n) \leftrightarrow (\psi_1, \cdots, \psi_n)}{\bigwedge_{i=1}^n L_{r_i}\phi_i \wedge \bigwedge_{j=2}^n M_{r'_j}\psi_j \rightarrow L_{\sum_{i=1}^n r_i - \sum_{j=2}^n r'_j}\psi_1}.$$
By induction hypothesis, we know that there is a guard index set G such that $\Sigma_v \vdash ((\phi_1, \cdots, \phi_n) \leftrightarrow (\psi_1, \cdots, \psi_n))(x_1, \cdots, x_n)$ where $x_1, \cdots, x_n$ are all the index variables that occur in $((\phi_1, \cdots, \phi_n) \leftrightarrow (\psi_1, \cdots, \psi_n))$ and $\tau : x_i \mapsto r_i$ is a realization. Now we define:
$$[G'] := [G, 0 \leq y_i \leq 1, 0 \leq z_j \leq 1, 0 \leq \textstyle\sum_{i=1}^n y_i - \sum_{j=2}^n z_j \leq 1]$$
where y_i's and z_j's are new and $1 \leq i \leq n$ and $2 \leq j \leq n$. Moreover, we define:
$$\tau' := \tau \cup \{(y_i, r_i), (z_j, r'_j) : 1 \leq i \leq n, 2 \leq j \leq n\}.$$
By Lemma 2.6.7, we know that
$$\Sigma_v \vdash (\textstyle\bigwedge_{i=1}^n L_{y_i}\phi_i \wedge \bigwedge_{j=2}^n M_{z'_j}\psi_j \rightarrow L_{\sum_{i=1}^n y_i + \sum_{j=2}^n z'_j}\psi_1)[G']$$
And the τ' is a realization.

(4) Assume that $\phi(x_1/r_1, \cdots, x_n/r_n)$ is a result by applying the rule:

$$\frac{\psi \to \phi, \psi}{\phi}(x_1/r_1, \cdots, x_n/r_n)$$

Here we also assume that $\phi(x_1/r_1, \cdots, x_n/r_n)$ and $\psi(x_1/r_1, \cdots, x_n/r_n)$ are instantiations of the formulas $\phi(x_1, \cdots, x_n)$ and $\psi(x_1, \cdots, x_n)$ respectively, which have the same index variable set. According to the induction hypothesis, there are two guard index sets G and G' such that

$$\Sigma_v \vdash (\psi \to \phi)[G] \text{ and } \Sigma_v \vdash \psi[G']$$

And $\tau : x_i \mapsto r_i$ is a realization of both $(\psi \to \phi)[G]$ and $\psi[G']$. By the rule MP_v, we know that $\Sigma_v \vdash \phi[G, G']$ and τ is a realization of the formula $\phi[G, G']$. This case also close our induction proof.

□

Corollary 2.6.12. *Σ_+ is decidable.*

Proof. First we know from the above completeness proof through filtration that Σ_+ has the finite model property. Now we need to show that the set of theorems of Σ_+ is recursively enumerable. In order to show that the logic Σ_+ is recursively enumerable, we only need to show that the set of theorems of Σ_v is recursively enumerable (by the above lemma). Compared to the decidability of normal modal logic, we need to consider only one more question: is the relationship among indexes of the different modalities is decidable? It is easy to see that the set of theorems of Σ_v is recursively enumerable. Now we use this result to show that the theorems of Σ_+ is recursively enumerable. We let one Turing machine to enumerate all the theorems derived from the system Σ_v and let the second Turing machine to instantiate the index variables in the theorems by the indexes in ψ. If ψ is provable in Σ_+, there is such a instantiation which is a realization (by the above theorem). Moreover, this can be decided. And hence, by Theorem 2.6.6, we know that ψ is provable in Σ_+. If it is not, it should be falsifiable in the finite canonical model in our proof of completeness. So, Σ_+ is decidable.

□

In the last part of this section, we will justify the reasoning with index parameters in the logic Σ_+. First we can expand the language $\mathcal{L}_I$ by including the following set of constants: $C_Q := \{c_r : r \in \mathcal{Q} \cap [0,1]\}$. Let Δ_Q be the set of all atomic formulas or the negation of atomic formulas that are true in $\langle \mathcal{Q} \cap [0,1], \leq, + \rangle$. And it is called the *diagram* of the index set $\mathcal{Q} \cap [0,1]$. For example, the following principles are in this diagram:

(1) $c_r + c_s = c_{r+s}$;

(2) $c_r \leq c_s$ if $r \leq s$ in $\mathcal{Q}$.

Let AX be $AX_I \cup \Delta_Q$ in this expanded language with C_Q and Σ_{v+c} be the corresponding deductive system in the expanded langauge. All the above proposition still hold in this expanded language. Now we justify our use of the following rule with a parameter in the logic Σ_+:

$$\frac{\gamma \to L_{r-x_\varepsilon}\phi}{\gamma \to L_r\phi}, \text{ where } x_\varepsilon \text{ is the only index variable in this rule.}$$

Assume that $r_1, \cdots, r_n$ (including r) are all the index constants occurring in $\gamma \to L_r\phi$. If $\gamma \to L_r\phi$ is a theorem in Σ_+, there is a guard index set G such that $AX \vdash G(r_1, \cdots, r_n)$ and $\mathcal{L}_{v+c} \vdash (\gamma \to L_r\phi)G(r_1, \cdots, r_n)$. In order to show this formula, it suffices to show the following instead:

$$\Sigma_{v+c} \vdash (\gamma \to L_{r-x_\varepsilon}\phi)[G(r_1, \cdots, r_n), 0 \leq x_\varepsilon \leq 1, 0 \leq r - x_\varepsilon \leq 1].$$

So, in order to show that $\Sigma_+ \vdash \phi(r_1, \cdots, r_n)$, it is suffices to show that

$$\Sigma_{v+c} \vdash (\gamma \to L_{r-x_\varepsilon}\phi)[G(r_1, \cdots, r_n), 0 \leq x_\varepsilon \leq 1, 0 \leq r - x_\varepsilon \leq 1] \text{ for some guard index set } G.$$

Usually this G is easy to get because it is just the basic guard index set for this formula. The illustration of how to use this principle is in Section 4 where we show the admissibility of the rule (B) in our deductive system Σ_+.

2.7. Moss's Conjecture

The Kreisel conjecture is a famous conjecture in proof theory that postulates that, if there is a uniform bound on the lengths of shortest proofs of instances of $S(\bar{n})$, then the universal generalization is provable in Peano arithmetic. Basing on the similarity of the rule (ARCH) and the Godel's ω rule in Peano arithmetic, Professor Moss pointed to me that it is quite possible that if there is a uniform bound on the set $\{s : s < r\}$ such that $\gamma \to L_s\phi$ is derivable in our system Σ_+, so is $\gamma \to L_r\phi$. In this section, we show that this conjecture is true by applying the above Fourier-Motzkin's elimination method. But, first we show that the probability measures on the canonical model in Section 3 can be computed *constructively*. There we only gave an existence proof.

Elements in Ω are called *atoms*. Let $n = |\Omega|$. Fix an atom Γ and its maximal consistent extension Γ^+ in the language $\mathcal{L}^+$. As you can see, the atoms in Ω are the building blocks in the semantics. We enumerate all the atoms Γ_i in Ω and denote them as x_i according to this ordering. For any formula ϕ in $\mathcal{L}[\psi]$, if ϕ is propositionally equivalent to the disjunction of different atoms: $\Gamma_1, \Gamma_2, \cdots, \Gamma_r$, then we denote ϕ as $x_1 + x_2 + \cdots + x_r$. Then the inequality system consists of the following inequalities:

(1) For any formula ϕ in $\mathcal{L}[\psi]$, if $\phi = x_1 + x_2 + \cdots + x_r$ and $\alpha^\Gamma_\phi < \beta^\Gamma_\phi$, then its inequality is

$$\alpha^\Gamma_\phi < x_1 + x_2 + \cdots + x_r < \beta^\Gamma_\phi.$$

(2) For any formula ϕ in $\mathcal{L}[\psi]$, if $\phi = x_1 + x_2 + \cdots + x_r$ and $\alpha^\Gamma_\phi = \beta^\Gamma_\phi$, then its inequality is

$$\alpha^\Gamma_\phi \le x_1 + x_2 + \cdots + x_r \text{ and } x_1 + x_2 + \cdots + x_r \le \beta^\Gamma_\phi.$$

(3) For each x_i $(1 \le i \le n)$, we add the inequalities:

$$x_i \ge 0 \text{ and } x_i \le 1.$$

(4) Actually this case is covered in the first case. In order to emphasize its importance, we make it explicit here:

$$x_1 + x_2 + \cdots + x_n \ge 1 \text{ and } -x_1 - x_2 - \cdots - x_n \ge -1.$$

To sum up, we get an inequality system like

$$S_\Gamma = \begin{cases} x_{i_1} + \cdots + x_{i_{\Gamma_i}} \geq r_i \\ \cdots \\ x_{k_1} + \cdots + x_{k_{\Gamma_k}} \geq r_k \\ -x_{i'_1} + \cdots - x_{i'_{\Gamma_i}} \geq -r'_i \\ \cdots \\ -x_{k'_1} + \cdots - x_{k'_{\Gamma_k}} \geq -r'_k \\ x_1 \geq 0 \\ \cdots \\ x_n \geq 0 \\ -x_1 \geq -1 \\ \cdots \\ -x_n \geq -1 \end{cases}$$

The *accuracy* of the system is the smallest natural number q such that any coefficients and constant terms are mutiple of its reciprocal $1/q$. Since we already know from the proof of the completeness of the system Σ_+ that this inequality system S_Γ has a solution, then we can use the Fourier-Motzkin's elimination to find the solutions of all $x_i's$ and hence a probability measure defined at the atom Γ. So we have shown that the probability measures at each atom of the canonical model can be defined constructively.

But this is not the end of the story. We can also show that the construction of "sandwich-making" in the proof of the completeness is constructive, too. At the beginning of the construction, we proved the following claim:

for any ϕ such that $\alpha_\phi^\Gamma < \beta_\phi^\Gamma$, then there is a rational r such that $\alpha_\phi^\Gamma < r < \beta_\phi^\Gamma$ and $\Gamma^+ \cup \{M_r\phi\}$ is consistent.

Moreover, we can prove a stronger result than that by constructively giving the rational r as a special case of Moss' conjecture.

LEMMA 2.7.1. *Let $\psi, \mathcal{L}[\psi]$ and Ω as those defined in the proof of the completeness of Σ_+. Then we can constructively find a rational number ε, which depends only on the syntactic form of ψ, such that, for any ϕ in $\mathcal{L}[\psi]$ and any Γ^+, if $\vdash \bigwedge \Gamma^+ \to L_{\beta^\Gamma_\phi - \epsilon}\phi$, then $\vdash \bigwedge \Gamma^+ \to L_{\beta^\Gamma_\phi}\phi$*

PROOF. The proof of the lemma is based on the observations about the elimination method of the inequality system like the above S_Γ. Since $\neg M_{\beta^\Gamma_\phi}\phi \in \Gamma^+$ and Γ_+ is consistent, $\Gamma \cup \{\neg M_{\beta^\Gamma_\phi}\phi\}$ has a model. We still use S_Γ to denote the inequality system obtained in the above way. Assume that the number of the unknowns in S_Γ is n and denote S_Γ as S_n. By appealing to the elimination method, we will get the following system with $(n-1)$ unknown variables:

$$S_{n-1} = \begin{cases} \sum_{j=2}^{n} a^1_{k,j}x_j - c^1_k \geq c^1_i - (\sum_{j=2}^{n} a^1_{i,j}x_j) \\ \sum_{j=2}^{n} a^1_{k',j}x_j - c^1_{k'} > c^1_{i'} - (\sum_{j=2}^{n} a^1_{i',j}x_j) \\ a_{r_-+1,2}x_2 + \cdots + a_{r_-+1,n}x_n \geq c_{r_-+1} \\ \cdots \\ a_{r,2}x_2 + \cdots + a_{r,n}x_n \geq c_r \end{cases}$$

for some $k's$ and $i's$. Finally, we will reach the base case: the number of the unknown variable is 1:

$max\{r_i : i \in I\}R\ x_nR\ min\{s_j : j \in J\}$ for some index set I and J where R is $>$ or $\geq$.

Observe that

(1) The number of the atoms (or just $|\Omega|$) is determined by the number of propositional letters occurring in ψ, the accuracy of ψ and the depth of ψ. And so is the number N of the inequalities in the inequality system S_Γ.
(2) Let q be the accuracy of the language $\mathcal{L}[\psi]$. Then $q!$ is the accuracy of all of the inequality system S_i $(1 \leq i \leq N)$. Indeed, all the reductions from S_{i+1} to S_i involves either division by a rational whose denominator is less than or equal to q or addition or subtraction, neither of these two operations will increase the

accuracy of the system. That is to say, the coefficients and constant terms of all the inequality systems S_i $(1 \leq i \leq N)$ are multiples of the rational $\frac{1}{q!}$.

Now we begin to estimate the ε. Consider the base case:

$$max\{r_i : i \in I\}R\; x_n R\; min\{s_j : j \in J\} \text{ for some index set } I \text{ and } J \text{ where } R \text{ is } > \text{ or } \geq.$$

If both of $R's$ are $\leq$, then we choose either $max\{r_i : i \in I\}$ or $min\{s_j : j \in J\}$. Note that both are still multiples of $\frac{1}{q!}$. If at least one of $R's$ is $<$, then we pick up the middle point of the two. At worst, it is a multiple of $\frac{1}{2 \cdot q!}$. Now we consider x_{n-1}. According to the elimination process, in order to find a solution of x_{n-1}, we only need to consider the following inequalities:

$$\begin{cases} x_{n-1} R b_1 x_n + c_1 \\ \cdots \\ x_{n-1} R b_k x_k + c_k \\ x_{n-1} S b_1' x_n + c_1' \\ \cdots \\ x_{n-1} S b_l' x_l + c_l' \end{cases}$$

where $R's$ are $\leq$ or $<$ and $S's$ are $\geq$ or $>$. Let $\alpha_{n-1} = max\{b_1' x_n + c_1', \cdots, S b_l' x_l + c_l'\}$ and $\beta_{n-1} = min\{b_1 x_n + c_1, \cdots, S b_k x_k + c_k\}$. Then we know that $\alpha_{n-1} R x_{n-1} R \beta_{n-1}$. If the first R is $\leq$, then we choose $x_{n-1} = \alpha_{n-1}$. Similarly for the second R. If both of them are $<$, then we choose the middle points of α_{n-1} and β_{n-1}. In the first case, x_{n-1} is a multiple of $\frac{1}{2(q!)^2}$ at worst. In the second case, x_{n-1} is a multiple of $\frac{1}{2^2(q!)^2}$ at worst. If we repeat this process, then we will get the following estimation:

$$x_i \text{ has a solution which is a multiple of } \frac{1}{2^i(q!)^i} \text{ for all } i : 1 \leq i \leq n.$$

We conclude that, if an inequality system of accuracy q has a solution, then it has a solution which is a multiple of $\frac{1}{2^n(q!)^n}$, i.e. x_i is a a multiple of $\frac{1}{2^n(q!)^n}$ for all $i : 1 \leq i \leq n$. Let $\varepsilon = \frac{1}{4^n(q!)^n}$.

Now we show the main lemma by translating the language of inequality system into the langauge of modal logic of probability. Suppose that $\nvdash \bigwedge \Gamma^+ \to L_{\beta^\Gamma_\phi}\phi$. Then $\Gamma^+ \cup \neg L_{\beta^\Gamma_\phi}\phi$ is consistent. According to our completeness, we know that it has a probability model. That is also to say, the above translated inequality system has a rational solution which is a multiple of $\frac{1}{2^n(q!)^n}$. Without loss of generality, we assume that in the above translation from formulas to inequalities:

$$x_1 + \cdots + x_r < \beta^\Gamma_\phi \text{ where } \phi \text{ is denoted by } x_1 + \cdots + x_r.$$

Note that β^Γ_ϕ is a multiple of $1/q$. According to the above solutions for the whole inequality system, $x_1 + \cdots + x_r$ is a multiple of $\frac{1}{2^n(q!)^n}$. Since $x_1 + \cdots + x_r < \beta^\Gamma_\phi$, $\beta^\Gamma_\phi - (x_1 + \cdots + x_r) > \frac{1}{2^n(q!)^n}$. Of course, $\beta^\Gamma_\phi - (x_1 + \cdots + x_r) > \frac{1}{4^n(q!)^n}$. That is to say, $x_1 + \cdots + x_r < \beta^\Gamma_\phi - \frac{1}{4^n(q!)^n}$. In the langauge of our deductive system, it says:

$$\Gamma^+ \cup \neg L_{\beta^\Gamma_\phi - \frac{1}{4^n(q!)^n}}\phi \text{ has a model and hence } \nvdash \bigwedge \Gamma^+ \to L_{\beta^\Gamma_\phi - \frac{1}{4^n(q!)^n}}\phi$$

So we finish the proof of the lemma.

□

Now we show Moss's theorem in the general form:

THEOREM 2.7.2. *For formulas γ and ϕ, we can constructively find a sufficiently small rational number ε, which depends only on the depth, the accuracy and the number of propositional letters of γ and ϕ such that*

$$\vdash_{\Sigma_+} \gamma \to L_{r-\varepsilon}\phi \Rightarrow \vdash_{\Sigma_+} \gamma \to L_r\phi.$$

PROOF. The first part of the proof is actually the repetition of the construction of the canonical model in the proof of the completeness of Σ_+. Consider the formula $\psi := \gamma \wedge \neg L_r\phi$. Use the same method before, we can get a local langauge $\mathcal{L}[\psi]$ and it gives rise to a carrier set Ω of maximal consistent sets of formulas in the local langauge. Then, in order to define α^Γ_ϕ for all formulas ϕ in the language and for all atoms Γ, we expand the local language to $\mathcal{L}_+$ by adding the depth by 1. Let q be the accuracy of the language $\mathcal{L}[\psi]$ and n be the number of the atoms in Γ, which depends only on the depth, the accuracy and the number of propositional letters of γ and ϕ. We choose $\varepsilon = \frac{1}{4^n(q!)^n}$. Suppose that $\nvdash_{\Sigma_+} \gamma \to L_r\phi$. Then $\gamma \wedge \neg L_r\phi$ is consistent. So it has a probability model. This is equivalent to saying that the corresponding inequality system has a solution. By the argument in the above lemma, we know that it has a solution which is a multiple of $\frac{1}{2^n(q!)^n}$. It means that $\gamma \wedge \neg L_r\phi$ is satisfiable in a probability model with transition probabilities which are multiples of $\frac{1}{2^n(q!)^n}$. This implies that $\gamma \wedge \neg L_{r-\varepsilon}$ is also satisfiable in this probability model. So $\vdash_{\Sigma_+} \gamma \to L_r\phi$ and hence finish the proof of Moss's conjecture.

□

This theorem also implies that we can replace the rule (ARCH) in Σ_+ by the following finitary rule:

$$(ARCH^f) : \frac{\gamma \to L_{r-\varepsilon}\phi}{\gamma \to L_r\phi}.$$

Let Σ_+^f denote the deductive system after this replacement. It is easy to see that the completeness proof still goes through because the rule $ARCH^f$ is stronger than $ARCH$. We only need to worry about the soundness of $ARCH^f$. This is just shown in the above argument by using Fourier-Motzkin's elimination. Indeed, our system Σ_+ can be strengthened to a finite system Σ_+^f. But we still choose to work with Σ_+ because it looks much more natural and easier to *use* with.

In the following, we will give an example to illustrate the above algorithm . Consider the following inequality system:

$$S_3 = \begin{cases} x_1 - 2x_3 \geq -1/4 \\ -x_1 - x_2 > -7/8 \\ 2x_2 + x_3 > 3/4 \\ -2x_1 + 3x_3 \geq 1/8 \end{cases}$$

By applying the elimination method, we will get the following system:

$$S_2 = \begin{cases} -x_2 - 2x_3 > -9/8 \\ -1/2x_3 \geq -3/16 \\ 2x_2 + x_3 > 3/4 \end{cases}$$

with the constraints for x_1:

$$\begin{cases} x_1 \geq -1/4 + 2x_3 \\ x_1 < -x_2 + 7/8 \\ x_1 \leq 3/2x_3 - 1/16 \end{cases}$$

By repeating the above process, we have:

$$S_1 = \begin{cases} -3x_3 > -6/4 \\ -1/2x_3 \geq -3/16 \end{cases}$$

with the constraints for x_2

$$\begin{cases} x_2 > 1/2x_3 - 3/8 \\ x_2 < 9/8 - 2x_3 \end{cases}$$

Since $3/8 < 1/2$, we choose $x_3 = 3/8$, which does not increase the accuracy of the system. Since $3/16 < x_2 < 3/8$, we choose the midpoint: $x_2 = 9/32$. By substituting these numbers, we get that $1/2 \leq x_1 < 25/32$ and $x_1 \leq 1/2$. So we can just pick up either of the bounds: $x_1 = 1/2$ and don't need to increase the accuracy. Obviously, all the solutions are multiple of the rational $\frac{1}{2^3 8^3}$ where 8 is the accuracy of the system S_3 and 3 is the number of the unknown variables. The above theorem tells us that, the following evaluation:

$$\begin{cases} x_1 = 1/2 \\ x_2 = 9/32 \\ x_3 = 3/8 \end{cases}$$

is also a solution of the following inequality system:

$$S_3 = \begin{cases} x_1 - 2x_3 \geq -1/4 \\ -x_1 - x_2 > -7/8 - \frac{1}{4^3 8^3} \\ 2x_2 + x_3 > 3/4 - \frac{1}{4^3 8^3} \\ -2x_1 + 3x_3 \geq 1/8 \end{cases}$$

CHAPTER 3

Failure of Strong Completeness

In our proof of the completeness of Σ_+, we use a filtration method from modal logic. The main reason for using filtration is the fact that Σ_+ is not compact, i.e., there is a set of formulas which is finitely satisfiable but not itself satisfiable. For example, $C := \{\neg L_{1/2}p\} \cup \{L_{1/2-1/2^n}p : n \in N\}$ is not compact. It seems that the noncompactness did cause some troubles to Aumann when he was trying to give a syntactic definition of consequence in his paper [4]. He defined consequence semantically instead, which is significantly different from what he did for knowledge in another paper [3], where he used a deductive system to define consequence syntactically. Although he did succeed in proving the existence of the *universal knowledge-belief system*, he did not achieve this goal syntactically as we just mentioned. Our work as well as Heifetz and Mongin's on probability logic did provide an axiomatization which is weakly complete with respect to the class of type spaces. By these deductive systems, we can define syntactically only consequence of a *finite* number of formulas.

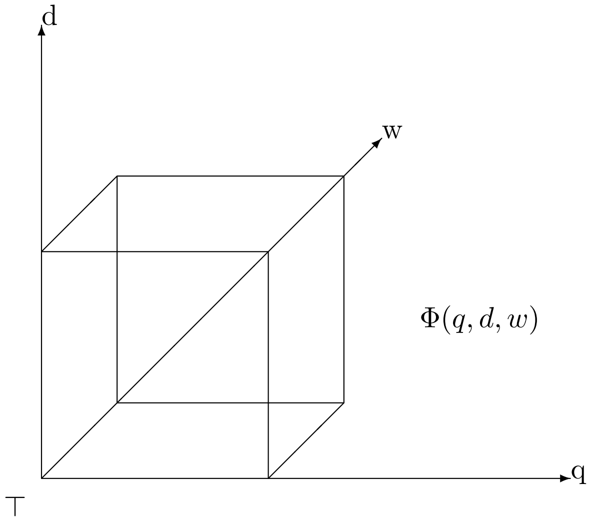
The first choice to define syntactically consequences of an *arbitrary* set of formulas is to avoid noncompact sets of formulas like the above C. *Admissible sets* are those maximal consistent sets of formulas that don't contain such noncompact set of formulas. It is easy to see that all satisfiable maximal consistent sets are admissible. Then we might expect that all admissible sets form a canonical countably-additive probability model. In the following section, we will show that the canonical model consisting of admissible sets is finitely additive but not countably additive. This implies that there is an admissible set that is not satisfiable in the class of type spaces. Moreover, it tells us that non-Archimedean property in indices isn't the failure of additivity related to (ARCH) for the lack of strong completeness or compactness of probability logics.

The second attempt is to restrict the indices to a finite set of rationals. We hoped that this restriction would prevent indices from getting close to non-Archimedean property. However, we can still show that there is a set of formulas with finite indices that derive such noncompact set like $C := \{\neg L_{1/2}p\} \cup \{L_{1/2-1/2^n}p : n \in N\}$. There is no hope of strong completeness or compactness for probability logics with finite indices, either. So probably the lack of compactness of probability logics ultimately does not come from non-Archemedean property in indices but comes from the use of explicit probabilities.

3.1. Admissible Sets Don't Necessarily Define Types

As we know, our language of probabilistic logic is three-dimensional. We use index vector (q, d, w) to denote its three dimensions: q is the accuracy index, d the depth index and w the width index. For a given formula ϕ, its accuracy index $q(\phi)$ is the least common multiple of all denominators of the indices that appear in ϕ, its depth index $d(\phi)$ is the depth of the formula and its width index $w(\phi)$ is the number of propositional letters that occur in ϕ. Here we assume much knowledge of the terminology and construction from the completeness proof of Σ_+. For a given formula ψ, assume that it is consistent and its index vector is $(q(\psi), d(\psi), w(\psi))$. Define $\Phi(q(\psi), d(\psi), w(\psi))$ to be the smallest set of formulas that satisfies the following conditions:

(1) the accuracy indices of any formula is a multiple of $1/q(\psi)$;
(2) the propositional letters that occur are among $p_1, \cdots, p_{w(\psi)}$;
(3) the depths of formulas are $\leq d(\psi)$.



Without loss of generality, we assume that all the accuracy indices in this section are factorials. Note that $\Phi(q(\psi), d(\psi), w(\psi))$ is finite because we take the quotient of propositional reasoning. This finite set $\Phi(q(\psi), d(\psi), w(\psi))$ gives rise to a set $\Omega(q(\psi), d(\psi), w(\psi))$ of maximal consistent sets of formulas, which are called *atoms* of $\Omega(q(\psi), d(\psi), w(\psi))$. Let A_q denote the set of rationals between 0 and 1 which is a multiple of $1/q$. Next we extend each atom Γ to a maximal consistent extension Γ^+ in the language $\mathcal{L}(q(\psi), d(\psi)+1, w(\psi))$ and define, for any formula ϕ in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))$,

$$\alpha_\phi^\Gamma := \max\{r \in A_{q(\psi)} : L_r\phi \in \Gamma^+\} \text{ and } \beta_\phi^\Gamma := \min\{r \in A_{q(\psi)} : M_r\phi \in \Gamma^+\}.$$

For each Γ, there is a maximal consistent extension Γ^∞ in the formula language $\mathcal{L}$ of our logic. Such a Γ_+ defines a probability measure $T(\Gamma)$ at Γ on the set $\{[\phi] : \phi$ is in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))\}$ of events where $[\phi] := \{\Delta \in \Omega(q(\psi), d(\psi), d(\psi)) : \phi \in \Delta\}$. We have shown that $\alpha_\phi^{\Gamma^\infty} = \beta_\phi^{\Gamma^\infty}$ for all formulas ϕ in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))\}$. And we define $T(\Gamma)([\phi]) := \alpha_\phi^{\Gamma^\infty}$. And then we can show that,

- if $\alpha_\phi^\Gamma = \beta_\phi^\Gamma$, then $T(\Gamma)([\phi]) = \alpha_\phi^\Gamma = \beta_\phi^\Gamma$;
- if $\alpha_\phi^\Gamma < \beta_\phi^\Gamma$, then $\alpha_\phi^\Gamma < T(\Gamma)([\phi]) < \beta_\phi^\Gamma$,

If we define $v(p) := \{\Delta \in \Omega(q(\psi), d(\psi), w(\psi)) : p \in \Delta\}$, then we can show that $M(q(\psi), d(\psi), w(\psi)) := \langle \Omega(q(\psi), d(\psi), w(\psi)), 2^{\Omega(q(\psi), d(\psi), w(\psi))}, T, v\rangle$ is a probability model for ψ.

More generally, we can show the truth lemma: for any formula ϕ in the language $\mathcal{L}(q(\psi)$, $d(\psi), w(\psi))$

$$M(q(\psi), d(\psi), w(\psi)), \Gamma \models \phi \text{ iff } \phi \in \Gamma.$$

Now we look at the model from the perspective of linear programming just as in [14]. We used this perspective to solve Moss's conjecture. First we enumerate all the atoms in $\Omega(q(\psi), d(\psi), w(\psi))$:

$$\Gamma_1, \Gamma_2, \cdots, \Gamma_n, \cdots, \Gamma_{N(q(\psi),d(\psi),w(\psi))}$$

where $N(q(\psi), d(\psi), w(\psi))$ is the number of atoms in $\Omega(q(\psi), d(\psi), w(\psi))$, which is totally determined by the index vector $(q(\psi), d(\psi), w(\psi))$. Next we want to define a *rational* probability measure at each atom Γ_i in the sense that transition probabilities from atoms to atoms are rational. The probability measures at all atoms are independent of each other. For any ϕ in the language, the probability measure of ϕ at the atom Γ_i is decided by its $\alpha_\phi^{\Gamma_i}$ and $\beta_\phi^{\Gamma_i}$, which is defined through the chosen maximal consistent extension Γ^+. Without loss of generality, we use a constructive algorithm to give a rational probability measure at the atom Γ_1. We use x_{ij} to denote a transition probability from the atom Γ_i to the atom Γ_j. Next we translate all the constraints in the forms $\alpha_\phi^{\Gamma_1}$ and $\beta_\phi^{\Gamma_1}$ to a linear inequality system.

$K(q(\psi), d(\psi), w(\psi))$ denotes the number of formulas in the set $\Phi(q(\psi), d(\psi), w(\psi))$. And the atoms that includes ϕ are: $\Gamma_{i_1}, \cdots, \Gamma_{i_{J(\phi)}}$ where $J(\phi)$ is a natural number. For each formula ϕ in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))$, if $\alpha_\phi^{\Gamma_1} = \beta_\phi^{\Gamma_1}$, then the inequality system includes the following two

- $x_{1i_1} + x_{1i_2} + \cdots + x_{1i_{j(\phi)}} \geq \alpha_\phi^{\Gamma_1}$ and
- $-x_{1i_1} - x_{1i_2} + \cdots - x_{1i_{j(\phi)}} \geq -\alpha_\phi^{\Gamma_1}$.

If $\alpha_\phi^{\Gamma_1} < \beta_\phi^{\Gamma_1}$, then the inequality system includes the following two

- $x_{1i_1} + x_{1i_2} + \cdots + x_{1i_{j(\phi)}} > \alpha_\phi^{\Gamma_1}$ and
- $-x_{1i_1} - x_{1i_2} + \cdots - x_{1i_{j(\phi)}} > -\beta_\phi^{\Gamma_1}$.

Since there are $M(q(\psi), d(\psi), w(\psi))$ formulas in the language, the inequality system must includes the above $2 \cdot K(q(\psi), d(\psi), w(\psi))$ inequalities. For each $x_{1k}(1 \leq k \leq N((q(\psi), d(\psi), w(\psi))$, it must satisfies the following constraints:

- $x_{ik} \geq 0$
- $-x_{ik} \geq -1$.

Since there are $N(q(\psi), d(\psi), w(\psi))$ atoms, there are $2 \times N(q(\psi), d(\psi), w(\psi))$ such inequalities. Besides, x_{1-} is a probability measure. So we have to include the following two more inequalities:

- $x_{11} + x_{12} + \cdots + x_{1,N(q(\psi),d(\psi),w(\psi))} \geq 1$;
- $-x_{11} - x_{12} - \cdots - x_{1,N(q(\psi),d(\psi),w(\psi))} \geq -1$;

If we put all these $2 \times K(q(\psi), d(\psi), w(\psi)) + 2 \times N(q(\psi), d(\psi), w(\psi)) + 2$ inequalities together, then we get the inequality system that corresponds to the probability formulas in Γ^+:

$$S = \begin{cases} x_{1i_1} + x_{1i_2} + \cdots + x_{1i_{j(\phi)}} & > \quad \alpha_\phi^{\Gamma_1} \\ -x_{1i_1} - x_{1i_2} + \cdots - x_{1i_{j(\phi)}} & > \quad -\beta_\phi^{\Gamma_1} \\ \cdots & \quad\ \cdots \\ x_{1i'_1} + x_{1i'_2} + \cdots + x_{1i_{j(\phi)'}} & > \quad \alpha_\phi^{\Gamma_1} \\ -x_{1i'_1} - x_{1i'_2} + \cdots - x_{1i_{j(\phi)'}} & > \quad -\beta_\phi^{\Gamma_1} \\ x_{ik} & \geq \quad 0 \\ -x_{ik} & \geq \quad -1 \\ \cdots\cdots & \\ x_{11} + x_{12} + \cdots + x_{1,N(q(\psi),d(\psi),w(\psi))} & \geq \quad 1 \\ -x_{11} - x_{12} - \cdots - x_{1,N(q(\psi),d(\psi),w(\psi))} & \geq \quad -1 \end{cases}$$

The completeness proof of Σ_+ has shown that the above linear inequality system has a solution. If we use the Fourier-Motzkin's elimination method, then we can show that the system has a rational solution $x_{1k} = r_{1k}^{(1)}$, i.e., all $r_{1k}^{(1)}$'s are rational and, moreover, there is a sufficiently small rational number ε_M such that all $r_{1k}^{(1)}$'s are multiples of ε_M. For example, we can take ε_M to be $\frac{1}{4^n(q(\psi)!)^n}$ where n is the number of atoms, i.e., $N(q(\psi), d(\psi), w(\psi))$. Here we pick $\varepsilon_M := 1/(q^{(1)})$ such that $q^{(1)}$ is a factorial and is divided by $q(\psi)$. For each ϕ

in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))$, if all the atoms that contain ϕ are $\Gamma_{i_1}, \cdots, \Gamma_{i_{J_\phi}}$, then we define:

$$r_\phi^{(1)} := r_{1,i_1}^{(1)} + \cdots + r_{1,i_{J_\phi}}^{(1)}.$$

It is easy to see that $r_{\bigwedge \Gamma_k}^{(1)} = r_{1k}^{(1)}$ according to this definition. Observe that all these $r_\phi^{(1)}$'s are multiples of $1/(q^{(1)})$. Define

$$\Theta_1 := \Gamma_1^+ \cup \{L_{r_\phi^{(1)}}\phi, M_{r_\phi^{(1)}}\phi : \phi \in \Phi(q(\psi), d(\psi), w(\psi))\}.$$

It easy to check that Θ_1 is satisfiable at Γ_1 in the canonical model $M(q(\psi), d(\psi), w(\psi)) = \langle \Omega(q(\psi), d(\psi), w(\psi)), 2^{\Phi(q(\psi), d(\psi), w(\psi))}, T, v\rangle$. Actually, the truth lemma can be generalized to the language $\mathcal{L}(q(\psi), d(\psi) + 1, w(\psi))$:

$$M(q(\psi), d(\psi), w(\psi)), \Gamma \models \phi \text{ iff } \phi \in \Gamma^+ \text{ for any formula } \phi \in \Phi(q(\psi), d(\psi) + 1, w(\psi)).$$

It follows that Θ_1 is consistent. Without loss of generality, we assume that $q^{(1)} > d(\psi), w(\psi)$. Then Θ_1 has a maximal consistent extension $\Gamma_1^{(1)}$ in the language $\mathcal{L}(q^{(1)}, q^{(1)}, q^{(1)})$. Note that, for any formula ϕ in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))$, both $L_{r_\phi^{(1)}}\phi$ and $L_{r_\phi^{(1)}}\phi$ are in $\Gamma_1^{(1)}$. For a maximal consistent extension $\Gamma_1^{(1)+}$ of $\Gamma_1^{(1)}$ in the language $\mathcal{L}(q^{(1)}, q^{(1)}+1, q^{(1)})$, we define $\alpha_\gamma^{\Gamma_1^{(1)}}$ and $\beta_\gamma^{\Gamma_1^{(1)}}$ similarly for all formulas γ in $\Phi(q^{(1)}, q^{(1)}, q^{(1)})$. The above observation means that, for any formula ϕ in the language $\mathcal{L}(q(\psi), d(\psi), w(\psi))$,

$$\alpha_\phi^{\Gamma_1^{(1)}} = \beta_\phi^{\Gamma_1^{(1)}} = r_\phi^{(1)}$$

As we pointed out in the proof of Moss's conjecture, such an ε_M depends only on the index vector $(q(\psi), d(\psi), w(\psi))$. This implies that we can just repeat the above procedure to all other Γ_i's ($1 \leq i \leq N$): (1) establishing a corresponding inequality system, (2) using the Fourier-Motzkin's elimination method to get a rational solution which are multiples of $1/(q^{(1)}$, (3) defining corresponding Θ_i's, (4) all the Θ_i's have maximal consistent extensions $\Gamma_i^{(1)}$ in the language $\mathcal{L}(q^{(1)}, q^{(1)}, q^{(1)})$ such that,

$$\alpha_\phi^{\Gamma_i^{(1)}} = \beta_\phi^{\Gamma_i^{(1)}} \text{ for all formulas } \phi \text{ in the language } \mathcal{L}(q(\psi), d(\psi), w(\psi))$$

which is defined through maximal consistent extensions $\Gamma_i^{(1)+}$ in the language $\mathcal{L}(q^{(1)}, q^{(1)} + 1, q^{(1)})$. Now we summarize the above results in the following theorem.

THEOREM 3.1.1. *There is a natural number $q^{(1)}$ satisfying the following conditions:*

(1) *$q^{(1)}$ is factorial and is a multiple of $q(\psi)$, and $q^{(1)} > d(\psi), w(\psi)$.*

(2) *for each atom $\Gamma_i \in \Omega(q(\psi), d(\psi), w(\psi))$, there is a maximal consistent extension $\Gamma_i^{(1)}$ in the language $\mathcal{L}(q^{(1)}, q^{(1)}, q^{(1)})$ such that,*

$$\alpha_\phi^{\Gamma_i^{(1)}} = \beta_\phi^{\Gamma_i^{(1)}} \text{ for all formulas } \phi \text{ in the language } \mathcal{L}(q(\psi), d(\psi), w(\psi))$$

which is defined through maximal consistent extensions $\Gamma_i^{(1)+}$ in the langauge $\mathcal{L}(q^{(1)}, q^{(1)}+1, q^{(1)})$.

More generally, we enumerate all atoms in $\Omega(q^{(k)}, q^{(k)}, q^{(k)})$:

$$\Gamma_{k,1}, \Gamma_{k,2}, \cdots, \Gamma_{k,N_k} \text{ where } N_k >> N.$$

And we can achieve a more general result for all k's.

Theorem 3.1.2. *There is a natural number $q^{(k+1)}$ satisfying the following conditions:*

(1) *$q^{(K+1)}$ is factorial and is a multiple of $q^{(k)}$.*

(2) *for each atom $\Gamma_{k,i} \in \Omega(q^{(k)}, q^{(k)}, q^{(k)}) (1 \leq i \leq N_k)$, there is a maximal consistent extension $\Gamma_{k,i}^{(k+1)}$ in the language $\mathcal{L}(q^{(k+1)}, q^{(k+1)}, q^{(k+1)})$ such that,*

$$\text{(C): } \alpha_\phi^{\Gamma_{k,i}^{(k+1)}} = \beta_\phi^{\Gamma_{k,i}^{(k+1)}} \text{ for all formulas } \phi \text{ in the language } \mathcal{L}(q^{(k)}, q^{(k)}, q^{(k)})$$

which is defined through maximal consistent extensions $\Gamma_{k,i}^{(k+1)+}$ in the langauge $\mathcal{L}(q^{(k+1)}, q^{(k+1)}+1, q^{(k+1)})$.

3.1.1. Admissible Sets of Formulas.

Definition 3.1.3. A maximal consistent extension Γ^∞ in the language $\mathcal{L}$ of our logic is *admissible* if, for all formulas ϕ and for all rationals r and r_n such that $\lim_{n\to\infty} r_n = r$ and $r_n < r$, $\{L_{r_n}\phi : n \in \mathcal{N}\} \cup \{\neg L_r\phi\} \not\subseteq \Gamma^\infty$.

Note that the condition in the definition is equivalent to saying that, whenever $\gamma \to L_s\phi \in \Gamma^\infty$ for all $s < r$, then $\gamma \to L_r\phi \in \Gamma^\infty$, i.e., Γ^∞ is closed under the rule (IND). Now we construct the infinite canonical model. For any atom $\Gamma_{k,i} \in \Omega(q^{(k)}, q^{(k)}, q^{(k)})$, there is a maximal consistent extension $\Gamma_{k,i}^{(k+1)}$ in the language $\mathcal{L}(q^{(k+1)}, q^{(k+1)}, q^{(k+1)})$ such that the property (C) is satisfied. This relation between $\Gamma_{k,i}$ and $\Gamma_{k,i}^{(k+1)}$ is denoted by $\Gamma_{k,i} \prec \Gamma_{k,i}^{(k+1)}$. Similarly, there is a maximal consistent extension $\Gamma_{k,i}^{(k+2)}$ in the langauge $\mathcal{L}(q^{(k+2)}, q^{(k+2)}, q^{(k+2)})$such that $\Gamma_{k,i}^{(k+1)} \prec \Gamma_{k,i}^{(k+2)}$. If we continue this kind of argument, then we will get the following chain:

$$\Gamma_{k,i} \prec \Gamma_{k,i}^{(k+1)} \prec \Gamma_{k,i}^{(k+2)} \prec \cdots \prec \Gamma_{k,i}^{(k+n)} \prec \cdots$$

Note that $q^{(k)} < q^{(k+1)} < \cdots < q^{(k+n)} < \cdots$. We also denote $\Gamma_{k,i}$ as $\Gamma_{k,i}^{(k)}$.

LEMMA 3.1.4. $\bigcup_{j=k}^{\infty} \Gamma_{k,i}^{(j)}$ *is admissible.*

PROOF. For any given ϕ in the language $\mathcal{L}$, there is a j such that $\phi \in \Phi(q^{(j)}, q^{(j)}, q^{(j)})$. This implies that either $\alpha_\phi^{\Gamma_{k,i}^{(j)}} = \beta_\phi^{\Gamma_{k,i}^{(j)}}$ or $\alpha_\phi^{\Gamma_{k,i}^{(j)}} < \beta_\phi^{\Gamma_{k,i}^{(j)}}$, both of which are defined through a maximal consistent extension $\Gamma_{k,i}^{(j)+}$ int the language $\mathcal{L}(q^{(j)}, q^{(j)}+1, q^{(j)})$. According to the above construction, $\alpha_\phi^{\Gamma_{k,i}^{(j+1)}} = \beta_\phi^{\Gamma_{k,i}^{(j+1)}}$. For any $m \geq j+1$, $\alpha_\phi^{\Gamma_{k,i}^{(m)}} = \beta_\phi^{\Gamma_{k,i}^{(m)}}$ because any maximal consistent extension can only shorten the distance between α and β but there is nothing to contract the distance (=0) between these two here. If we look at all the accuracy indices r in the formulas of the form $L_r\phi$, we can see an increasing sequence which becomes a constant sequence after a sufficiently large number of steps:

$$\cdots \leq \alpha_\phi^{\Gamma_{k,i}^{(j)}} \leq \alpha_\phi^{\Gamma_{k,i}^{(j+1)}} = \alpha_\phi^{\Gamma_{k,i}^{(j+2)}} = \alpha_\phi^{\Gamma_{k,i}^{(j+3)}} = \cdots$$

So $\bigcup_{j=k}^{\infty} \Gamma_{k,i}^{(j)}$ is admissible.

□

THEOREM 3.1.5. *(Existence Theorem) For any language $\mathcal{L}(q,d,w)$, any atom $\Gamma \in \Omega(q,d,w)$ has at least one* admissible *maximal consistent extension in the formal language of our logic.*

PROOF. For any given atom $\Gamma \in \Omega(q,d,w)$, there is a k such that $\Gamma \in \Omega(q^{(k)}, q^{(k)}, q^{(k)})$. Assume that Γ is $\Gamma_{k,i}$ in the enumeration of all atoms in $\Omega(q^{(k)}, q^{(k)}, q^{(k)})$. Then $\Gamma = \Gamma_{k,i} \subseteq \bigcup_{j=k}^{\infty} \Gamma_{k,i}^{(j)}$, which is admissible.

□

One may ask whether *all* admissible maximal consistent extensions in the language $\mathcal{L}$ can be obtained in the above way of "stabilizing" restrictions of the set in different local languages through the Fourier-Motzkin elimination method. The answer is no. Actually most of them cannot be finitely approximated in this way. In the above construction, we used a *uniform* algorithm to find the accuracy index for all the atoms at this level. For example, as stated in Theorem 3.1.1, we found a natural number $q^{(k+1)}$ such that

- for **each** atom $\Gamma_{k,i} \in \Omega(q^{(k)}, q^{(k)}, q^{(k)})(1 \leq i \leq N_k)$, there is a maximal consistent extension $\Gamma_{k,i}^{(k+1)}$ in the language $\mathcal{L}(q^{(k+1)}, q^{(k+1)}, q^{(k+1)})$ such that,

  (C): $\alpha_\phi^{\Gamma_{k,i}^{(k+1)}} = \beta_\phi^{\Gamma_{k,i}^{(k+1)}}$ for all formulas ϕ in the language $\mathcal{L}(q^{(k)}, q^{(k)}, q^{(k)})$

For a group of linear inequality systems like in Section 2.6, if all of them have solutions, they definitely have rational solutions. It is not necessarily true that *all rational* solutions of *all* systems are multiples of a single rational unit of the form $1/q$ for some natural number q. This is just as obvious as the Archimedean property. Most of them will misses the above construction in a uniform way. Our reasoning here is an adaption to our problem of Cantor's diagonalization method to show that there is no bijection between $\mathcal{N}$ and $\mathcal{R}$. If we cannot use the above method to find all the admissible maximal consistent sets in $\mathcal{L}$, then how can we finitely approximate any given admissible set? Now we give a brief description of this finite approximation. Consider the admissible set Γ^∞. Define

$$\Gamma^\infty(n, n, n) := \Phi(n, n, n) \cap \Gamma^\infty.$$

In other words, $\Gamma^\infty(n, n, n)$ is the set of formulas with index vectors $(q, d, w)(q \leq n, d \leq n, w \leq w)$ in Γ^∞. Similarly, for any $\phi \in \Phi(n, n, n)$, we can define $\alpha_\phi^{\Gamma^\infty(n,n,n)}$ and $\beta_\phi^{\Gamma^\infty(n,n,n)}$ through a maximal consistent extension $\Gamma^{\infty+}(n, n, n)$ in the language $\mathcal{L}(n, n + 1, n)$ which should be also a subset of Γ^∞. We already know that

$$\cdots \leq \alpha_\phi^{\Gamma^\infty(n-1,n-1,n-1)} \leq \alpha_\phi^{\Gamma^\infty(n,n,n)} \leq \alpha_\phi^{\Gamma^\infty(n+1,n+1,n+1)} \leq \cdots$$

Also we know that, since Γ^∞ is admissible, there is an N_ϕ such that, if $n \geq N_\phi$, then $\alpha_\phi^{\Gamma^\infty(n,n,n)}$=$\alpha_\phi^{\Gamma^\infty(n+1,n+1,n+1)}$. Since the number of formulas in $\Phi(n, n, n)$ is finite, we can find a common N^n such that, for all formulas $\phi \in \Phi(n, n, n)$ and all $n \geq N^n$,

$$\alpha_\phi^{\Gamma^\infty(n,n,n)} = \alpha_\phi^{\Gamma^\infty(n+1,n+1,n+1)}.$$

This also implies that, for all formulas $\phi \in \Phi(n, n, n)$,

$$\alpha_\phi^{\Gamma^\infty(N^n,N^n,N^n)} = \beta_\phi^{\Gamma^\infty(N^n,N^n,N^n)}$$

In other words, this equation will be preserved under any maximal consistent extensions of $\Gamma^{\infty}(N^n, N^n, N^n)$. Now we can see how this admissible set miss the above uniform construction. Without loss of generality, we assume that $n = q^{(k)}$ for some k and $N^n \leq q^{(k=1)}$. As we know, the rational solution to the linear inequality system that corresponds to $\Gamma^{\infty}(n, n, n)$ is not necessarily unique, it is quite possible that $(\Gamma^{\infty}(n, n, n))^{(k+1)} \cap \Phi(N^n, N^n, N^n)$ is not the same as $\Gamma^{\infty}(N^n, N^n, N^n)$. This also means that the admissible set miss the uniform construction at this step. Although we can apply the uniform maximal consistent extensions to all atoms in $\Omega(q^{(k+1)}, q^{(k+1)}, q^{(k+1)})$, Γ^{∞} can still miss this construction at this step, too. If Γ^{∞} keep missing the method all the time, it also means that we can obtain it in the above uniform way. Indeed, it can miss the uniform method all the time because we can probably make the additional formulas in the finite approximation of Γ^{∞} behave differently from they do in the uniform finite approximation. So it is not necessarily true that all maximal consistent extensions can be obtained in the above uniform construction. However, this negative fact does not affect our proof of the completeness because the most important fact that we need is the above Existence Theorem.

Now we construct the canonical model. The state space Ω^{∞} is the set of all admissible maximal consistent sets of formulas in the langauge $\mathcal{L}$. Define $[\phi] := \{\Gamma^{\infty} \in \Omega^{\infty} : \phi \in \Gamma^{\infty}\}$. The σ-algebra $\mathcal{A}^{\infty}$ on the state space is defined by the σ-algebra generated by the set $\Phi^{\infty} := \{[\phi] : \phi \text{ is a formula in } \mathcal{L}\}$. It is easy to see that Φ^{∞} is an algebra or a field. Define $v^{\infty}(p) := [p]$. It remains to define the transition probability function T^{∞} on the space.

Lemma 3.1.6. *For any formula ϕ, $\alpha_{\phi}^{\Gamma^{\infty}} = \beta_{\phi}^{\Gamma^{\infty}}$. Moreover, $L_{\alpha_{\phi}^{\Gamma^{\infty}}}\phi \in T^{\infty}$ and $M_{\alpha_{\phi}^{\Gamma^{\infty}}}\phi \in T^{\infty}$*

Proof. The proof of the first part was already shown before. The second part follows from the fact that Γ^{∞} is admissible.

$\square$

For any $\Gamma^{\infty} \in \Omega^{\infty}$, define $T^{\infty}(\Gamma^{\infty})([\phi]) = \alpha_{\phi}^{\Gamma^{\infty}}$. Now we want to go over some results in measure theory. Also we give a detailed proof of these results though.

DEFINITION 3.1.7. A *probability measure on an algebra* $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0,1]$ satisfying the properties

(1) $\mu(\emptyset) = 0, \mu(\Omega) = 1$;
(2) If $A_1, \cdots, A_n, \cdots$ is a disjoint sequence of sets in $\mathcal{A}$ *whose union is in* $\mathcal{A}$, then

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

We already shown the finite additivity of $T^\infty(\Gamma^\infty)$. Assume that $\{[\phi_i]\}_i$ is a disjoint sequence of sets in Φ^∞ whose union $[\phi]$ is in Φ^∞. Suppose that we could show that

$$T^\infty(\Gamma^\infty)(\bigcup[\phi_i]) = \sum_{i=1}^{\infty} T^\infty(\Gamma^\infty)([\phi_i]).$$

Then we could define an outer measure $\mu^*(\Gamma)^\infty$ at Γ^∞ on the state space Ω^∞ as follows:

$$\text{for any subset } E \subseteq \Omega^\infty \text{ , } \mu^*(\Gamma^\infty)(E) = \inf\{\sum_{i=1}^{\infty} T^\infty(\Gamma^\infty)([\phi_i]) : E \subseteq \bigcup_{i=1}^{\infty}[\phi_i]\}$$

This outer measure would define a type at Γ^∞ on the σ-algebra generated by the algebra of all sets $[\phi]$. However, $T^\infty(\Gamma^\infty)(\bigcup[\phi_i])$ is not necessarily equal to $\sum_{i=1}^{\infty} T^\infty(\Gamma^\infty)([\phi_i])$.

LEMMA 3.1.8. *For any formulas γ_1 and γ_2, $\vdash \gamma_1 \to \gamma_2$ iff $[\gamma_1] \subseteq [\gamma_2]$. Especially, for any finite conjunction $\bigwedge_{i=1}^{N} \phi_i$, we have that $\vdash \bigwedge_{i=1}^{N} \phi_i \to \phi$.*

PROOF. Suppose that $\nvdash \gamma_1 \to \gamma_2$. Then $\gamma_1 \wedge \neg\gamma_2$ is consistent. By the above Existence Theorem, there is an admissible set Δ^∞ such that $\gamma_1 \wedge \neg\gamma_2 \in \Delta^\infty$. This implies that $[\gamma_1] \subseteq [\gamma_2]$. The other direction is straightforward. The second part follows from the fact that $\bigcup_{i=1}^{\infty}[\phi_i] = [\phi]$ □

THEOREM 3.1.9. $T^\infty(\Gamma^\infty)(\bigcup_{i=1}^{\infty}[\phi_i]) \geq \sum_{i=1}^{\infty} T^\infty(\Gamma^\infty)([\phi_i])$.

However, it is not generally true that $T^\infty(\Gamma^\infty)(\bigcup_{i=1}^{\infty}[\phi_i]) \leq \sum_{i=1}^{\infty} T^\infty(\Gamma^\infty)([\phi_i])$. Since $T^\infty(\Gamma^\infty)$ is finitely additive, the above inequality is equivalent to the following statement:

- If $[\phi_1] \supseteq [\phi_2] \supseteq \cdots$ and $\bigcap_{i=1}^{\infty}[\phi_i] = \emptyset(= [\bot])$, then $T^\infty(\Gamma^\infty)(\bigcap_{i=1}^{\infty}[\phi_i]) = 0$.

In the following section, we show that this statement does not necessarily hold. So admissible sets don't form a probability model.

3.2. Guided Maximal Consistent Extension

For any local language $\mathcal{L}(q,d,w)$, we can define the set $\Omega(q,d,w)$. The elements of $\Omega(q,d,w)$ can be enumerated as follows:

$$\Gamma_1(q,d,,w), \Gamma_2(q,d,w), \cdots, \Gamma_{N(q,d,w)}(q,d,w)$$

$\gamma_i(q,d,w)$ denotes $\bigwedge \Gamma_i(q,d,w)$. We can define a probability space $S(q,d,w)$ on $\Omega(q,d,w)$: $S(q,d,w) := \langle \Omega(q,d,w), 2^{\Omega(q,d,w)}, P(q,d,w)\rangle$ where $P(q,d,w)$ is any given probability measure on $\Omega(q,d,w)$. For simplicity, we denote $P(q,d,w)(\Gamma_i(q,d,w))(1 \leq i \leq N(q,d,w))$ by $p_i^{(d)}$. So $p_1^{(d)} + \cdots + p_{N(q,d,w)}^{(d)} = 1$. For any formula $\phi \in \Phi(q,d,w)$, we can show that

$$\phi \leftrightarrow \bigvee_{\phi\in\Gamma_i(q,d,w)} \gamma_i(q,d,w) \text{ is a tautology.}$$

It is easy to see that $P(q,d,w)([\phi]) = \sum_{\phi\in\Gamma_i(q,d,w)} p_i^{(d)}$. Define $l_\phi^{S(q,d,w)}$ to be the largest multiple of $\frac{1}{q}$ which is less than or equal to $P(q,d,w)([\phi])$ and $m_\phi^{S(q,d,w)}$ to be the smallest multiple of $\frac{1}{q}$ which is greater than or equal to $P(q,d,w)([\phi])$. Then such a probability space $S(q,d,w)$ will determine two probability formulas about ϕ:

$$L_{l_\phi^{S(q,d,w)}}\phi \text{ and } M_{m_\phi^{S(q,d,w)}}\phi.$$

Define $\Xi(q,d,w)$ to be the set of all these formulas, i.e.

$$\Xi(q,d,w) := \{L_{l_\phi^{S(q,d,w)}}\phi, M_{m_\phi^{S(q,d,w)}}\phi : \phi \in \Phi(q,d,w)\}.$$

LEMMA 3.2.1. *The above defined $\Xi(q,d,w)$ is consistent.*

PROOF. It suffices to show that $\Xi(q,d,w)$ is contained in some atom in $\Omega(q,d+1,w)$. In some sense, this lemma can be regarded as a reverse of our completeness theorem where, given a maximal consistent set of formulas in some local langauge, we need to find a probability measure at this set. In order to show that, we construct this atom from the bottom.

We start this construction from $\Omega(1,0,w)$, which is the set of formulas of all Boolean combinations of propositional letters occurring in $\Omega(q,d,w)$. So the accuracy index does not play any role for the set of propositional formulas. The reason why we chose 1 as the accuracy index is to emphasize that the local language is the least accurate. Enumerate all the atoms in $\Omega(1,0,w)$:

$$\Gamma_1(1,0,w), \Gamma_2(1,0,w), \cdots, \Gamma_{N(1,0,w)}(1,0,w).$$

Fix an atom $\Gamma_1(1,0,w) \in \Omega(1,0,w)$, we want to define a probability measure on $\Omega(1,0,w)$ at this atom which agrees with the constraints imposed by $\Xi(q,d,w)$. We only need to define the transition probabilities from one atom to another. Define

$$T(\Gamma_1(1,0,w))((\Gamma_i(1,0,w)) = \sum_{\Gamma_i(1,0,w) \subseteq \Gamma_j(q,d,w)} p_j^{(d)}.$$

It is easy to check that such defined $T(\Gamma_1(1,0,w))$ is a probability measure on the state space $\Omega(1,0,w)$. For other atoms $\Gamma_i(1,0,w) (i \neq 1)$, we define the probability measures at these atoms according to their maximal consistent extensions Γ_i^{∞} in the infinite language obtained through Γ_i^{+}, which is a maximal consistent extension in the language $\mathcal{L}(q,1,w)$, by the construction method in our completeness proof. Define $M(1,0,w) := \langle \Omega(1,0,w), 2^{\Omega(1,0,w)}, T, \nu \rangle$ where $\nu(p) = \{\Gamma(1,0,w) \in \Omega(1,0,w) : p \in \Gamma(1,0,w)\}$. So $M(1,0,w)$ is a type space (or probability model).

Claim 3.2.2. *$\Xi(1,0,w)$ is satisfiable at $\Gamma_1(1,0,w)$.*

For any formula $\phi \in \Phi(1,0,w)$, $T(\Gamma_1(1,0,w))([\phi]) = \sum_{\phi \in \Gamma_j(1,0,w)} T(\Gamma_1(1,0,w))(\Gamma_j(1,0,w))$ $= \sum_{\phi \in \Gamma_i(q,d,w)} p_i^{(d)}$. It follows that $l_\phi^{S(q,d,w)} \leq T(\Gamma_1(1,0,w))([\phi]) \leq m_\phi^{S(q,d,w)}$. Therefore $L_{l_\phi^{S(q,d,w)}}\phi$ and $M_{m_\phi^{S(q,d,w)}}\phi$ are satisfiable at $\Gamma_1(1,0,w)$. Hence $\Xi(1,0,w)$ is satisfiable at $\Gamma_1(1,0,w)$.

According to the definition of ν, we know that $\Gamma_1(1,0,w)$ is satisfiable at $\Gamma_1(1,0,w)$ too. So $\Gamma(1,0,w) \bigcup \Xi(1,0,w)$ is satisfiable at $\Gamma_1(1,0,w)$ and hence is consistent. There is a maximal consistent extension $\Gamma(q,1,w)$ such that $\Gamma(q,1,w) \supseteq \Gamma_1(1,0,w) \cup \Xi(1,0,w)$ and $\Gamma(q,1,w) \in \Omega(q,1,w)$.

Enumerate all the atoms in $\Omega(q,1,w)$:

$$\Gamma_1(q,1,w), \Gamma_2(q,1,w), \cdots, \Gamma_{N(q,1,w)}(q,1,w).$$

Without loss of generality, $\Gamma_1(q,1,w) = \Gamma(q,1,w)$. Define the probability measure $T(\Gamma_1$ $(q,1,w))$ at the atom $\Gamma_1(q,1,w)$ as follows:

$$T(\Gamma_1(q,1,w))(\Gamma_i(q,1,w)) = \sum_{\Gamma_i(q,1,w)\subseteq\Gamma_j(q,d,w)} p_j^{(d)}.$$

It is easy to see that it is a probability measure on the state space $\Omega(q,1,w)$. Moreover, for any formula $\psi \in \Phi(1,0,w)$,

$$\alpha_\psi^{\Gamma(q,1,w)} = l_\psi^{S(q,d,w)} \text{ and } \beta_\psi^{\Gamma(q,1,w)} = m_\psi^{S(q,d,w)}.$$

This also implies that the above defined $T(\Gamma(q,1,d)$ satisfied the following property:

- if $\alpha_\psi^{S(q,1,w)} = \beta_\psi^{S(q,1,w)}$, then $T(\Gamma_1(q,d,w))([\psi]) = \alpha_\psi^{S(q,1,w)}$;
- if $\alpha_\psi^{S(q,1,w)} < \beta_\psi^{S(q,1,w)}$, then $\alpha_\psi^{S(q,1,w)} < T(\Gamma_1(q,d,w))([\psi]) < \beta_\psi^{S(q,1,w)}$;

For other atoms $\Gamma_i(p,1,w)(i \neq 1)$, we define the probability measures at these atoms according to their maximal consistent extensions Γ_i^∞ in the infinite language obtained through Γ_i^+, which is a maximal consistent extension in the language $\mathcal{L}(q,1,w)$, by the construction method in our completeness proof. For each atom $\Gamma_i(q,1,w)$ and each propositional letter p, define $\nu(p) := \{\Gamma_i(q,1,w) \in \Omega(q,1,w) : p \in \Gamma_i(q,1,w)\}$. Then we get a type space $M(q,1,w) := \langle \Omega(q,1,w), 2^{\Omega(q,1,w)}, T, \nu\rangle$. By a similar induction (to that in the proof of the truth lemma in Σ_+), we can show the following claim:

CLAIM 3.2.3. *For any atom $\Gamma_i(q,1,w)$ and any formula $\phi \in \Phi(q,1,w)$,*

$$M(q,1,w), \Gamma_i(q,1,w) \models \phi \textit{ iff } \phi \in \Gamma_i(q,1,w).$$

Here we only prove this claim at the atom $\Gamma_i(q,1,w)$. We only show the case that $\phi := L_r\psi$. Other cases are straightforward. Assume that $M(q,1,w), \Gamma_1(q,1,w) \models \phi$. It implies that $T(\Gamma_1(q,1,w))([[\psi]]) \geq r$. If $\alpha_\psi^{\Gamma_1(q,1,w)} = \beta_\psi^{\Gamma_1(q,1,w)}$, then, by induction hypothesis, $T(\Gamma_1(q,1,w))([[\psi]]) = T(\Gamma_1(q,1,w))([\psi])$ and $l_\psi^{S(q,d,w)} = m_\psi^{S(q,d,w)}$. Since $l_\psi^{S(q,d,w)} \leq T(\Gamma_1(q,1,w))([\psi]) \leq m_\psi^{S(q,d,w)}$, $T(\Gamma_1(q,1,w))([[\psi]]) = \alpha_\psi^{\Gamma_1(q,1,w)}$. It follows that $L_r\psi \in \Gamma_1(q,1,w)$. If $\alpha_\psi^{\Gamma_1(q,1,w)} < \beta_\psi^{\Gamma_1(q,1,w)}$, $\alpha_\psi^{\Gamma_1(q,1,w)} < T(\Gamma_1(q,1,w))([[\psi]]) < \beta_\psi^{\Gamma_1(q,1,w)}$ since $T(\Gamma_1(q,1,w))([[\psi]]) = T(\Gamma_1(q,1,w))([\psi])$ (by I. H.). It follows from the assumption that $T(\Gamma_1(q,1,w))([[\psi]]) \geq r$ that $r \leq \alpha_\psi^{\Gamma_1(q,1,w)}$. Hence $L_r\psi \in \Gamma_1(q,1,w)$. For the other direction, assume that $L_r\psi \in \Gamma_1(q,1,w)$. This implies that $r \leq \alpha_\psi^{\Gamma_1(q,1,w)}$. Since $\alpha_\psi^{\Gamma_1(q,1,w)} \leq T(\Gamma_1(q,1,w))([\psi])$ and $T(\Gamma_1(q,1,w))([[\psi]]) = T(\Gamma_1(q,1,w))([\psi])$, $T(\Gamma_1(q,1,w))([[\psi]]) \geq r$. That is to say, $M(q,1,w), \Gamma_1(q,1,r) \models L_r\psi$.

CLAIM 3.2.4. $\Xi(q,1,w)$ *is satisfiable at* $\Gamma_1(q,1,w)$.

Given any formula $\phi \in \Phi(q,1,w)$, we need to show that both $L_{l_\psi^{S(q,d,w)}}\phi$ and $M_{l_\psi^{S(q,d,w)}}\phi$ are satisfiable at $\Gamma_1(q,1,w)$. By the above claim, we know that $[[\phi]] = [\psi]$. This implies that $T(\Gamma_1(q,1,w))([[\phi]]) = T(\Gamma_1(q,1,w))([\phi]) = \sum_{\phi\in\Gamma_j(q,d,w)} p_j^{(d)}$, which is greater than $l_\psi^{S(q,1,w)}$. So $L_{l_\psi^{S(q,d,w)}}\phi$ is satisfiable at the atom $\Gamma_1(q,d,w)$. Similarly, we can show that $M_{l_\psi^{S(q,d,w)}}\phi$ is satisfiable at $\Gamma_1(q,1,w)$.

From the above claim, we know that $\Xi(q,1,w)\bigcup\Gamma_1(1,0,w)$, which is just the set of formulas of depth 0, is satisfiable at $\Gamma_1(q,1,w)$ in the type space $M(q,1,w)$. Therefore, $\Xi(q,1,w)\bigcup\Gamma_1(1,0,w)$ is consistent and is contained in a maximal consistent extension $\Gamma(q,2,w)$ which is an atom in $\Omega(q,2,w)$. By repeating the above argument, we can construct a canonical type space $M(q,2,w)$ where the probability measure at $\Gamma(q,2,w)$ agrees with the probability measure $P(q,d,w)$ such that $\Gamma(1,0,w)\bigcup\Xi(q,2,w)$ is satisfiable at $\Gamma(q,2,w)$. Hence $\Xi(q,2,w)$ is consistent. Eventually, we can find a maximal consistent set $\Gamma(q,d+1,w) \in \Omega(q,d+1,w)$ such that $\Gamma(1,0,w)\bigcup\Xi(q,d,w)$ is included in $\Gamma(q,d+1,w)$. So we have shown that $\Xi(q,d,w)$ is consistent.

□

COROLLARY 3.2.5. *For any consistent formula* ϕ *of degree 0, i.e.* $\phi \in \Phi(1,0,w)$, $\{\phi\} \cup \Xi(q,d,w)$ *is consistent.*

PROOF. Fix the consistent formula ϕ of degree 0. It is contained in a maximal consistent extension $\Gamma(1,0,w) \in \Omega(1,0,w)$. We define a probability measure $T(\Gamma(1,0,w))$ at $\Gamma(1,0,w)$ that agrees with the probability measure $P(q,d,w)$. Just as above, we can define a canonical type space on the state space $\Omega(1,0,w)$. By repeating this argument, we can get a maximal consistent extension $\Gamma(q,d+1,w) \in \Omega(q,d+1,w)$ that includes $\Xi(q,d,w)$ and $\Gamma(1,0,w)$ which contains ϕ. So $\{\phi\} \cup \Xi(q,d,w)$ is consistent.

□

COROLLARY 3.2.6. *For any atom $\Gamma(q,d,w) \in \Omega(q,d,w)$, the propositional part and the probability part of its normal form are independent of each other in the sense that, for any consistent formula ϕ, the conjunction of the probability part and ϕ is consistent.*

Fix an index vector (q,d,w). For this index vector, we have a canonical type space $M(q,d,w) := \langle \Omega(q,d,w), 2^{\Omega(q,d,w)}, T, \nu \rangle$.[1] We have already shown the following truth lemma: for any formula $\phi \in \Phi(q,d,w)$ and $\Delta(q,d,w) \in \Omega(q,d,w)$,

$$M(q,d,w), \Delta(q,d,w) \models \phi \text{ iff } \phi \in \Delta(q,d,w).$$

Enumerate all elements in $\Omega(q,d,w)$:

$$\Gamma_1(q,d,w), \Gamma_2(q,d,w), \cdots, \Gamma_{N(q,d,w)}(q,d,w).$$

Now we consider guided maximal extensions of $\Gamma_1(q,d,w)$. Let $(q',d',w') \succ (q,d,w)$. Next we define a probability measure $P(q',d',w')$ on the state space $\Omega(q',d',w')$ satisfying the following conditions:

(1) $\sum_{\Delta(q',d',w') \in \Omega(q',d',w')} P(q',d',w')(\Delta(q',d',w')) = 1$;
(2) for each atom $\Gamma_i(q,d,w) \in \Omega(q,d,w)(1 \leq i \leq N(q,d,w)), T(\Gamma_1(q,d,w))(\Gamma_i(q, d,w)) = \sum_{\Gamma_i(q,d,w) \subseteq \Delta(q',d',w')} P(q',d',w')(\Delta(q',d',w'))$.

The first condition makes sure that such defined $P(q',d',w')$ is a probability measure on $\Omega(q',d',w')$. The second condition guarantees that atoms in $\Omega(q',d',w')$ with a probability measure satisfying this condition are maximal consistent extensions of $\Gamma_1(q,d,w)$. Define $S(q',d',w') := \langle \Omega(q',d',w'), 2^{\Omega(q',d',w')}, P(q',d',s') \rangle$. Just as above, we can define, for any formula $\phi \in \Phi(q',d',w')$, $l_\phi^{S(q',d',w')}$, $m_\phi^{S(q',d',w')}$ and $\Xi(q',d',w')$ similarly.

THEOREM 3.2.7. *(Guided Maximal Consistent Extension Theorem) The proposition consists of the following two parts:*

- *For the atom $\Gamma_1(q,d,w) \in \Omega(q,d,w)$, there is a maximal consistent extension $\Gamma(q',d'+1,w')$ such that $\Xi(q',d',w') \cup \Gamma_1(q,d,w) \subseteq \Gamma(q',d'+1,w')$.*

[1]Since T is not unique, such a canonical type space on $\Omega(q,d,w)$ is not unique, either.

- *Let* $\Gamma(q', d', w') = \Gamma(q', d' + 1, w') \cap \Phi(q', d', w')$. *If we set the probability measure at* $\Gamma(q', d', w')$: $T(\Gamma(q', d', w')) = P(q', d', w')$ *and define the probability measures at other atoms in* $\Omega(q', d', w')$ *by a similar procedure to that in the proof of the completeness of* Σ_+, *such a defined canonical type space* $M(q', d', w') := \langle \Omega(q', d', w'), 2^{\Omega(q', d', w')}, T, \nu \rangle$ *satisfies the following property (truth lemma): for any formula* $\phi \in \Phi(q', d', w')$ *and* $\Delta(q', d', w') \in \Omega(q', d', w')$,

$$M(q', d', w'), \Delta(q', d', w') \models \phi \textit{ iff } \phi \in \Delta(q', d', w').$$

PROOF. Let γ_0 be the conjunction of all formulas of degree 0 in the set $\Gamma_1(q, d, w)$. By the corollary 3.2.5, we know that $\{\gamma_0\} \cup \Xi(q', d', w')$ is contained in an atom $\Gamma(q', d'+1, w') \in \Omega(q', d' + 1, w')$. In order to show that $\Xi(q', d', w') \cup \Gamma_1(q, d, w) \subseteq \Gamma(q', d' + 1, w')$, we only need to show that $\Gamma_1(q, d, w) \subseteq \Gamma(q', d' + 1, w')$. But this is guaranteed by the second condition of our definition of $P(q', d', w')$. According to the above lemma 3.2.1, it suffices to show that each conjunct of the normal form of $\Gamma_1(q, d, w)$ is contained in $\Gamma(q', d' + 1, w')$. Since the propositional part of the normal form is tautologically equivalent to γ_0, it is contained in $\Gamma(q', d' + 1, w')$.

CLAIM 3.2.8. *Each probability formula* $L_r\phi$ *which is a conjunct of the probability part of the normal form is also in* $\Gamma(q', d' + 1, w')$.

Assume that $L_r\phi$ is a conjunct of the probability part of the normal form.

$$\begin{aligned} T(\Gamma_1(q, d, w)([\phi])) &= \sum_{\phi \in \Delta(q,d,w)(\in \Omega(q,d,w))} T(\Gamma_1(q, d, w))(\Delta(q, d, w)) \\ &= \sum_{\phi \in \Delta(q,d,w)(\in \Omega(q,d,w))} \sum_{\Delta(q,d,w) \in \Delta(q',d',w')(\in \Omega(q',d',w'))} P(q', d', w')(\Delta(q', d', w')) \\ &= \sum_{\phi \in \Delta(q',d',w')} P(q', d', w')(\Delta(q', d', w')) \end{aligned}$$

Observe that $\alpha_\phi^{\Gamma_1(q,d,w)}$ is the largest multiple of $\frac{1}{q}$ that is less than or equal to $T(\Gamma_1(q, d, w)$ $([\phi])$, which is equal to $l_\phi^{S(q',d',w')}$ because $T(\Gamma_1(q, d, w)([\phi])) = \sum_{\phi \in \Delta(q',d',w')} P(q', d', w')$

$(\Delta(q',d',w'))$. Similarly, we can show that $\beta_\phi^{\Gamma_1(q,d,w)} = m_\phi^{S(q',d',w')}$. It follows immediately that $L_{\alpha_\phi^{\Gamma_1(q,d,w)}}\phi \in \Gamma(q',d'+1,w')$. Since $L_r \in \Gamma_1(q,d,w)$, $r \leq \alpha_\phi^{\Gamma_1(q,d,w)}$ and hence $L_r\phi \in \Gamma(q',d'+1,w')$. We finish the proof the claim and hence show that $\Gamma_1(q,d,w) \subseteq \Gamma(q',d'+1,w')$.

For the second part, we only need to show that, for any formula $\phi \in \Phi(q',d',w')$ and any atom $\Gamma_i(q',d',w') \in \Omega(q',d',w')(1 \leq i \leq N(q',d',w'))$,

(1) if $\alpha_\phi^{\Gamma_i(q',d',w')} = \beta_\phi^{\Gamma_i(q',d',w')}$, then $T(\Gamma_i(q',d',w'))([\phi]) = \alpha_\phi^{\Gamma_i(q',d',w')}$; and
(2) if $\alpha_\phi^{\Gamma_i(q',d',w')} < \beta_\phi^{\Gamma_i(q',d',w')}$, then $\alpha_\phi^{\Gamma_i(q',d',w')} < T(\Gamma_i(q',d',w'))([\phi]) < \beta_\phi^{\Gamma_i(q',d',w')}$.

We may assume that $\Gamma(q',d',w') = \Gamma_1(q',d',w')$. Here we only consider $\Gamma(q',d',w')$. According to our above argument, we know that $\alpha_\phi^{\Gamma(q',d',w')} = l_\phi^{S(q',d',w')}$ and $\beta_\phi^{\Gamma(q',d',w')} = m_\phi^{S(q',d',w')}$. If $\alpha_\phi^{\Gamma(q',d',w')} = \beta_\phi^{\Gamma(q',d',w')}$, then $T(\Gamma(q',d',w'))([\phi]) = l_\phi^{S(q',d',w')} = m_\phi^{S(q',d',w')} = \alpha_\phi^{\Gamma(q',d',w')}$. If $\alpha_\phi^{\Gamma(q',d',w')} < \beta_\phi^{\Gamma(q',d',w')}$, then $l_\phi^{S(q',d',w')} < m_\phi^{S(q',d',w')}$. It follows immediately that $l_\phi^{S(q',d',w')} < T(\Gamma(q',d',w'))([\phi]) < m_\phi^{S(q',d',w')}$. That is to say, $\alpha_\phi^{\Gamma(q',d',w')} < T(\Gamma(q',d',w'))([\phi]) < \beta_\phi^{\Gamma(q',d',w')}$

□

COROLLARY 3.2.9. *Each atom Γ in $\Omega(q,d,w)$ has at least two descendants (or maximal consistent extensions) in $\Omega(q,d+1,w)$.*

PROOF. Assume that $\Gamma \in \Omega(q,d,w)$. Then there is an atom $\Gamma_{d-1} \in \Omega(q,d,w)$ such that $\neg M_0(\bigwedge \Gamma_{d-1}) \in \Gamma$. For, otherwise, $M_0\top \in \Gamma$ and hence $\bot \in \Gamma$. Let Γ_d be a maximal consistent extension of Γ_{d-1} in $\Omega(q,d,w)$. According to the above theorem, both $\Gamma \cup \{M_0(\bigwedge \Gamma_d)\}$ and $\Gamma \cup \{\neg M_0(\bigwedge \Gamma_d)\}$ are consistent. So Γ has at least two descendants.

□

We have already known that the following set is consistent:

$$C := \{\neg M_0 p_1\} \cup \{M_{\frac{1}{2^n}} p_1 : n \in \mathcal{N}\}.$$

So there is a maximal consistent set Γ^∞ that contains C as a subset. Recall that Γ^∞ does not have a probability model. Fir any index vector (q,d,w), define

$$\Gamma(q,d,w) := \Gamma^\infty \cap \Phi(q,d,w).$$

Especially we have a sequence of maximal consistent sets in the local language $\mathcal{L}(2^n, 2^n, 2^n)$:

$$\Gamma(1,1,1) \subseteq \Gamma(2,2,2) \subseteq \Gamma(2^2,2^2,2^2) \subseteq \cdots \subseteq \Gamma(2^n,2^n,2^n) \subseteq \cdots$$

such that $\bigcup_n \Gamma(2^n,2^n,2^n) = \Gamma^\infty$. Now we apply the above theorem about guided maximal consistent extensions to find an admissible maximal $\mathcal{L}$-consistent set Γ_1 such that

$$\tfrac{1}{4} = T(\Gamma_1)([\Gamma(2,2,2)]) = T(\Gamma_1)([\Gamma(2^2,2^2,2^2)]) = \cdots = T(\Gamma_1)([\Gamma(2^n,2^n,2^n)]) = \cdots$$

First we consider the algebra $(\Omega(1,1,1), 2^{\Omega(1,1,1)})$. Since $\Gamma(1,1,1)$ is a maximal $\mathcal{L}(1,1,1)$-consistent, $\Gamma(1,1,1) \in \Omega(1,1,1)$. Set an arbitrary probability measure $P(1,1,1)$ on this algebra such that:

- $P(1,1,1)(\Gamma(1,1,1)) = \frac{1}{4}$;
- For all atoms Δ in $\Omega(1,1,1)$, $P(1,1,1)(\Delta)$ is a multiple of $\frac{1}{2^{m_1}}$ for some positive natural number m_1.

By applying the above theorem, we know that there is a maximal $\mathcal{L}(2^{m_1}, 2^{m_1}, 2^{m_1})$-consistent set $\Gamma_{2^{m_1}} \in \Omega(2^{m_1}, 2^{m_1}, 2^{m_1})$ such that

- for any formula ϕ in the language $\mathcal{L}(1,1,1)$, $\alpha_\phi^{\Gamma_{2^{m_1}}} = \beta_\phi^{\Gamma_{2^{m_1}}} = P(1,1,1)([\phi])$; and
- $\alpha^{\Gamma_{2^{m_1}}}_{\bigwedge \Gamma(1,1,1)} = \beta^{\Gamma_{2^{m_1}}}_{\bigwedge \Gamma(1,1,1)} = \frac{1}{4}$;

Now consider the algebra $(\Omega(2^{m_1}, 2^{m_1}, 2^{m_1}), 2^{\Omega(2^{m_1}, 2^{m_1}, 2^{m_1})})$. Note that $(\Omega(1,1,1), 2^{\Omega(1,1,1)})$ is a subalgebra of $(\Omega(2^{m_1}, 2^{m_1}, 2^{m_1}), 2^{\Omega(2^{m_1}, 2^{m_1}, 2^{m_1})})$. Set a probability measure $P(2^{m_1}, 2^{m_1}, 2^{m_1})$ on this algebra that satisfies the following conditions:

- $P(2^{m_1}, 2^{m_1}, 2^{m_1})$ agrees with $P(1,1,1)$ on $2^{\Omega(1,1,1)}$;
- $P(2^{m_1}, 2^{m_1}, 2^{m_1})(\Gamma(2^{m_1}, 2^{m_1}, 2^{m_1})) = \frac{1}{4}$;
- for all atoms Δ in $\Omega(2^{m_1}, 2^{m_1}, 2^{m_1})$, $P(2^{m_1}, 2^{m_1}, 2^{m_1})(\Delta)$ is a multiple of $\frac{1}{2^{m_2}}$ for some $m_2 > m_1$.

By applying the above theorem, we know that there is a maximal $\mathcal{L}(2^{m_2}, 2^{m_2}, 2^{m_2})$-consistent set $\Gamma_{m_2} \in \Omega(2^{m_2}, 2^{m_2}, 2^{m_2})$ such that

- for any formula ϕ in the language $\mathcal{L}(2^{m_1}, 2^{m_1}, 2^{m_1})$, $\alpha_\phi^{\Gamma_{2^{m_2}}} = \beta_\phi^{\Gamma_{2^{m_2}}} = P(2^{m_1}, 2^{m_1}, 2^{m_1})([\phi])$;

- $\alpha^{\Gamma_{2^{m_2}}}_{\bigwedge\Gamma(2^{m_1},2^{m_1},2^{m_1})} = \beta^{\Gamma_{2^{m_2}}}_{\bigwedge\Gamma(2^{m_1},2^{m_1},2^{m_1})} = \frac{1}{4}$;
- $\Gamma_{2^{m_1}} \subseteq \Gamma_{2^{m_2}}$.

By repeating above process, we get a sequence of maximal $\mathcal{L}(2^{m_i}, 2^{m_i}, 2^{m_i})$-consistent set $\Gamma_{2^{m_i}}$:

$$\Gamma_{2^{m_1}} \subseteq \Gamma_{2^{m_2}} \subseteq \cdots$$

that satisfies the following conditions:

- for each $\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})$, $\alpha^{\Gamma_{2^{m_i}}}_{\wedge\Gamma(2^{m_i},2^{m_i},2^{m_i})} = \beta^{\Gamma_{2^{m_i}}}_{\wedge\Gamma(2^{m_i},2^{m_i},2^{m_i})} = \frac{1}{4}$;
- any formula $\phi \in \Phi(2^{m_i}, 2^{m_i}, 2^{m_i})$, $\alpha^{\Gamma_{2^{m_i}}}_{\phi} = \beta^{\Gamma_{2^{m_i}}}_{\phi}$

Define $\Gamma_1 := \bigcup_i \Gamma_{2^{m_i}}$. From the above observations, it follows that Γ_1 is an admissible maximal consistent set in $\mathcal{L}$. According to our definition on canonical models,

$$\tfrac{1}{4} = T(\Gamma_1)([\Gamma(2,2,2)]) = T(\Gamma_1)([\Gamma(2^2,2^2,2^2)]) = \cdots = T(\Gamma_1)([\Gamma(2^n,2^n,2^n)]) = \cdots$$

LEMMA 3.2.10. $\bigcap_i[\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})] = \emptyset$.

PROOF. Suppose that $\bigcap_i[\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})] \neq \emptyset$. Assume that $\Delta \in \bigcap_i[\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})]$. That is to say, $\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i}) \subseteq \Delta$. It follows that $\{\neg M_0 p_1\} \cup \{M_{\frac{1}{2^n}} p_1 : n \in \mathcal{N}\} \subseteq \Delta$. So Δ is not admissible. This contradicts the assumption that Δ is an admissible maximal consistent set in the language $\mathcal{L}$.

□

THEOREM 3.2.11. *For such a Γ_1, $T(\Gamma_1)$ is finitely additive but not σ-additive.*

PROOF. Suppose that $T(\Gamma_1)$ were sigma-additive. This would imply that

$$T(\Gamma_1)(\textstyle\bigcap_i[\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})]) = \lim_{i\to\infty} T(\Gamma_1)([\Gamma(2^{m_i}, 2^{m_i}, 2^{m_i})]) = \frac{1}{4}.$$

But this contradicts the fact that $T(\Gamma_1)(\emptyset) = 0$.

□

COROLLARY 3.2.12. *The canonical model defined on the set of admissible sets of formulas is not a probability model.*

THEOREM 3.2.13. *All satisfiable maximal consistent sets of formulas form a probability model. Moreover, it is the biggest (or universal) probability model in the sense that any other probability model can be embedded into it.*

PROOF. The following proof is adapted from [4] or [18] Assume that Ω is the set of all satisfiable maximal consistent sets of formulas. $[\phi]$ denotes the set $\{s \in \Omega : \phi \in s\}$. Obviously, the set of $[\phi]$'s is an algebra, which is denoted as $\mathcal{A}_0$. $\mathcal{A}$ denotes the σ-algebra generated by $\mathcal{A}_0$. For any $s \in \Omega$, define $T(s)([\phi]) = sup\{r \in [0,1] : L_r\phi \in s\}$.

CLAIM 3.2.14. *If s is satisfiable at some state w of a probability model $M = \langle \Omega_M, \mathcal{A}_M, T_M, v_M \rangle$, then, for any formula, $T(s)([\phi]) = T_M(w)([[\phi]])$ where $[[\phi]] = \{w' \in \Omega_M : M, w' \models \phi\}$.*

But this is straightforward because both of them are equal to $sup\{r \in [0,1] : L_r\phi \in s\}$. Moreover, since $T_M(w)$ is a probability measure on the algebra that consists of sets of the form $[[\phi]]$, $T(s)$ is also a probability measure on the algebra $\mathcal{A}_0$. Hence it can be uniquely extended to a probability measure on the σ-algebra $\mathcal{A}$. In order to prove that such defined T is a type function, it remains to show that $T(\cdot, A)$ is $\mathcal{A}$-measurable for any $A \in \mathcal{A}$.

$B^r(A)$ denotes the set $\{w \in \Omega : T(w, A) \geq r\}$ for $r \in [0,1]$ and $A \in \mathcal{A}$. Define $\mathcal{A}' = \{A \in \mathcal{A} : B^r(A) \in \mathcal{A}, \forall r \in [0,1]\}$. It is easy to see that $B^r(A) \in \mathcal{A}$ if $A \in \mathcal{A}_0$. In other words, $\mathcal{A}' \supseteq \mathcal{A}_0$.

CLAIM 3.2.15. *$\mathcal{A}'$ is a monotone class.*

Let $(A_i)_i$ be a decreasing sequence of events in $\mathcal{A}'$. Since $B^r(\bigcap_i A_i) = \bigcap_i B^r(A_i)$, $\bigcap_i A_i \in \mathcal{A}'$. So $\mathcal{A}'$ is closed under the formation of monotone intersections. Let $(A_i)_i$ be an increasing sequence of event in $\mathcal{A}'$. Also we have that $\bigcup_i A_i \in \mathcal{A}'$. Indeed, $B^r(\bigcup_i A_i) = \bigcap_n \bigcup_m B^{r-\frac{1}{n}}(A_m) \in \mathcal{A}$. So $\mathcal{A}'$ is also closed under the formation of monotone unions.

CLAIM 3.2.16. *$\mathcal{A}' = \mathcal{A}$ and hence $T(\cdot, A)$ is $\mathcal{A}$-measurable for any $A \in \mathcal{A}$.*

According to Halmos's monotone class theorem, $\mathcal{A} \subseteq \mathcal{A}'$. So $\mathcal{A} = \mathcal{A}'$. In other words, for any $A \in \mathcal{A}$ and any rational $r \in [0, 1]$, $B^r(A) \in \mathcal{A}$. This implies that $T(\cdot, A)$ is $\mathcal{A}$-measurable for any $A \in \mathcal{A}$.

Next we show that it is the biggest probability model. Let $M' = \langle \Omega', \mathcal{A}', T', v' \rangle$ be a probability model. Define $d : \Omega' \to \Omega$ as $d(s') = \{\phi : M', s' \models \phi\}$. Note that $d(s') \in \Omega$. It follows that, for any $s' \in \Omega'$, $M', s' \models L_r\phi$ if and only if $L_r\phi \in d(s')$ and hence $M', s' \models \psi$ if and only if $\psi \in d(s')$. Indeed, d is a type morphism from M' to its image in M. So we have finished the proof of the whole theorem.

□

COROLLARY 3.2.17. *The set of satisfiable maximal consistent sets of formulas is a proper subset of the set of admissibles.*

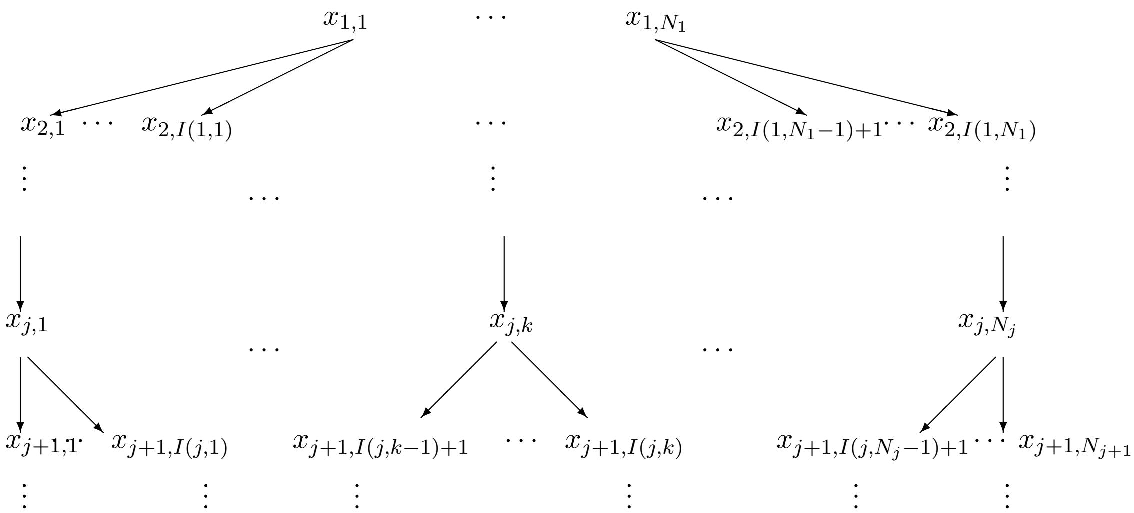
PROOF. The γ_1 that we constructed in the above counterexample is admissible but not satisfiable. For, otherwise, it should define a probability measure on the algebra consisting of $[\phi]$'s. But we have shown that this is impossible.

□

3.3. No Compactness or Strong Completeness for Probability Logics with Finite Indices

One might expect that if we would weaken the language of our logic to include only finite indices, we could prove a strong completeness. With strong completeness, we could easily prove the completeness of Σ_+ plus some other higher-order probability formulas with respect to their corresponding class of type spaces. Moreover, with strong completeness, we could give a syntactic definition of consequence in this restricted language for Aumann's work on interactive epistemology: probability. All of his results there would also be proved in this language with finite indices. However, in this section, we show that there is no compactness or strong completeness for probability logics with finite indices. So this approach with a restricted syntax towards interactive epistemology does not work well. First we define linear systems of probabilities over trees. Next we construct such a system that is not compact. In the third part, we show that, whenever any consistent set of formulas contains this noncompact linear system, it is finitely satisfiable but not satisfiable.

3.3.1. System of Linear Inequalities and Equations of Probabilities. We define I to be a function from $\mathcal{N}^+ \times \mathcal{N}^+$ to $\mathcal{N}^+$ that is strictly increasing in the second coordinate, where $\mathcal{N}^+$ is the set of positive natural numbers. N_j is defined inductively: $N_j = I(j-1, N_{j-1})$ for $j \geqslant 2$. Now we consider a genealogy tree $\mathcal{T}$. X is the set of nodes. The first generation consists of $x_{1,1}, \cdots, x_{1,N_1}$, which are independent of each other. Each node in the tree has at least one descendant, which is indicated by arrows in the following graph. The population in the j-th generation is N_j. Note that I is the function to determine the number of descendants. $\mathcal{T}$ can be illustrated as follows:



A *term* t is a finite subset of X such that any two elements are not in the same branch. Its *family set* consists of the elements in t and all their descendants. The *order* of t is defined to be $max\{j : x_{j,k} \in t\}$, i.e., the lowest generation order of elements in t. It is denoted by $o(t)$. Two terms t and t' are *equal* if their family sets at the $max\{o(t), o(t')\}th$ generation are the same. We denote $t \equiv t'$.

Remark. We can interpret the above definitions in the context of probability theory. Ω denotes the set of all (infinite) branches of the tree $\mathcal{T}$. Each branch is represented by an infinite sequence of nodes. For example, the leftmost branch is $(x_{1,1}, x_{2,1}, x_{3,1}, \cdots)$. Define $\mathcal{T}_n$ as the subtree of $\mathcal{T}$ which consists of all notes of depth $\leq n$ and Ω_n the set of (finite) branches in $\mathcal{T}_n$. Note that any branch in Ω_n is represented by a finite sequence of length n. For example, the leftmost branch in this tree is $(x_{1,1}, x_{2,1}, \cdots, x_{n,1})$. $z_k : \Omega \to X$ is *the k-th coordinate function.* A *cylinder of rank n* is a set of the form:

$$A = \{w \in \Omega : (z_1(w), z_2(w), \cdots, z_n(w)) \in H\} \text{ for some } H \subseteq \Omega_n.$$

$\mathcal{A}_0$ denotes the set of cylinders of all ranks. It is easy to check that $\mathcal{A}_0$ is a field. Now we can interpret our above definition of terms in this "generalized sequence space". Cylinders of rank n corresponds exactly to terms of order n. For example, let $I(1,1) = 2$, the cylinder $\{w \in \Omega : (z_1(w), z_2(w)) = (x_{1,1}, x_{2,1})\} \cup \{w \in \Omega : (z_1(w), z_2(w), z_3(w)) = (x_{1,1}, x_{2,2}, x_{3,I(2,1)+1})\}$ corresponds to the term $\{x_{2,1}, x_{3,I(2,1)+1}\}$.

Let $t = \{x_{j_1,k_1}, \cdots, x_{j_m,k_m}\}$ and $t' = \{x_{j'_1,k'_1}, \cdots, x_{j'_n,k'_n}\}$. Since we will associate each node with a probability, we simply write $t \equiv t'$ more explicitly as:

$$(*)\text{: } x_{j_1,k_1} + \cdots + x_{j_m,k_m} = x_{j'_1,k'_1} + \cdots + x_{j'_n,k'_n}$$

Let $\mathcal{S}_=$ denote the set of all such linear equations (*) as well as the following:

$$x_{j,1} + \cdots + x_{j,N_j} = 1 \text{ for all } j \geq 1.$$

Note that $\mathcal{S}_=$ is uniquely determined by the genealogy tree. $\mathcal{S}_0$ is the set of all linear inequalities:

$$0 \leq x_{j,k} \leq 1 \text{ for all } j \geq 1 \text{ and } k \leq N_j.$$

Fix a natural number $p \geq 2$. This is a crucial assumption for our following conjecture. Define $\mathcal{Q}_p := \{\frac{q}{p} : p, q \in \mathcal{N}^+\} \cap [0,1]$. Let $\mathcal{S}_\leq$ be a countable (possibly countably infinite) set of linear inequalities which is one of the following forms:

- $\frac{q}{p} < x_{j_1,k_1} + \cdots + x_{j_m,k_m} < \frac{q+1}{p}$ for some term $t = \{x_{j_1,k_1}, \cdots, x_{j_m,k_m}\}$ and some natural number q such that $\frac{q}{p}, \frac{q+1}{p} \in \mathcal{Q}_p$;
- $x_{j_1,k_1} + \cdots + x_{j_m,k_m} = \frac{q'}{p}$ for some q' such that $\frac{q'}{p} \in \mathcal{Q}_p$.

$S_{\leq}$ can be regarded as a constraint set. Let $\mathcal{S}$ denote the union of $\mathcal{S}_{=}, \mathcal{S}_0$ and $\mathcal{S}_{\leq}$. Any subset of the linear system $\mathcal{A}$ is called a *linear probability tree system.*

REMARK. The first part of $\mathcal{S}_{=}$ says that each cylinder may have different representations in terms of *terms* and probability measures on the field $\mathcal{A}_0$ is finitely additive. The second part tells us that, since all cylinders of the form $\{w \in \Omega : (z_1(w), \cdots, z_n(w)) \in \Omega_n\}$ is the same as Ω, they must have the probability 1. The system $\mathcal{S}_0$ says that probabilities of cylinders should be between 0 and 1. In short, both $\mathcal{S}_{=}$ and $\mathcal{S}_0$ make sure that the solution to the system $\mathcal{S}$ is a (finitely additive) probability measure on the field $\mathcal{A}_0$ of cylinders. $\mathcal{S}_{\leq}$ is the essential part of $\mathcal{S}$. Given a finite set $\mathcal{Q}_p$ of rationals, it says that for any cylinder, we only know that its probability is either equal to one rational from $\mathcal{Q}_p$ or is strictly between two consecutive numbers from $\mathcal{Q}_p$.

3.3.2. No Compactness for Linear System of Probabilities over Trees. In this subsection, we construct a linear probability tree system that is not compact. In the following, we fix $p = 2$. Then $\mathcal{Q}_2 = \{0, 1/2, 1\}$.

LEMMA 3.3.1. *There is a linear probability tree system $\mathcal{S}_2$ such that*

- *all indices in $\mathcal{S}_2$ are in $\mathcal{Q}_2$;*
- *$\mathcal{S}_2$ has a real solution;*
- *$0 < x_1 < \frac{1}{32}$ is deducible from $\mathcal{S}_2$.*

PROOF. We assume that $\mathcal{S}$ includes $0 < x_1$. In order to get $x_1 < \frac{1}{32}$, it suffices to show that, for some x_2,

$$x_1 + x_2 < \frac{1}{16} \tag{3.1}$$

$$x_1 - x_2 < 0 \tag{3.2}$$

Note that neither of these two inequalities are specified by our language above. For the first inequality, it suffices to show that, for some x_3,

$$x_1 + x_2 + x_3 < \frac{1}{8} \tag{3.3}$$

$$x_1 + x_2 - x_3 < 0 \tag{3.4}$$

Similarly, for the first part, it suffices to show that, for some x_4,

$$x_1 + x_2 + x_3 + x_4 < \frac{1}{4} \tag{3.5}$$

$$x_1 + x_2 + x_3 - x_4 < 0 \tag{3.6}$$

Similarly, for the first part, it suffices to show that, for some x_5,

$$x_1 + x_2 + x_3 + x_4 + x_5 < \frac{1}{2} \tag{3.7}$$

$$x_1 + x_2 + x_3 + x_4 - x_5 < 0 \tag{3.8}$$

Note that the first one is within our language. Next we deal with all these inequalities,

$$x_1 + x_2 + x_3 + x_4 - x_5 < 0 \tag{3.9}$$

$$x_1 + x_2 + x_3 - x_4 < 0 \tag{3.10}$$

$$x_1 + x_2 - x_3 < 0 \tag{3.11}$$

$$x_1 - x_2 < 0 \tag{3.12}$$

These four inequalities are deducible from the following eight inequalities:

$$x_1 + x_2 + x_3 + x_4 + x_5^{(5)} < \frac{1}{2} \tag{3.13}$$

$$\frac{1}{2} < x_5 + x_5^{(5)} \tag{3.14}$$

$$x_1 + x_2 + x_3 + x_5^{(5)} + x_4^{(5)} < \frac{1}{2} \tag{3.15}$$

$$\frac{1}{2} < x_4 + x_5^{(5)} + x_4^{(5)} \tag{3.16}$$

$$x_1 + x_2 + x_5^{(5)} + x_4^{(5)} + x_3^{(5)} < \frac{1}{2} \tag{3.17}$$

$$\frac{1}{2} < x_3 + x_5^{(5)} + x_4^{(5)} + x_3^{(5)} \tag{3.18}$$

$$x_1 + x_5^{(5)} + x_4^{(5)} + x_3^{(5)} + x_2^{(5)} < \frac{1}{2} \tag{3.19}$$

$$\frac{1}{2} < x_2 + x_5^{(5)} + x_4^{(5)} + x_3^{(5)} + x_2^{(5)} \tag{3.20}$$

For example, (9) is derivable from (13) and (14). $\mathcal{S}_2$ denotes the system consists of (13)-(20), (7) and the inequality: $0 < x_1$. It is easy to see that the system is within our language and $0 < x_1 < \frac{1}{32}$ is provable from $\mathcal{S}_2$. Finally, we need to show that $\mathcal{S}_2$ has a real solution. Consider the inequalities from (1) to (8), we assign the following vales to variables: $x_1 = \frac{1}{64}, x_2 = \frac{1}{32}, x_3 = \frac{1}{16}, x_4 = \frac{1}{8}, x_5 = \frac{1}{4}$. Consider the inequalities (13) and (14), it is desirable that $x_5^{(5)} = \frac{33}{128}$. Consider (15) and (16), it is natural to assign $x_4^{(5)} = \frac{16}{128}$. Similarly, we assign $x_3^{(5)} = \frac{8}{128}, x_2^{(5)} = \frac{4}{128}$. Indeed, these form a solution to the system $\mathcal{S}_2$.

□

THEOREM 3.3.2. *For any n, there is a linear probability tree system $\mathcal{S}_2$ such that*

- *all indices in $\mathcal{S}_2$ are in $\mathcal{Q}_2$;*
- *$\mathcal{S}_2$ has a real solution;*
- *$0 < x_1 < \frac{1}{2^n}$ is deducible from $\mathcal{S}_2$.*

PROOF. It follows from a similar algorithm proof as in the above lemma.

□

THEOREM 3.3.3. *There is a linear probability tree system $\mathcal{S}$ such that*

- *all indices in $\mathcal{S}$ are in $\mathcal{Q}_2$;*

- *$\mathcal{S}$ is finitely satisfiable;*
- *$0 < x_1 < \frac{1}{2^n}$ is deducible from $\mathcal{S}$ for all n.*

Proof. Here we demonstrate a typical step in the construction of our required $\mathcal{S}_2$. Define $\mathcal{S}_2^{(1)} = \{0 < x_1 < \frac{1}{2}\}$. $\mathcal{S}_2^{(2)}$ denotes the system consists of the following inequalities:

$$0 < x_1 \tag{3.21}$$

$$x_1 + x_2 < \frac{1}{2} \tag{3.22}$$

$$x_1 + x_2^{(2)} < \frac{1}{2} \tag{3.23}$$

$$\frac{1}{2} < x_2 + x_2^{(2)} \tag{3.24}$$

We already know that $\mathcal{S}_2^{(2)}$ has a real solution and $0 < x_1 < \frac{1}{2^2}$ is provable in this system. Define $\mathcal{S}^{(2)} = \mathcal{S}_2^{(1)} \cup \mathcal{S}_2^{(2)}$. Similarly, $\mathcal{S}_2^{(3)}$ denotes the system consists of the following inequalities:

$$0 < x_1 \tag{3.25}$$

$$x_1 + x_2 + x_3 < \frac{1}{2} \tag{3.26}$$

$$x_1 + x_2 + x_3^{(3)} < \frac{1}{2} \tag{3.27}$$

$$\frac{1}{2} < x_3 + x_3^{(3)} \tag{3.28}$$

$$x_1 + x_2^{(3)} + x_3^{(3)} < \frac{1}{2} \tag{3.29}$$

$$\frac{1}{2} < x_2 + x_2^{(3)} + x_3^{(3)} \tag{3.30}$$

We already know that $\mathcal{S}_2^{(3)}$ has a real solution and $0 < x_1 < \frac{1}{2^3}$ is provable in this system. Define $\mathcal{S}^{(3)}$ as the set consists of all elements in $\mathcal{S}^{(2)}$ and $\mathcal{S}_2^{(3)}$ as well as the following equation:

$$x_2^{(2)} = x_2^{(3)} + x_3^{(3)}.$$

It is easy to see that $\mathcal{S}^{(3)}$ has a real solution and $0 < x_1 < \frac{1}{8}$ is derivable there. $\mathcal{S}_2^{(4)}$ consists of the following linear inequalities:

$$\begin{aligned}
&(3.31) & 0 &< x_1 \\
&(3.32) & x_1 + x_2 + x_3 + x_4 &< \frac{1}{2} \\
&(3.33) & x_1 + x_2 + x_3 + x_4^{(4)} &< \frac{1}{2} \\
&(3.34) & \frac{1}{2} &< x_4 + x_4^{(4)} \\
&(3.35) & x_1 + x_2 + x_4^{(4)} + x_3^{(4)} &< \frac{1}{2} \\
&(3.36) & \frac{1}{2} &< x_3 + x_4^{(4)} + x_3^{(4)} \\
&(3.37) & x_1 + x_4^{(4)} + x_3^{(4)} + x_2^{(4)} &< \frac{1}{2} \\
&(3.38) & \frac{1}{2} &< x_2 + x_4^{(4)} + x_3^{(4)} + x_2^{(4)}
\end{aligned}$$

${\mathcal{S}'}_2^{(4)}$ is gotten from $\mathcal{S}_2^{(4)}$ by replacing $x_2^{(4)}$ by $x_2^{(3)}$. Define $\mathcal{S}^{(4)}$ as the set consisting of all inequalities in $\mathcal{S}^{(3)}$ and ${\mathcal{S}'}_2^{(4)}$ as well as the following inequality:

$$x_3^{(3)} = x_3^{(4)} + x_4^{(4)}.$$

It is easy to see that $\mathcal{S}^{(4)}$ has a real solution and $0 < x_1 < \frac{1}{16}$ is derivable in this system. Similarly, $\mathcal{S}^{(5)}$ can be constructed from $\mathcal{S}^{(4)}$ and $\mathcal{S}_2^{(5)}$ (constructed as above) by some replacements and by adding the following equation:

$$x_4^{(4)} = x_4^{(5)} + x_5^{(5)}.$$

There is a pattern for the kind of relationships among $x_j^{(i)}$'s.

$$
\begin{tikzcd}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_1^{(6)} & x_2^{(6)} & x_3^{(6)} & x_4^{(6)} & x_5^{(6)} & x_6^{(6)} & \\
x_1^{(5)} \arrow[u] & x_2^{(5)} \arrow[u] & x_3^{(5)} \arrow[u] & x_4^{(5)} \arrow[u] & x_5^{(5)} \arrow[u] \arrow[ur] & & \\
x_1^{(4)} \arrow[u] & x_2^{(4)} \arrow[u] & x_3^{(4)} \arrow[u] & x_4^{(4)} \arrow[u] \arrow[ur] & & & \\
x_1^{(3)} \arrow[u] & x_2^{(3)} \arrow[u] & x_3^{(3)} \arrow[u] \arrow[ur] & & & & \\
x_1^{(2)} \arrow[u] & x_2^{(2)} \arrow[u] \arrow[ur] & & & & & \\
x_1^{(1)} \arrow[u] \arrow[ur] & & & & & &
\end{tikzcd}
$$

$\mathcal{T}_a$ denotes this tree. By repeating this inductive process, we will achieve a sequence of finite systems:

$$\mathcal{S}^{(1)} \subset \mathcal{S}^{(2)} \subset \mathcal{S}^{(3)} \subset \cdots \subset \mathcal{S}^{(n)} \subset \cdots$$

such that, for every n,

- $\mathcal{S}^{(n)}$ has a real solution;
- $0 < x_1 < \frac{1}{2^n}$ is derivable in this system.

$\mathcal{S}$ denotes $\bigcup_{n=1}^{\infty} \mathcal{S}^{(n)}$. It is easy to check that $\mathcal{S}$ satisfies all the properties that we need. So we have finished the proof of the theorem.

$\square$

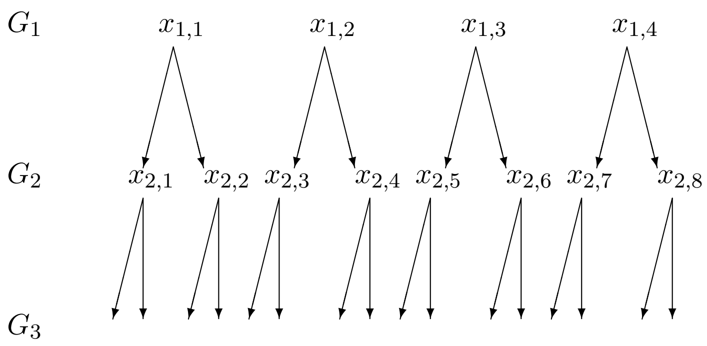
So our conjecture does not hold.

Corollary 3.3.4. *(Non-compactness) There is a linear probability tree system $\mathcal{S}$ that is finitely solvable but not solvable.*

Proof. Consider the system $\mathcal{S}$ that we construct in the above theorem. Recall that it is finitely solvable. However, it is not solvable, since $0 < x_1 < \frac{1}{2^n}$ for all n is derivable in this system.

$\square$

3.3.3. No Compactness or Strong Completeness for Probability Logic with Finite Indexes. In this section, we show that there is no hope of compactness and strong completeness for probability logics even with finite indices by embedding the system of linear inequalities in the last section into a certain structure of formulas. Enumerate all propositional letters: $p_0, p_1, p_2, \cdots$. Define $G_n := \{\bigwedge_{i=0}^{n} s_i p_i : s_i \text{ is } \neg \text{ or blank }\}$. G_n will be the n-th generation in our following tree. Elements of G_n are called *n-atoms*. An $n+1$-atom x is an *immediate descendant* of an n-atom x' if x' is a conjunct of x. An n-atom x is a *descendant* of a m-atom y if $n > m$ and y is a conjunct of x. Actually the second condition implies the first one. It is well-known that all these atoms form a binary tree, which is denoted by $\mathcal{T}^{(2)}$. We illustrate the tree as follows.



Note that each node in the n-th generation is an n-atom and the tree $\mathcal{T}^{(2)}$ consists of the four subtrees $\mathcal{T}_1^{(2)}, \mathcal{T}_2^{(2)}, \mathcal{T}_3^{(2)}$ and $\mathcal{T}_4^{(2)}$ with roots $x_{1,1}, x_{1,2}, x_{1,3}$ and $x_{1,4}$, respectively. Now we embed the nodes of the system of linear inequalities in the last section to the tree $\mathcal{T}^{(2)}$ in the following way:

(1) the tree $\mathcal{T}_a$ in the counterexample is faithfully embedded into the subtree $\mathcal{T}_4^{(2)}$;
(2) $x_1, x_2, \cdots, x_n, \cdots$ are mapped into the other three subtrees $\mathcal{T}_1^{(2)}, \mathcal{T}_2^{(2)}$ and $\mathcal{T}_3^{(2)}$ such that x_i's are in different branches of $\mathcal{T}^{(2)}$.

We still use x_i and x_i^j to denote their images, which are atoms and hence Boolean formulas of propositional letters. Next we translate the linear inequalities in the system $\mathcal{S}$ into a set of formulas. There are tree types of linear inequalities or equations in $\mathcal{S}$:

(1) $\frac{j}{2} < x_{i_1} + \cdots + x_{i_n} < \frac{j+1}{2}$ where $\frac{j}{2}, \frac{j+1}{2} \in \mathcal{Q}_2$;

(2) $x_{i_1} + \cdots + x_{i_n} \in \mathcal{Q}_2$;

(3) $x_{i_1} + \cdots + x_{i_n} = x_{j_1} + \cdots + x_{j_m}$

They are translated into formulas as follows:

(1) for the first type, it is translated to $((L_{\frac{j}{2}} \wedge \neg M_{\frac{j}{2}})(x_{i_1} \vee \cdots \vee x_{i_n})) \wedge ((M_{\frac{j+1}{2}} \wedge \neg L_{\frac{j+1}{2}})(x_{i_1} \vee \cdots \vee x_{i_n}))$

(2) for the second type, it is translated directly into $((L_{\frac{j}{2}} \wedge M_{\frac{j}{2}})(x_{i_1} \vee \cdots \vee x_{i_n}))$ for $\frac{j}{2} \in \mathcal{Q}_2$;

(3) for the third type, it only appears in the tree $\mathcal{T}_a$ and it is already incorporated in the relationships among atoms of $\mathcal{T}_4^{(2)}$. For example, if $x_1^{(1)} = x_1^{(2)} + x_2^{(2)}$, then, according to this mapping, $x_1^{(1)} \leftrightarrow x_1^{(2)} \vee x_2^{(2)}$ is a propositional tautology.

Let $\Gamma_{\mathcal{S}}$ denote this translated set of formulas from the system $\mathcal{S}$.

THEOREM 3.3.5. *(Non-Compactness) There is a set of formulas in the language of* Σ_+ *with finite indices that is finitely satisfiable but not satisfiable.*

PROOF. Consider the above $\Gamma_{\mathcal{S}}$. Since each finite subset of $\mathcal{S}$ has a solution, $\Gamma_{\mathcal{S}}$ is finitely satisfiable. According to this faithful translation, we can see immediately that

$$\{\neg M_0 x_1\} \cup \{M_{\frac{1}{2^n}} x_1 : n \geq 1\} \text{ is derivable from } \Gamma_{\mathcal{S}} \text{ in } \Sigma_+.$$

So $\Gamma_{\mathcal{S}}$ does not have a probability model.

□

COROLLARY 3.3.6. *Strong completeness does not hold for probability logics with finite indices. In other words, if* Γ *is a (possibly infinite) set of formulas in the language of* Σ_+ *that is consistent in* Σ_+*, it is not necessarily true that it has a probability model.*

PROOF. Since strong completeness = weak completeness + compactness, there is no hope of strong completeness for probability logics even with finite indices, either.

□

CHAPTER 4

Probability Logic for Harsanyi Type Spaces

In game theory and economics, there is a special kind of type spaces which are used to describe introspection of agents. In a *Harsanyi type space*, each agent is certain of his degree of belief at every state of this space. It is shown in Heifetz and Samet [18] that there is a universal Harsanyi type space within a general measure-theoretical context. In this section, we provide a deductive system which is complete with respect to the class of Harsanyi type spaces. Moreover, we underscore the relative simplicity of one-person belief system by showing that each atom in a finite language has only one maximal consistent extension if we keep the depth and width of the language the same. Viglizzo [37] and Viglizzo and Moss [29] study Harsanyi type spaces from the perspective of coalgebra.

4.1. Normal forms

Given any finite local language $\mathcal{L}(q,d,w)$, we define $F(q,n,w)(n \leq d)$, the set of *normal forms of depth* n, as follows:

(1) $n = 0$. $F(1,0,w)$ is the set of formulas $\pi_1 p_1 \wedge \pi_2 p_2 \wedge \cdots \wedge \pi_w p_w$ where each $\pi_i (1 \leq i \leq w)$ is either blank or $\neg$;

(2) $n > 0$. Then $F(q,n,w)$ is the set of formulas $\bigwedge\{p_i : 1 \leq i \leq w \text{ and } p_i \in \Gamma(q,n,w)\} \wedge \bigwedge\{\neg p_i : 1 \leq i \leq w \text{ and} \neg p_i \in \Gamma(q,n,w)\} \wedge \bigwedge\{L_r\phi, M_s\phi : L_r\phi, M_s\phi \in \Gamma(q,d,w)\} \wedge \bigwedge\{\neg L_r\phi, \neg M_s\phi : \neg L_r\phi, \neg M_s\phi \in \Gamma(q,d,w)\}$ for some atom $\Gamma(q,d,w) \in \Omega(q,d,w)$.

The first part $\bigwedge\{p_i : 1 \leq i \leq w \text{ and } p_i \in \Gamma(q,n,w)\} \wedge \bigwedge\{\neg p_i : 1 \leq i \leq w \text{ and} \neg p_i \in \Gamma(q,n,w)\}$ is called *the propositional part* of the normal form and $\bigwedge\{L_r\phi, M_s\phi : L_r\phi, M_s\phi \in \Gamma(q,d,w)\} \wedge \bigwedge\{\neg L_r\phi, \neg M_s\phi : \neg L_r\phi, \neg M_s\phi \in \Gamma(q,d,w)\}$ for some atom $\Gamma(q,d,w) \in \Omega$ (q,d,w) is called *the probability part* of the normal form.

LEMMA 4.1.1. *For any atom $\Gamma(q,d,w)$, the conjunction $\gamma(q,d,r)$ of all formulas in $\Gamma(q,d,w)$ is tautologically equivalent to a normal form in the language $\mathcal{L}(q,d,w)$. Moreover, any formula $\phi \in \Phi(q,d,w)$ is tautologically equivalent to a disjunction of normal forms in the language $\mathcal{L}(q,d,w)$.*

PROOF. The second part follows from the first part immediately. So we only show the first part. There are two types of *atomic formulas*:

(1) literals, i.e. propositional letters or their negations;
(2) formulas of the forms $L_r\phi, M_r\phi, \neg L_r\phi$ or $\neg M_r\phi$ (all of them are formulas in Φ (q,d,w)).

Any formula $\bigwedge\{\pi_i p_i : 1 \leq i \leq w\} \wedge \bigwedge\{\pi_{L_r\phi} L_r\phi : L_r\phi \in \Phi(q,d,w)\} \wedge \bigwedge\{\pi_{M_r\phi} M_r\phi : M_r\phi \in \Phi(q,d,w)\}$ where $\pi_i, \pi_{L_r\phi}, \pi_{M_r\phi}$ are either blank or $\neg$ is called *a statement.* Note that a statement is not necessarily consistent. By using propositional reasoning, we can show that the conjunction $\gamma(q,d,w)$ of formulas in the atom $\Gamma(q,d,w)$ is tautologically equivalent to a conjunction of statements: $s_1 \vee s_2 \vee \cdots \vee s_k$. Note that we regard each formula as an equivalence class where the equivalence relation is the tautological equivalence. Under this equivalence, $\gamma(q,d,w) \in \Gamma(q,d,w)$. One of the statements, say s_1, is in $\Gamma(q,d,w)$ because otherwise $\neg\gamma(q,d,w) \in \Gamma(q,d,w)$, which contradicts the fact that $\gamma(q,d,w) \in \Gamma(q,d,w)$. Moreover, $s_i \to \neg s_j$ is a tautology for any i and j such that $1 \leq i,j \leq k$ and $i \neq j$. Then only the state s_1 is in $\Gamma(q,d,w)$. Suppose that $s_i \in \Gamma(q,d,w)$ for some $i \neq 1$. This would imply that $\Gamma(q,d,w)$ is inconsistent because $s_i \to \neg s_1$ is a tautology and $s_1 \in \Gamma(q,d,w)$. Therefore, for $i \neq 1$, $\neg s_i \in \Gamma(q,d,w)$. So we have shown that there is one and only one statement that is in $\Gamma(q,d,w)$.

Since $\gamma(q,d,w)$ is tautologically equivalent to the disjunction $s_1 \vee s_2 \vee \cdots \vee s_k$, $s_1 \to \gamma(q,d,w)$ is a tautology. Moreover, $s_1 \in \Gamma(q,d,w)$. This implies that $\gamma(q,d,w) \to s_1$ is a tautology because $\gamma(q,d,w)$ is the conjunction of all the formulas in $\Gamma(q,d,w)$ and s_1 is one of them. So we have shown that $\gamma(q,d,w)$ is tautologically equivalent to a statement s_1, which is a normal form according to the definition.

□

DEFINITION 4.1.2. In the above lemma, we have shown that, for any atom $\Gamma(q,d,w)$, $\gamma(q,d,w)$ is tautologically equivalent to a statement and also to a normal form s_1. This normal form is called *the normal form* of $\Gamma(q,d,w)$ (or of $\gamma(q,d,w)$). For any formula ϕ, $\phi_{=0}$ denotes the propositional part of its normal form and $\phi_{>0}$ denotes its probability part.

4.2. Correspondence and Completeness

In this section we will show that the axiom system Σ_+ plus the following two axiom schemes:

- $(4_p) : L_r\phi \rightarrow L_1 L_r\phi$
- $(5_p) : \neg L_r\phi \rightarrow L_1\neg L_r\phi$

is sound and completeness with respect to the Harsanyi type spaces. Let Σ_H denote this system. Assume that $\langle \Omega, \mathcal{A}, T\rangle$ is a type space and $\mathcal{A}$ is generated by a countable subfield $\mathcal{A}_0$. $[T(w)]$ denotes $\{w' : T(w) = T(w')\}$. Indeed, for any w,

$$\begin{aligned}
[T(w)] &= \bigcap_{A\in\mathcal{A}_0} \{w' \in \Omega : T(w)(A) = T(w')(A)\} \\
&= \bigcap_{A\in\mathcal{A}_0} \bigcap_{r\in\mathcal{Q}\cap[0,1]} \{w' \in \Omega : T(w')(A) \geq r \leftrightarrow T(w)(A) \geq r\} \\
&= \bigcap_{A\in\mathcal{A}_0} \bigcap_{r\in\mathcal{Q}\cap[0,1], T(w)(A)\geq r} \{w' \in \Omega : T(w')(A) \geq r\}
\end{aligned}$$

and is hence measurable. Note that the last equality comes from the fact that $\mathcal{A}_0$ is a field. If each type is certain of its type, i.e.,

$$T(w)([T(w)]) = 1 \text{ for all } w \in S,$$

the type space is called *a Harsanyi type space*. First we show a correspondence result.

THEOREM 4.2.1. *(Correspondence) Let $S = \langle \Omega, \mathcal{A}, T\rangle$ be a type space and $\mathcal{A}_0$ be a countable subfield generating $\mathcal{A}$. Then both 4_p and 5_p are valid in S if and only if S is a Harsanyi type space.*

PROOF. This theorem follows directly from the proof that $[T(w)]$ is measurable. □

The main result of this section is the following theorem:

THEOREM 4.2.2. *Σ_H is sound and complete with respect to the class of Harsanyi type spaces.*

The soundness of the system is clear. We concentrate on the completeness. In [17], they gave a proof sketch of the completeness proof. But it seems that they missed many crucial steps that are needed for the completeness. Here we give a detailed proof of the completeness. Assume that ψ is consistent. We need to show that it is satisfiable in a Harsanyi type space. Just as in the proof of the completeness of Σ_+, we define a local langauge $\mathcal{L}[\psi]$. Let $\Phi(\psi)$ denote the finite set of formulas in this language. It gives rise to a set Ω of maximal consistent sets, whose elements are called *atoms*.

PROOF. Assume that ψ is consistent and is in $\Phi(q,d,w)$. Enumerate all the atoms in $\Omega(q,d,w)$:

$$\Gamma_1(q,d,w), \Gamma_2(q,d,w), \cdots, \Gamma_{N(q,d,w)}(q,d,w).$$

The conjunction of formulas in any atom $\Gamma_i(q,d,w) \in \Omega(q,d,w)$ is denoted as $\gamma_i(q,d,w)$. It has a normal form $\phi^i_{=0} \wedge \phi^i_{>0}$ where $\phi^i_{=0}$ is the propositional part and $\phi^i_{>0}$ is the probability part.

CLAIM 4.2.3. *Any maximal consistent extension $\Gamma_i(q,d+1,w) \in \Omega(q,d+1,w)$ of $\Gamma_i(q,d,w)$ contains the formula $L_1(\phi^i_{>0})$.*

For any maximal consistent extension $\Gamma_i(q,d+1,w) \in \Omega(q,d+1,w)$, $L_1(\phi^i_{>0}) \in \Gamma_i(q,d+1,w)$ since $\phi^i_{>0} \in \Gamma_i(q,d,w) \subseteq \Gamma_i(q,d+1,w)$.

We divide the set of atoms in $\Omega(q,d,w)$ into many groups according to their probability parts. Let $0 = i_0 < i_1 < \cdots < i_k < \cdots < i_N = N(q,d,w)$. Define:

$$G_k := \{\Gamma_j(q,d,w) \in \Omega(q,d,w) : i_{k-1} + 1 \leq j \leq i_k\} \text{ for each } 1 \leq k \leq N.$$

Without loss of generality, we assume that

- all the atoms in each group have the same probability parts;
- atoms in different groups have different probability parts.

Next we define probability measures at all atoms according to their representatives at each group. Given any $\Gamma_{i_k}(q,d,w)$, we know that it has a maximal consistent extension denoted by $\Gamma_{i_k}(q,d,w) \in \Omega(q,d+1,w)$. Moreover, there is a maximal consistent extension $\Gamma^{\infty}_{i_k}(q,d,w)$ in the language $\mathcal{L}$ that includes $\Gamma_{i_k}(q,d,w)$. Now, as usual, we define the probability according to $\Gamma^{\infty}_{i_k}(q,d,w)$. For any formula γ in the formal language $\mathcal{L}$, we have shown that $\alpha_{\gamma}^{\Gamma^{\infty}_{i_k}(q,d,w)} = \beta_{\gamma}^{\Gamma^{\infty}_{i_k}(q,d,w)}$. Recall that, for any formula $\phi \in \Phi(q,d,w)$, $[\phi] = \{\Delta(q,d,w) \in \Omega(q,d,w) : \phi \in \Delta(q,d,w)\}$. Define:

$$T(\Gamma_{i_k}(q,d,w))([\phi]) = \alpha_{\phi}^{\Gamma^{\infty}_{i_k}(q,d,w)}$$

We have shown that this indeed defines a probability measure at $\Gamma_{i_k}(q,d,w)$ satisfying the following properties: for any formula $\phi \in \Phi(q,d,w)$,

- if $\alpha_{\phi}^{\Gamma_{i_k}(q,d+1,w)} = \beta_{\phi}^{\Gamma_{i_k}(q,d+1,w)}$, then $T(\Gamma_{i_k}(q,d,w))([\phi]) = \alpha_{\phi}^{\Gamma_{i_k}(q,d+1,w)}$;
- if $\alpha_{\phi}^{\Gamma_{i_k}(q,d+1,w)} < \beta_{\phi}^{\Gamma_{i_k}(q,d+1,w)}$, $\alpha_{\phi}^{\Gamma_{i_k}(q,d+1,w)} < T(\Gamma_{i_k}(q,d,w))([\phi]) < \beta_{\phi}^{\Gamma_{i_k}(q,d+1,w)}$;

This is to say, $S(\Gamma_{i_k}(q,d,w)) := \langle \Omega(q,d,w), 2^{\Omega(q,d,w)}, T(\Gamma_{i_k}(q,d,w)) \rangle$ is a probability space.

Claim 4.2.4. $T(\Gamma_{i_k}(q,d,w))(G_k) = 1$

Observe that $G_k = [\phi^{i_k}_{>0}]$. Since $\Gamma^{\infty}_{i_k}(q,d,w)$ is a maximal consistent extension of $\Gamma_{i_k}(q,d,w)$, $L_1(\phi^{i_k}_{>0}) \in \Gamma^{\infty}_{i_k}(q,d,w)$. So $T(\Gamma_{i_k}(q,d,w))(G_k) = T(\Gamma_{i_k}(q,d,w))(\phi^{i_k}_{>0}) = \alpha_{\phi^{i_k}_{>0}}^{\Gamma^{\infty}_{i_k}(q,d,w)} = 1$.

Define probability measures at other atoms according to their representatives. For any j such that $i_{k_1} + 1 \leq j \leq i_k$, define:

$$T(\Gamma_j(q,d,w))([\phi]) = T(\Gamma_{i_k})([\phi]) \text{ for any formula } \phi \in \Phi(q,d,w).$$

Note that $T(\Gamma_j(q,d,w))([\phi^j_{>0}]) = 1$ because $[\phi^j_{>0}] = [\phi^{i_k}_{>0}]$. After defining T for all atoms, we take $M(q,d,w) := \langle \Omega(q,d,w), 2^{\Omega(q,d,w)}, T, \nu \rangle$ where $\nu(p) := \{\Delta(q,d,w) \in \Omega(q,d,w) : p \in \Delta(q,d,w)\}$. $[[\phi]]$ denote the set $\{\Delta(q,d,w) \in \Omega(q,d,w) : M(q,d,w), \Delta(q,d,w) \models \phi\}$.

CLAIM 4.2.5. *For any formula $\phi \in \Phi(q,d,w)$ and any atom $\Delta(q,d,w) \in \Omega(q,d,w)$, $M(q,d,w), \Delta(q,d,w) \models \phi$ iff $\phi \in \Delta(q,d,w)$. Equivalently, $[\phi] = [[\phi]]$.*

This is exactly the truth lemma. As usual, we prove by induction on the formula ϕ. It is easy to see that the claim holds for base case and Boolean ones. Here we only show the nontrivial case. Assume that $M(q,d,w), \Delta(q,d,w) \models L_r\phi$. It follows that $T(\Delta(q,d,w))([[\phi]]) \geq r$. By induction hypothesis, this is equivalent to saying that $T(\Delta(q,d,w))([\phi]) \geq r$. Suppose that $\Delta(q,d,w) \in G_k$. It follows that $T(\Gamma_{i_k}(q,d,w)) = T(\Delta(q,d,w))$. Hence $T(\Gamma_{i_k}(q,d,w))([\phi]) \geq r$. If $\alpha_\phi^{\Gamma_{i_k}(q,d+1,w)} = \beta_\phi^{\Gamma_{i_k}(q,d+1,w)}$, then $\alpha_\phi^{\Gamma_{i_k}(q,d+1,w)} = T(\Gamma_{i_k}(q,d,w))$ $([\phi]) \geq r$. This implies that $L_r\phi \in \Gamma_{i_k}(q,d+1,w)$ and hence $L_r\phi \in \Gamma_{i_k}(q,d,w)$ because $\Gamma_{i_k}(q,d,w) = \Gamma_{i_k}(q,d+1,w) \cap \Phi(q,d,w)$. Moreover, since $\Delta(q,d,w)$ and $\Gamma_{i_k}(q,d,w)$ have the same probability parts, $L_r\phi \in \Delta(q,d,w)$. If $\alpha_\phi^{\Gamma_{i_k}(q,d+1,w)} < \beta_\phi^{\Gamma_{i_k}(q,d+1,w)}$, then $r \leq \alpha_\phi^{\Gamma_{i_k}(q,d+1,w)}$ because r is a multiple of $1/q$. Similarly, this implies that $L_r\phi \in \Gamma_{i_k}(q,d+1,w)$ and hence $L_r\phi \in \Gamma_{i_k}(q,d,w)$ because $L_r\phi \in \Phi(q,d,w)$. Since $\Delta(q,d,w)$ and $\Gamma_{i_k}(q,d,w)$ are in the same group and hence have the same probability parts, $L_r\phi \in \Delta(q,d,w)$.

Now for the other direction. Assume that $L_r\phi \in \Delta(q,d,w)$. It follows that $L_r\phi \in \Gamma_{i_k}(q,d,w)$ and hence $L_r\phi \in \Gamma_{i_k}^\infty(q,d,w)$. This implies that $T(\Gamma_{i_k})([\phi]) = \alpha_\phi^{\Gamma_{i_k}^\infty(q,d,w)} \geq r$. By induction hypothesis, this is equivalent to saying $T(\Gamma_{i_k})([[\phi]]) \geq r$. That is to say, $M(q,d,w), \Delta(q,d,w) \models L_r\phi$. Now we finish the proof of the claim.

CLAIM 4.2.6. *$M(q,d,w)$ is a Harsanyi type space.*

Recall that, for any atom $\Delta(q,d,w)$, $[T(\Delta(q,d,w))]$ denotes the set $\{\Delta'(q,d,w) \in \Omega(q,d,w) : T(\Delta'(q,d,w)) = T(\Delta(q,d,w))\}$. We need to show that $T(\Delta(q,d,w))([T(\Delta(q,d,w))]) = 1$ for all atom $\Delta(q,d,w)$. Assume that $\Delta(q,d,w) \in G_k$. In the above definition of T, we make the probability measures at atoms in each group the same and probability measures at atoms from different groups are different. This implies that $[T(\Delta(q,d,w))] = G_k$. We already know that $G_k = [\phi_{>0}^{i_K}]$. According to Claim 4.2.4, $T(\Gamma_{i_k})([\phi_{>0}^{i_K}]) = T(\Gamma_{i_k})(G_k) = 1$. Since $T(\Delta(q,d,w)) = T(\Gamma_{i_k}(q,d,w))$, $T(\Delta(q,d,w))([T(\Delta(q,d,w))]) = T(\Delta(q,d,w))(G_k) = T(\Gamma_{i_k}(q,d,w))(G_k) = 1$. So we have shown that, for any atom $\Delta(q,d,w) \in \Omega(q,d,w)$,

$T(\Delta(q,d,w))([T(\Delta(q,d,w))]) = 1$. That is to say, $M(q,d,w)$ is a Harsanyi type space.

Since ψ is consistent, it is contained in an atom $\Gamma_i(q,d,w)$ for some i such that $1 \leq i \leq N(q,d,w)$. By the truth lemma, we know that $M(q,d,w), \Gamma_i(q,d,w) \models \psi$. In other words, ψ is satisfiable in a Harsanyi type space. So we finish the proof of the completeness.

□

THEOREM 4.2.7. *It is decidable whether ϕ is a theorem of Σ_H.*

PROOF. The proof here is similar to the decidability proof of Σ_+. We only need to add the following two axioms to Σ_I:

$$L_r\phi \to L_1 L_r\phi \text{ and } \neg L_r\phi \to L_1\neg L_r\phi \text{ where } r \text{ is a parameter.}$$

Through this system with index parameters, we can show that the set of the theorems of Σ_H is recursively enumerable. Moreover, by the above completeness theorem, we know that Σ_H has finite model property. These imply that it is decidable whether ϕ is a theorem of Σ_H.

□

4.3. Basic Properties of Σ_H

LEMMA 4.3.1. *The proposition consists of four parts:*

(1) $\vdash_{\Sigma_H} L_s\psi \wedge L_r\phi \to L_r(\phi \wedge L_s\psi)$
(2) $\vdash_{\Sigma_H} L_s\psi \wedge \neg L_r\phi \to \neg L_r(\phi \wedge L_s\psi)$
(3) $\vdash_{\Sigma_H} \neg L_s\psi \wedge L_r\phi \to L_r(\phi \wedge \neg L_s\psi)$
(4) $\vdash_{\Sigma_H} \neg L_s\psi \wedge \neg L_r\phi \to \neg L_r(\phi \wedge \neg L_s\psi)$

PROOF. Here we use the rule (B). First note that $\vdash_{\Sigma_+} (\phi,\psi) \leftrightarrow (\phi \vee \psi, \phi \wedge \psi)$. Reason inside Σ_H:

$$\begin{aligned} L_r\phi \wedge L_s\psi &\rightarrow L_r\phi \wedge L_1(L_s\psi) \wedge M_1(\phi \vee L_s\psi) \\ &\rightarrow L_r(\phi \wedge L_s\psi) \end{aligned}$$

In both (1) and (2), we apply the rule (B). Dually, we have:

$$\begin{aligned} \neg L_r\phi \wedge L_s\psi &\rightarrow \neg L_r\phi \wedge M_1(L_s\psi) \wedge L_1(\phi \vee L_s\psi) \\ &\rightarrow \neg L_r(\phi \wedge L_s\psi) \end{aligned}$$

In (3), we use implicitly the following argument:

$$\begin{aligned} L_s\psi &\rightarrow L_1 L_s\psi \\ &\rightarrow L_1(\phi \vee L_s\psi) \end{aligned}$$

The proofs of 3 and 4. are similar.

$\square$

THEOREM 4.3.2. *The proposition consists of two parts:*

(1) $\vdash_{\Sigma_H} (\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j) \wedge L_r\phi \rightarrow L_r(\phi \wedge (\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j))$

(2) $\vdash_{\Sigma_H} (\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j) \wedge \neg L_r\phi \rightarrow \neg L_r(\phi \wedge (\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j))$

PROOF. This proposition follows directly from the above lemma by the following fact:

$$\vdash_{\Sigma_H} (\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j) \rightarrow L_1(\bigwedge_{i=1}^m L_{s_i}\psi_i \wedge \bigwedge_{j=1}^n \neg L_{t_j}\psi'_j)$$

$\square$

LEMMA 4.3.3. $\vdash_{\Sigma_H} (L_r\phi \vee L_s\psi) \rightarrow L_1(L_r\phi \vee L_s\psi)$

PROOF. Reason inside Σ_H:

$$\begin{array}{rcl} L_r\phi & \to & L_1 L_r\phi \\ & \to & L_1(L_r\phi \vee L_s\psi) \\ L_s\psi & \to & L_1(L_r\phi \vee L_s\psi) \\ L_r\phi \vee L_s\psi & \to & L_1(L_r\phi \vee L_s\psi) \end{array}$$

□

LEMMA 4.3.4. $\vdash_{\Sigma_H} (L_r\phi \wedge L_s\psi) \to L_1(L_r\phi \wedge L_s\psi)$, *and* $\vdash_{\Sigma_H} (L_r\phi \wedge \neg L_s\psi) \to L_1(L_r\phi \wedge \neg L_s\psi)$

PROOF. The proof of the second part is similar to that of the first one. We only show the first part. Reason inside Σ_H:

$$\begin{array}{rcl} L_r\phi \wedge L_s\psi & \to & L_1 L_r\phi \wedge L_1 L_s\psi \\ & \to & L_1 L_r\phi \wedge L_1 L_s\psi \wedge M_1(L_r\phi \vee L_s\psi) \\ & \to & L_1(L_r\phi \wedge L_s\psi) \end{array}$$

□

THEOREM 4.3.5. *If ϕ is a Boolean combination of formulas of the form $L_r\psi$, then $\phi \to L_1\phi$ is a theorem in Σ_H.*

PROOF. Assume that ϕ is a Boolean combination of formulas of the form $L_r\gamma$ and its disjunctive normal form is $\bigvee_{i=1}^{I} \bigwedge_{j=1}^{k_i} \pi_j^i L_{r_j^i} \phi_j^i$ where π_j^i is either blank or $\neg$. Reason inside Σ_H:

$$
\begin{aligned}
\bigvee_{i=1}^{I}\bigwedge_{j=1}^{k_i}\pi_j^i L_{r_j^i}\phi_j^i &\rightarrow \bigvee_{i=1}^{I} L_1(\bigwedge_{j=1}^{k_i}\pi_j^i L_{r_j^i}\phi_j^i) \\
&\rightarrow L_1(\bigvee_{i=1}^{I}\bigwedge_{j=1}^{k_i}\pi_j^i L_{r_j^i}\phi_j^i)
\end{aligned}
$$

□

Fix the index vector (q, d, w) of the local language. Define

- $\Gamma_{=0} := \{\gamma_{=0} : \gamma_{=0}$ is the propositional part of some atom $\Gamma(q, d, w)\}$;
- $\Gamma_{>0} := \{\gamma_{>0} : \gamma_{>0}$ is the probability part of some atom $\Gamma(q, d, w)\}$;
- $\Phi_{=0} := \{\Phi_{=0} : \phi_{=0}$ is the propositional part of some formula $\phi \in \Phi(q, d, w)\}$;
- $\Phi_{>0} := \{\phi_{>0} : \phi_{>0}$ is the probability part of some formula $\phi \in \Phi(q, d, w)\}$.

Observe that, for any $\phi_{>0} \in \Phi_{>0}$, $\vdash_{\Sigma_H} \phi_{>0} \rightarrow L_1\phi_{>0}$.

LEMMA 4.3.6. *For any $\gamma_{=0} \in \Gamma_{=0}$ and $\gamma_{>0} \in \Gamma_{>0}$, if $\gamma_{=0} \wedge \gamma_{>0}$ is inconsistent, then $\gamma_{>0} \rightarrow M_0\gamma_{=0}$ is provable in Σ_H.*

PROOF. Assume that $\gamma_{=0} \wedge \gamma_{>0}$ is inconsistent. Reason inside Σ_H:

$$
\begin{aligned}
\gamma_{>0} &\rightarrow \neg\gamma_{=0} \\
L_1\gamma_{>0} &\rightarrow L_1\neg\gamma_{=0} \\
\gamma_{>0} &\rightarrow L_1\gamma_{>0} \\
\gamma_{>0} &\rightarrow L_1\neg\gamma_{=0} \\
\gamma_{>0} &\rightarrow M_0\gamma_{=0}
\end{aligned}
$$

□

LEMMA 4.3.7. *Let Γ_1, Γ_2 and Γ_3 be three atoms in $\Omega(q, d, w)$ and γ_1, γ_2 and γ_3 be the conjunctions of formulas in these three atoms, respectively. Their normal forms are $\phi^1_{=0} \wedge \phi^1_{>0}$, $\phi^2_{=0} \wedge \phi^2_{>0}$ and $\phi^3_{=0} \wedge \phi^3_{>0}$. Then the following three propositions hold:*

(1) *if both* $\phi^2_{>0}$ *and* $\phi^3_{>0}$ *are different from* $\phi^1_{>0}$*, then* $\gamma_1 \to M_0(\gamma_2 \vee \gamma_3)$ *is provable in* Σ_H*;*

(2) *if only one of them, say,* $\phi^2_{>0}$*, is different from* $\phi^1_{>0}$*, then*

   (a) $\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3)$ *is provable in* Σ_H *whenever* $\alpha^{\Gamma_1}_{\phi^3_{=0}} = \beta^{\Gamma_1}_{\phi^3_{=0}}$*;*

   (b) $\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge \neg M_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge \neg L_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2 \vee \gamma_3)$ *is provable in* Σ_H *whenever* $\alpha^{\Gamma_1}_{\phi^3_{=0}} < \beta^{\Gamma_1}_{\phi^3_{=0}}$*;*

(3) *if none of these two is different from* $\phi^1_{>0}$*, then*

   (a) $\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3)$ *is a theorem of* Σ_H *whenever* $\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}} = \beta^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}$*;*

   (b) $\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge \neg M_{\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3) \wedge \neg L_{\beta^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}}(\gamma_2 \vee \gamma_3)$ *is a theorem of* Σ_H *whenever* $\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}} < \beta^{\Gamma_1}_{\phi^2_{=0} \vee \phi^3_{=0}}$*;*

PROOF. Assume that both $\phi^2_{>0}$ and $\phi^3_{>0}$ are different from $\phi^1_{>0}$. Reason inside Σ_H:

$$\begin{aligned}
\gamma_1 &\to \neg\phi^2_{>0} \\
&\to L_1(\neg\phi^2_{>0}) \\
&\to M_0(\phi^2_{>0}) \\
&\to M_0(\phi^3_{>0}) \\
&\to M_0(\phi^2_{>0} \vee \phi^3_{>0}) \\
&\to M_0(\gamma_2 \vee \gamma_3)
\end{aligned}$$

Next we assume that only one of them, say, $\phi^2_{>0}$, is different from $\phi^1_{>0}$. From above, we know that $\gamma_1 \to M_0\gamma_2$ is provable in Σ_H.

$$
\begin{aligned}
\gamma_1 &\rightarrow \phi^3_{>0}\\
&\rightarrow L_1(\phi^3_{>0})\\
&\rightarrow L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\phi^3_{=0}\wedge\phi^3_{>0})\\
&\rightarrow L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_3)\\
&\rightarrow L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\\
&\rightarrow M_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_3)\\
&\rightarrow M_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\\
&\rightarrow L_{\alpha^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\wedge M_{\beta^{\Gamma_1}_{\phi^3_{=0}}}(\gamma_2\vee\gamma_3)
\end{aligned}
$$

Assume that none of these two is different from $\phi^1_{>0}$. Then $\phi^1_{>0}=\phi^2_{>0}=\phi^3_{>0}$. Reason inside Σ_H:

$$
\begin{aligned}
\gamma_1 &\rightarrow \phi^2_{>0}\\
&\rightarrow L_1(\phi^2_{>0})\\
&\rightarrow L_1(\phi^2_{>0})\\
&\rightarrow L_1(\phi^3_{>0})\\
&\rightarrow L_{\alpha_{\phi^2_{=0}\vee\phi^3_{=0}}}(\phi^2_{>0}\wedge(\phi^2_{=0}\vee\phi^3_{=0}))\\
&\rightarrow L_{\alpha_{\phi^2_{=0}\vee\phi^3_{=0}}}((\phi^2_{>0}\wedge\phi^2_{=0})\vee(\phi^3_{>0}\wedge\phi^3_{=0}))\\
&\rightarrow L_{\alpha_{\phi^2_{=0}\vee\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\\
&\rightarrow M_{\beta_{\phi^2_{=0}\vee\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\\
&\rightarrow L_{\alpha_{\phi^2_{=0}\vee\phi^3_{=0}}}(\gamma_2\vee\gamma_3)\wedge M_{\beta_{\phi^2_{=0}\vee\phi^3_{=0}}}(\gamma_2\vee\gamma_3)
\end{aligned}
$$

$\square$

THEOREM 4.3.8. *Probabilities of formulas in the extensions of higher depth are* uniquely *determined by those in their restrictions of lower depth. Let* $\Gamma(q,d+1,w)$ *be a maximal* *consistent extension of* $\Gamma(q,d,w) \in \Omega(q,d,w)$ *by increasing its depth by 1 and the normal form of* $\Gamma(q,d,w)$ *is* $\phi_{=0} \wedge \phi_{>0}$. *Define* $\Gamma(q,1,d) := \Gamma(q,d,w) \cap \Phi(q,1,d)$. *Assume that* ϕ *is a formula in* $\Phi(q,d,w)$ *and it is tautologically equivalent to the disjunction of the following normal forms:*

$$\phi^1_{=0} \wedge \phi^1_{>0}, \phi^2_{=0} \wedge \phi^2_{>0}, \cdots, \phi^n_{=0} \wedge \phi^n_{>0}.$$

In addition, we assume that the first $m(\leq n)$ *probability parts are the same as* $\phi_{>0}$ *and other probability parts are different, i.e.*

$$\phi^i_{>0} = \phi^m_{>0}(1 \leq i \leq m) \text{ and } \phi^j_{>0} \neq \phi^m_{>0}(m+1 \leq j \leq n)$$

Then $\alpha_\phi^{\Gamma(q,d+1,w)} = \alpha^{\Gamma(q,1,w)}_{\bigvee_{i=1}^m \phi^i_{=0}}$ *and* $\beta_\phi^{\Gamma(q,d+1,w)} = \beta^{\Gamma(q,1,w)}_{\bigvee_{i=1}^m \phi^i_{=0}}$. *That is to say,* $\Gamma(q,d,w)$ *has one and only one maximal consistent extension in* $\Omega(q,d+1,w)$, *which is* $\Gamma(q,d+1,w)$.

THEOREM 4.3.9. *Assume that* $\Gamma_1(q,d,w)$ *and* $\Gamma_2(q,d,w)(d \geq 1)$ *are two atoms in* $\Omega(q,d,w)$, *and* $\Gamma_1(q,d+k,w)$ *and* $\Gamma_2(q,d+k,w)$ *are maximal consistent extensions in* $\Omega(q,d+k,w)$ *of* $\Gamma_1(q,d,w)$ *and* $\Gamma_2(q,d,w)$, *respectively. If the normal forms of* $\Gamma_1(q,d,w)$ *and* $\Gamma_2(q,d,w)$ *have the same probability parts, then the normal forms of* $\Gamma_1(q,d+k,w)$ *and* $\Gamma_2(q,d+k,w)$ *also have the same probability parts.*

COROLLARY 4.3.10. *Any atom* $\Gamma(q,d,w) \in \Omega(q,d,w)(d \geq 1)$ *has one and only one maximal consistent extension in* $\Omega(q,d+k,w)$ *for any* $k \geq 0$. *More precisely, it is the probability part of the normal form of* $\Gamma(q,d,w)$ *that uniquely determines the probability part(s) of its maximal consistent extension(s) in* $\Omega(q,d+k,w)$.

4.4. A Conservation Result

THEOREM 4.4.1. *For any formula* ψ *of depth* ≤ 1, *if* $\vdash_{\Sigma_H} \psi$, *then* $\vdash_{\Sigma_+} \psi$.

PROOF. It suffices to show that, for any formula of depth 1, if it is consistent in Σ_+, then so is it in Σ_H. Assume that ψ is a Σ_+-consistent formula of depth 1. Then it is contained in a maximal Σ_+-consistent set $\Gamma_0(q,1,w) \in \Omega(q,1,w)$ of formulas for some q and w where $\Omega(q,1,w)$ is the set of all maximal Σ_+-consistent set of formulas in $\Phi(q,1,w)$. Recall that $\Omega(1,0,w)$ denotes the set of all formulas of depth 0 with propositional letters

$p_1, \cdots, p_w$ and $\Gamma_0(q,0,w)$ denotes the set $\Gamma_0(q,1,w) \cap \Phi(q,0,w)$.i.e. the set of formulas of depth 0 in $\Gamma_0(q,1,w)$.

Now we define a Harsanyi type space on $\Omega(q,0,w)$. Consider $\Gamma_0(q,0,w)$, which is an element of $\Omega(q,0,w)$. We have shown that there is a probability measure $T(\Gamma(q,0,w))$ at $\Gamma(q,0,w)$ such that, for any formula $\phi \in \Phi(q,0,w)$,

(1) if $\alpha_\phi^{\Gamma_0(q,1,w)} = \beta_\phi^{\Gamma_0(q,1,w)}$, then $T(\Gamma_0(q,0,w))([\phi]) = \alpha_\phi^{\Gamma_0(q,1,w)}$;
(2) if $\alpha_\phi^{\Gamma_0(q,1,w)} < \beta_\phi^{\Gamma_0(q,1,w)}$, then $\alpha_\phi^{\Gamma_0(q,1,w)} < T(\Gamma_0(q,0,w))([\phi]) < \beta_\phi^{\Gamma_0(q,1,w)}$;

For other atoms $\Delta(q,0,w) \in \Omega(q,0,w)$, define $T(\Delta(q,0,w)) = T(\Gamma_0(q,0,w))$ and further the canonical model $M(q,0,w) := \langle \Omega(q,0,w), 2^{\Omega(q,0,w)}, T, \nu \rangle$ where $\nu(p) = \{\Delta(q,0,w) \in \Omega(q,0,w) : p \in \Delta(q,0,w)\}$. $M(q,0,w)$ is a Harsanyi type space. Indeed, for any atom $\Delta(q,0,w)$, $[T(\Delta(q,0,w))] = \Omega(q,0,w)$ and hence $T(\Delta(q,0,w))([T(\Delta(q,0,w))]) = 1$.

Claim 4.4.2. *$M(q,0,w), \Gamma_0(q,0,w) \models \psi$.*

It suffices to show that, for any formula $\phi \in \Phi(q,1,d)$, $M(q,0,d), \Gamma_0(q,0,d) \models \phi$ iff $\phi \in \Gamma_0(q,1,d)$. It is easy to check that this is true for the base case and the Boolean cases. Now we show the nontrivial case. Assume that $M(q,0,d), \Gamma_0(q,0,d) \models L_r\phi'$. $T(\Gamma_0(q,0,d))([[\phi']]) \geq r$. Note that ϕ' is a formula of depth 0. Obviously, $[[\phi']] = [\phi']$. So $T(\Gamma_0(q,0,d))([\phi']) \geq r$. If $\alpha_{\phi'}^{\Gamma_0(q,1,w)} = \beta_{\phi'}^{\Gamma_0(q,1,w)}$, then $\alpha_{\phi'}^{\Gamma_0(q,1,w)} = T(\Gamma_0(q,0,d))([\phi']) \geq r$. This implies that $L_r\phi' \in \Gamma_0(q,1,d)$. If $\alpha_{\phi'}^{\Gamma_0(q,1,w)} < \beta_{\phi'}^{\Gamma_0(q,1,w)}$, then $\alpha_{\phi'}^{\Gamma_0(q,1,w)} < T(\Gamma_0(q,0,d))$ $([\phi']) < \beta_{\phi'}^{\Gamma_0(q,1,w)}$. This also implies that $L_r\phi' \in \Gamma_0(q,1,d)$. For the other direction, assume that $L_r\phi' \in \Gamma_0(q,1,w)$. It follows that $r \leq \alpha_{\phi'}^{\Gamma_0(q,1,d)}$. Moreover, $T(\Gamma_0(q,0,d))([\phi']) \geq \alpha_{\phi'}^{\Gamma_0(q,1,d)}$. So $T(\Gamma_0(q,0,d))([\phi']) \geq r$ and hence by induction hypothesis $T(\Gamma_0(q,0,d))([[\phi']])$ $\geq r$.

We have shown that ψ is satisfiable in a Harsanyi type space. According to the above completeness theorem, ψ is Σ_H-consistent.

$\square$

THEOREM 4.4.3. *The probability part and the propositional part of any normal forms in* Σ_H *are independent of each other in the following sense: for any normal form* $\gamma(q,d,r)$*(*$\gamma_{=0}$ *and* $\gamma_{>0}$ *are its propositional part and probability part, respectively) in* Σ_H *and for any* Σ_H*-consistent formula* ϕ *of depth 0,* $\phi \wedge \gamma_{>0}$ *is consistent.*

PROOF. Fix a consistent formula ϕ of depth 0 and a normal form $\gamma(q,d,w)$ that is tautologically equivalent to the conjunction of all the formulas in some atom $\Gamma(q,d,w) \in \Omega(q,d,w)$. Let $\gamma_{=0}$ and $\gamma_{>0}$ be its propositional part and probability part, respectively. Since ϕ is consistent, it is a conjunct of the propositional part $\phi_{=0}$ of some normal form. It suffices to show that $\phi_{=0} \wedge \gamma_{>0}$ is Σ_H-consistent. $\Gamma(q,1,d)$ denotes $\Gamma(q,d,w) \cap \Phi(q,1,w)$. Suppose that $\gamma_{=0}(q,1,w)$ and $\gamma_{>0}(q,1,w)$ are the propositional part and the probability part of the normal form of $\Gamma(q,1,d)$. From one previous independence result, we know that $\phi_{=0} \wedge \gamma_{>0}(q,1,w)$ is Σ_+-consistent. According to the above conservation result, $\phi_{=0} \wedge \gamma_{>0}(q,1,w)$ is Σ_H-consistent. Moreover, it is a normal form of some atom $\Gamma'(q,1,w)$. But $\Gamma'(q,1,w)$ has a unique maximal consistent extension $\Gamma'(q,d,w)$ in $\Omega(q,d,w)$. Since the normal forms of $\Gamma(q,1,w)$ and $\Gamma'(q,1,w)$ have the same probability parts, the normal forms of $\Gamma(q,d,w)$ and $\Gamma'(q,d,w)$ have the same probability parts, too. That is to say, the probability part of the normal form of $\Gamma'(q,d,w)$ is also $\gamma_{>0}$. In addition, its propositional part is $\phi_{=0}$. So, $\phi_{=0} \wedge \gamma_{>0}$ is Σ_H-consistent.

□

The following figure illustrates the maximal consistent extensions in Σ_H. The first step maximal consistent extensions from $\top$ in Σ_H is the same as that in Σ_+ because the set of maximal Σ_H-consistent sets of formulas of depth ≤ 1 is the same as that of maximal Σ_+-consistent sets of formulas of depth ≤ 1. But after that, any atom has only one maximal consistent extension, which is illustrated in the figure by demonstrating that each node from the second step has only one descendant.

$$
\begin{tikzcd}[column sep=small]
 & & \top \arrow[dll] \arrow[dl] \arrow[dr] \arrow[drr] & & \\
\Gamma_1(q,1,w) \arrow[d] & \Gamma_2(q,1,w) \arrow[d] & \cdots & \Gamma_{N-1}(q,1,w) \arrow[d] & \Gamma_N(q,1,w) \arrow[d] \\
\Gamma_1(q,2,w) & \Gamma_2(q,2,w) & & \Gamma_{N-1}(q,2,w) & \Gamma_N(q,2,w) \\
\vdots \arrow[d] & \vdots \arrow[d] & \cdots & \vdots \arrow[d] & \vdots \arrow[d] \\
\Gamma_1(q,d,w) & \Gamma_2(q,d,w) & \cdots & \Gamma_{N-1}(q,d,w) & \Gamma_N(q,d,w)
\end{tikzcd}
$$

CHAPTER 5

Adding Knowledge to Belief

Aumann's knowledge-belief semantic systems are conservative extensions of Harsanyi type spaces by including knowledge operators in the languages. They have played an important role in game theory with incomplete information [2]. In this section, we add knowledge operators to our probability logic to get a complete axiomatization with respect to Aumann's semantics for knowledge and belief. Just as we did for the probability logic of Harsanyi type spaces, we show that, in our knowledge-belief deductive system, each atom in a finite language has only one maximal consistent extension if we keep the accuracy and the width of the language the same. This result is an indication of the relative simplicity of the one-person interactive epistemology. Moreover, despite the infinitary flavor of common knowledge, it can also be finitely axiomatized [15]. The forthcoming handbook chapter [5] offers a panorama of epistemic logic.

5.1. Interactive Epistemology: Knowledge

The syntax of our logic for knowledge is similar to that of modal logic. We start with a fixed infinite set $P := \{p_1, p_2, \cdots\}$ of propositional letters. We also use $p, q, \cdots$ to denote propositional letters. The set of formulas Φ is built from propositional letters as usual by connectives $\neg, \wedge$ and a modality operator K, which is the initial letter of the word knowledge. In other words, a formula ϕ is formed by the following syntax:

$$\phi := p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid K\phi.$$

A *knowledge frame* F is a tuple $\langle \Omega, R \rangle$ where Ω is a Kripke frame and R is an equivalence relation. A *knowledge model* M on the frame F is a tuple $\langle \Omega, R, v \rangle$ where $\langle \Omega, R, v \rangle$ is a knowledge frame and v is a valuation such that $v(p) \in 2^{\Omega}$ for any propositional letter p. The forcing relation $\models$ between states and formulas are defined inductively as follows:

- $M, w \models p$ if $w \in v(p)$ for propositional letters p;
- $M, w \models \phi_1 \wedge \phi_2$ if $M, w \models \phi_1$ and $M, w \models \phi_2$;
- $M, w \models \neg\phi$ if $M, w \not\models \phi$;
- $M, w \models K\phi$ if, for any w' such that wRw', $M, w' \models \phi$.

ϕ is *valid in the knowledge model* M if $M, w \models \phi$ for all states $w \in M$. ϕ is *valid in a knowledge frame* F if it is valid in all models on the frame. ϕ is *valid in a class of knowledge frames* if it is valid in all the frames in the class. It is well known that the following set of axiom schemata and inference rules provides a sound and complete axiomatization for the logic of knowledge with respect to the above class of knowledge frames:

- All instances of propositional tautologies;
- **K** : $K(\phi \to \psi) \to (K\phi \to K\psi)$;
- **T** : $K\phi \to \phi$;
- **4** : $K\phi \to KK\phi$;
- **5** : $\neg K\phi \to K\neg K\phi$;
- Modus Ponens: From ϕ and $\phi \to \psi$ infer ψ;
- Generalization: From ϕ infer $K\phi$.

The names **K**, **T**, **4**, **5** are standard in modal logic. **T** says that if the agent knows something, it should be true. **4** and **5** says that the agent knows that he knows or not. For completeness, we will present the proof of the soundness and completeness of the above system with respect to the class of knowledge frames.

THEOREM 5.1.1. *ϕ is valid in the class of knowledge frames if and only if* $\vdash_{S5} \phi$[1]

PROOF. It is easy to show the soundness. Here we only prove the completeness. Assume that ϕ is consistent. Now we need to show that ϕ is satisfiable in a knowledge model. Let $Sub(\phi)$ be the smallest set of formulas that includes all subformulas of ϕ and is closed under simple negation, i.e. $\neg\psi \in Sub(\phi)$ iff $\psi \in Sub(\Phi)$. Define Ω as the set of all maximal consistent subsets s of $Sub(\phi)$. Elements of Ω are called *atoms*. And they are also states of

[1] This is actually weak completeness. From this system, strong completeness also holds.

the following canonical model. Let k be the set function from Ω to $2^{Sub(\phi)}$: for each atom s, $k(s)$ is the subset of s of formulas of the form $K\psi$ or $\neg K\psi$. Define sRt if $k(s) = k(t)$. It is easy to check that R is indeed an equivalence relation on Ω. $v(p) = \{s \in \Omega : p \in s\}$. Let $M := \langle \Omega, R, v \rangle$. It remains to show the following truth lemma:

$$\text{(Truth Lemma) For any } \psi \in Sub(\phi),\ M, s \models \psi \text{ iff } \psi \in s.$$

We prove the truth lemma on the structural complexity of ψ. The only case that we need to take seriously is $\psi := K\psi'$. Assume that $M, s \models K\psi'$. Now we show that $K\psi' \in s$. Suppose that $K\psi' \notin s$. It follows that $\neg K\psi' \in s$. Therefore, $k(s) \cup \{\neg\psi'\}$ is consistent because, otherwise, $\vdash_{S5} k(s) \to \psi'$ and $\vdash_{S5} k(s) \to K(\psi')$ and hence $K\psi' \in s$, which contradicts the assumption that $K\psi' \notin s$. So, $K\psi' \in s$. Next we show the other direction. Assume that $K\psi' \in s$. We need to show that $M, s \models K\psi'$. There is an atom s' such that sRs', i.e. $k(s) = k(s')$. Then $K\psi' \in s'$. Since $K\psi' \to \psi'$ is an axiom of $S5$, $\psi' \in s'$. By induction hypothesis, we know that $M, s' \models \psi'$. So $M, s \models K\psi'$. We have finished the proof of the truth lemma and hence the theorem.

□

5.2. Aumann's Knowledge-belief System

Now we combine the syntaxes of $S5$ and Σ_+ to get a sound and complete axiomatization of the logic of knowledge and probability with respect to the class of knowledge-probability frames. Formally, a formula ϕ is formed by the following syntax:

$$\phi := p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid K\phi \mid L_r\phi$$

where r is a rational between 0 and 1. Σ_{S5} denote the system $\Sigma_+ + S5$. Since there is a well-known one-to-one correspondence between equivalence relations and partitions, we will use them alternatively. Let R be an equivalence relation on Ω and Π be the associated partition of Ω. A *choice function* c on the partition Π (or on R) is a function from Ω to Ω such that

$$\text{for any states } w_1 \in \Pi(w_2),\ c(w_1) = c(w_2) \in \Pi(w_2).$$

Definition 5.2.1. A *knowledge-probability* frame F is a tuple $\langle \Omega, R, c, \mathcal{A}, T \rangle$ where

- $\langle \Omega, R \rangle$ is a knowledge frame;
- $\langle \Omega, \mathcal{A}, T \rangle$ is a probability frame;
- the choice function c is $\mathcal{A}$-measurable.

A *knowledge-probability model* M is a tuple $\langle \Omega, R, \mathcal{A}, T, v \rangle$ where $\langle \Omega, R, \mathcal{A}, T \rangle$ is a knowledge-probability frame and v is a valuation such that $v(p) \in \mathcal{A}$ for all propositional letters p.

The following is the main theorem in this section. A variant of the same result can be found in [14].

THEOREM 5.2.2. *A formula ϕ is valid in the class of knowledge-probability frame iff $\vdash_{\Sigma_{S5}} \phi$.*

PROOF. (Sketch) The soundness is easy to check. Here we only show the completeness. Assume that ϕ is consistent. Here we extend the definition of $dp(\phi)$ to include the following one more clause:

- $dp(K\psi) = dp(\psi) + 1$.

Similarly define the local language $\mathcal{L}[\phi]$ to be the same as in [17] except that we count the depth in K. Note that formulas here are actually equivalence classes under the tautological equivalence. Such a defined language $\mathcal{L}[\phi]$ gives rise to a set of maximal consistent subsets and it is denoted by Ω, which will be the carrier set of our following canonical model. Elements in Ω are also called atoms. Using propositional reasoning, we can show that the conjunction of formulas in each atom Γ is equivalent to a conjunctive formula $\phi_0 \wedge \phi_K \wedge \phi_P$ where ϕ_0 is a conjunction of propositional letters, ϕ_K a conjunction of formulas of the forms $K\psi$ or $\neg K\psi$ and ϕ_P a conjunction of formulas of the forms $L_r\psi$ or $\neg L_r\psi$. This modularity in atoms reflects the modularity of our axiomatization of the knowledge-probability system. Define the knowledge accessibility relation R as usual: $\Gamma_1 R \Gamma_2$ if $k(\Gamma_1) = k(\Gamma_2)$. In order to define the type function T, we constantly increase just the accuracy of the local langauge $\mathcal{L}[\phi]$. When the accuracy goes to infinity, we will get a definition of T at this level. The proof of the completeness is also a combination of those for knowledge and for probability. $\square$

Aumann's interactive epistemology deals with the logic of knowledge and belief when there is more than one agent. First we handle the simplest case when there is only one agent. As you may expect, there should be some conditions that connect knowledge with probability.

DEFINITION 5.2.3. A *knowledge-belief* fame F is a tuple $\langle \Omega, R, c, \mathcal{A}, T\rangle$ where $\langle \Omega, R\rangle$ is a knowledge frame and $\langle \Omega, \mathcal{A}, T\rangle$ is a probability frame such that

- c is a $\mathcal{A}$-measurable function, or for any $A \in \mathcal{A}$, $\{w : \Pi(w) \subseteq A\} \in \mathcal{A}$;
- $T(w)(\Pi(w)) = 1$ for all $w \in \Omega$;
- For any $s \in \Pi(t)$, $T(s) = T(t)$.

A *knowledge-belief model* M on the frame F is a tuple $\langle \Omega, R, \mathcal{A}, T, v\rangle$ where v is a valuation such that $v(p) \in \mathcal{A}$ for all propositional letters p.

Let Σ_{kb} be the system Σ_{S5} with the following additional axioms which deal with the interaction between knowledge and belief:

- $K\phi \to L_1\phi$;
- $L_r\phi \to L_1 L_r\phi$;
- $\neg L_r\phi \to L_1\neg L_r\phi$.

The first axiom says that, if the agent knows something, he believes it with certainty. The second and the third say that the agent knows his probability distribution. Our main task is to show the soundness and completeness of Σ_{kb}.

THEOREM 5.2.4. *ϕ is valid in the class of all knowledge-belief frames if and only if $\vdash_{\Sigma_{kb}} \phi$.*

PROOF. Soundness is straightforward. Now we only show the completeness. Assume that ϕ is consistent. We want to prove that ϕ is satisfiable in a knowledge-belief frame. As usual, we define a local language $\mathcal{L}[\phi]$ and it induces a set $\Phi[\phi]$ of formulas in this local language. Define Ω to be the set of all maximal consistent subsets of $\Phi[\phi]$. Elements of Ω are called atoms. For each atom s, $k(s)$ denotes the subset of s consisting of all formulas of the forms $K\psi, \neg K\psi, L_r\psi$ or $\neg L_r\psi$ and $l(s)$ denotes the subset of s consisting of formulas only of the forms $L_r\psi$ or $\neg L_r\psi$. Define sRt if $k(s) = k(t)$. It is easy to see that R is an equivalence relation. Let Π denote the partition associated with this equivalence relation R. It is easy to see that sRt if and only if $s \in \Pi(t)$. The definition of T is totally determined by the subset $l(s)$ for any atom s. In order to define T at all atoms s with the same $l(s)$, we choose one, say s_0, among these atoms and fix it. Then, by applying the same technique as in the completeness proof for Σ_+ and Σ_H, we can define a probability measure $T(s_0)$ and set probability measures at all other atoms s with the same $l(s) = l(s_0)$ to be the same as $T(s_0)$. That is to say, $T(s) = T(s_0)$ if $l(s) = l(s_0)$. Such defined partition Π and type function T satisfy the following frame conditions:

- for any atoms s and $t \in \Omega$, if $s \in \Pi(t)$, then $T(s) = T(t)$;
- for any atom $s \in \Omega$, $T(s)(\Pi(s)) = 1$.

The first part follows from the fact that, for any atoms s and t, $k(s) = k(t)$ implies $l(s) = l(t)$. For the second part, it suffices to show that $L_r \to L_1 L_r\psi, \neg L_r\psi \to L_1\neg L_r\psi, K\psi \to L_1 K\psi$ and $\neg K\psi \to L_1\neg K\psi$ are provable in Σ_{kb}. These are obvious. Define the canonical model $M := \langle \Omega, R, 2^\Omega, T, v\rangle$ where $v(p) := \{s \in \Omega : p \in s\}$. It remains to show the truth lemma:

$$\text{(Truth Lemma) for any formula } \psi \in \Phi(\phi),\ M, s \models \psi \text{ iff } \psi \in s.$$

The proof for the base case and Boolean cases are straightforward and the proof for the probability case is the same as that in Σ_H. Here we only show the case: $\psi := K\psi'$. Assume that $M, s \models K\psi'$. We want to show that $K\psi' \in s$.

Claim 5.2.5. *If $K\psi' \notin s$, $k(s) \cup \{\neg\psi'\}$ is consistent.*

Proof of the claim Suppose that $k(s) \cup \{\neg\psi'\}$ is inconsistent. It follows that $\bigwedge k(s) \to \psi'$ is provable and hence $K(\bigwedge k(s)) \to K\psi'$ is provable. Note that $\bigwedge k(s) \to K(\bigwedge k(s))$ is provable because $K(\psi_1) \wedge K(\psi_2) \to K(\psi_1 \wedge \psi_2), K\chi \to KK\chi, \neg K\chi \to K\neg K\chi, L_r\chi \to KL_r\chi$ and $\neg L_r\chi \to K\neg L_r\chi$ are all provable. These implies that $\bigwedge k(s) \to K\psi'$ is provable. Since $K\psi' \in \Phi(\phi)$, $K\psi' \in s$. So we have finished the proof of the claim.

Suppose that $K\psi' \notin s$. By the above claim, we know that $k(s) \cup \{\neg\psi'\}$ is consistent. So there is an atom s' such that $k(s) \cup \{\neg\psi'\} \subseteq s'$. This implies that sRs' and $\psi' \notin s'$ and hence by induction hypothesis $M, s' \not\models \psi'$. So there is an atom s' such that sRs' and $M, s' \not\models \psi'$. Therefore, $M, s \not\models K\psi'$, which contradicts our assumption that $M, s \models K\psi'$. We conclude that $K\psi' \in s$.

The other direction is much easier to show. Assume that $K\psi' \in s$. We want to show that $M, s \models K\psi'$. For any atom s' such that $sRs', i.e. k(s) = k(s')$ and $\psi' \in s'$. Indeed, $K\psi' \in s'$ and $\psi' \in s'$ because of the truth axiom $K\psi' \to \psi'$. That is to say, $M, s \models K\psi'$. We have finished the proof of the other direction of the truth lemma and hence of the whole theorem.

□

Recall that Σ_H is the probability logic for Harsanyi type spaces. After we have shown the above theorem, it follows immediately that Σ_{kb} is conservative over Σ_H.

Corollary 5.2.6. *Let $\mathcal{L}_H$ and $\mathcal{L}_{kb}$ be the languages for the logics Σ_H and Σ_{kB}, respectively. For any formula ϕ in the smaller language $\mathcal{L}_H$, $\vdash_{\Sigma_{kb}} \phi$ if and only if $\vdash_{\Sigma_H} \phi$.*

Just like in our logic for Harsanyi type spaces, we show that each maximal consistent set Γ of formulas has one and only one maximal consistent extension in the language that is expanded only by increasing the depth by 1.

Lemma 5.2.7. *If ϕ is a Boolean combination of formulas of the forms $L_r\psi$ or $K\psi'$, then $\phi \to K\psi$ (and hence $\phi \to L_1\psi$) is provable in Σ_{kb}.*

Proof. In order to show this lemma, we need the following reasoning:

$$\begin{aligned} K\psi &\to KK\psi \\ &\to L_1K\psi \end{aligned}$$

□

Recall that $\Omega(q,d,w)$ is the set of all maximal consistent sets of formulas in the language $\mathcal{L}(q,d,w)$. By using propositional reasoning, we can easily show that, for each atom $\Gamma \in \Omega(q,d,w)$, the conjunction of formulas in it has the following normal form $\gamma_0 \wedge \gamma_k \wedge \gamma_p$ where γ_0 is a conjunction of propositional letters or their negations, γ_k the conjunction of formulas of the forms $K\psi'$ or their negations, and γ_p the conjunction of formulas of the forms $L_r\psi'$ or their negations.

THEOREM 5.2.8. *Let Γ_1, Γ_2 and Γ_3 be three atoms in $\Omega(q,d,w)$ and γ_1, γ_2 and γ_3 be the conjunctions of formulas in these three atoms, respectively. Their normal forms are $\phi_0^1 \wedge \phi_k^1 \wedge \phi_p^1$, $\phi_0^2 \wedge \phi_k^2 \wedge \phi_b^2$ and $\phi_0^3 \wedge \phi_k^3 \wedge \phi_p^3$. Then the following three propositions hold:*

(1) *if both $\phi_k^2 \wedge \phi_b^2$ and $\phi_k^3 \wedge \phi_p^3$ are different from $\phi_k^1 \wedge \phi_p^1$, then $\gamma_1 \to M_0(\gamma_2 \vee \gamma_3)$ is provable in Σ_{kb};*

(2) *if only one of them, say, $\phi_k^2 \wedge \phi_p^2$, is different from $\phi_k^1 \wedge \phi_p^1$, then*
   (a) *$\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3)$ is provable in Σ_{kb} whenever $\alpha^{\Gamma_1}_{\phi_0^3} = \beta^{\Gamma_1}_{\phi_0^3}$;*
   (b) *$\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge \neg M_{\alpha^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge \neg L_{\beta^{\Gamma_1}_{\phi_0^3}}(\gamma_2 \vee \gamma_3)$ is provable in Σ_{kb} whenever $\alpha^{\Gamma_1}_{\phi_0^3} < \beta^{\Gamma_1}_{\phi_0^3}$;*

(3) *if none of these two is different from $\phi_k^1 \wedge \phi_p^1$, then*
   (a) *$\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}}(\gamma_2 \vee \gamma_3)$ is a theorem of Σ_{kb} whenever $\alpha^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3} = \beta^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}$;*
   (b) *$\gamma_1 \to L_{\alpha^{\Gamma_1}_{\phi^2_{=0} \vee \phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge \neg M_{\alpha^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge M_{\beta^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}}(\gamma_2 \vee \gamma_3) \wedge \neg L_{\beta^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}}(\gamma_2 \vee \gamma_3)$ is a theorem of Σ_{kb} whenever $\alpha^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3} < \beta^{\Gamma_1}_{\phi_0^2 \vee \phi_0^3}$;*

PROOF. The proof here is similar to that in probability logic for Harsanyi type spaces. We exploit the theorem that $K\psi \to L_1 K\psi$ is provable in Σ_{kb}. $\square$

The following lemmas is a generalization of Aumann's results to include probability formulas.

LEMMA 5.2.9. *$K(\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5) \leftrightarrow (K\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5)$ is provable in Σ_{kb}.*

PROOF. For the direction from left to right, we need the axioms: $K\phi_2 \to KK\phi_2, \neg K\phi_3 \to K\neg K\phi_3, L_r\phi_4 \to KL_r\phi_4$ and $\neg L_s\phi_5 \to K\neg L_s\phi_5$.

For the other direction, we use metareasoning. Assume that $K(\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5)$. Of course, $\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5$. If $K\phi_2, \neg K\phi_3, L_r\phi_4$ or $\neg L_s\phi_5$, then we are done. Otherwise, $\neg K\phi_2, K\phi_3, \neg L_r\phi_4$ and $L_s\phi_5$. This implies that $K(\neg K\phi_2 \wedge K\phi_3 \wedge \neg L_r\phi_4 \wedge L_s\phi_5)$. Since $K(\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5)$, $K(\phi_1 \wedge (\neg K\phi_2 \wedge K\phi_3 \wedge \neg L_r\phi_4 \wedge L_s\phi_5))$. Immediately, we have $K(\phi_1)$ and hence $K\phi_1 \vee K\phi_2 \vee \neg K\phi_3 \vee L_r\phi_4 \vee \neg L_s\phi_5$.

□

COROLLARY 5.2.10. *Let ϕ be a conjunction of formulas of the forms $K\psi$, $L_r\psi'$ or their negations. Then, for any formula ϕ', $K(\phi' \vee \phi) \leftrightarrow (K\phi' \vee \phi)$ is provable in Σ_{kb}.*

THEOREM 5.2.11. *(Normal Form Theorem) Any formula in the language $\mathcal{L}_{kb}$ is equivalent to a conjunction of formulas of the following form*

$$\gamma_0 \wedge \bigwedge_{i=1}^{m} L_{r_i}\phi_i \wedge \bigwedge_{j=1}^{n} \neg L_{s_j}\psi_j \wedge K\gamma_1 \wedge \neg K\gamma_2$$

where γ_0, γ_1 and γ_2 are Boolean combinations of propositional letters.

PROOF. Given a formula ϕ in the language $\mathcal{L}_{kb}$, we know by using propositional reasoning that it has a disjunctive normal form. Here we concentrate only on one of its conjunctions

$$\gamma_0 \wedge \bigwedge_{i=1}^{m} L_{r_i}\phi_i \wedge \bigwedge_{j=1}^{n} \neg L_{s_j}\psi_j \wedge K\phi_1 \wedge \neg K\phi_2$$

Note that ϕ_1 and ϕ_2 are not necessarily Boolean combinations of propositional letters. It suffices to show that $K\phi_1$ is equivalent to $K\gamma_1$ and $\neg K\phi_2$ is equivalent to $\neg K\gamma_2$ for some formulas γ_1 and γ_2 which are of depth 2. According to propositional reasoning, ϕ_1 is equivalent to a conjunction of disjunctions:

$$\gamma \vee \bigvee_{i=1}^{m} L_{r_i}\phi_i \vee \bigvee_{j=1}^{n} \neg L_{s_j}\psi_j \vee \bigvee_{k=1}^{m'} K\delta_{t_k} \vee \bigvee_{l=1}^{n'} \neg K\tau_{u_l}$$

where γ is of depth 0. From the above lemmas, it follows that $K\phi_1$ is equivalent to the corresponding conjunction of the following disjunctions:

$$K\gamma \vee \bigvee_{i=1}^{m} L_{r_i}\phi_i \vee \bigvee_{j=1}^{n} \neg L_{s_j}\psi_j \vee \bigvee_{k=1}^{m'} K\delta_{t_k} \vee \bigvee_{l=1}^{n'} \neg K\tau_{u_l}$$

Next we reduce $K\delta_{t_k}$ and $\neg K\tau_{u_l}$. Note that the lengths of δ_{t_k} and τ_{u_l} are both smaller than that of ϕ_1. So we can repeat the above procedure. Finally, we can reduce to formulas all of whose K formulas of depth 1. Similarly, we can reduce $K\phi_2$ to formulas all of whose K-formulas are of depth 1. After we finish all these reductions to get a formula whose K formulas are of depth 1, its disjunction normal form satisfies the requirement.

□

COROLLARY 5.2.12. *Each atom $\Gamma(q,d,w) \in \Omega(q,d,w)(d \geq 1)$ has one and only one maximal consistent extension in $\Omega(q,d+k,w)(k \geq o)$.*

PROOF. Suppose that $\Gamma(q,d,w)$ has two maximal consistent extensions $\Gamma_1(q,d+k,w)$ and $\Gamma_2(q,d+k,w)$ in $\Omega(q,d+k,w)$. According to the above normal form theorem, the conjunctions of formulas in $\Gamma_1(q,d+k,w)$ and $\Gamma_2(q,d+k,w)$ are equivalent to

- $\gamma_0^1 \wedge \bigwedge_i L_{r_i}\phi_i^1 \wedge \bigwedge_j \neg L_{s_j}\psi_j^1 \wedge K\gamma_1^1 \wedge \neg K\gamma_2^1$, and
- $\gamma_0^2 \wedge \bigwedge_i L_{r_i}\phi_i^2 \wedge \bigwedge_j \neg L_{s_j}\psi_j^2 \wedge K\gamma_1^2 \wedge \neg K\gamma_2^2$, respectively.

Since both are maximal consistent extensions of $\Gamma(q,d,w)$, $K\gamma_1^1 \wedge \neg K\gamma_2^1 \leftrightarrow K\gamma_1^2 \wedge \neg K\gamma_2^2$ is provable in Σ_{kb}. Moreover, their probability parts are determined by the probability part of $\Gamma(q,d,w)$. Therefore the conjunction of formulas in $\Gamma_1(q,d+k,w)$ is equivalent to the conjunction of formulas in $\Gamma_2(q,d+k,w)$. So $\Gamma(q,d,w) \in \Omega(q,d,w)(d \geq 1)$ has one and only one maximal consistent extension in $\Omega(q,d+k,w)(k \geq o)$.

□

This corollary shows the simplicity of one-agent epistemology.

5.3. Adding Common Knowledge

Now we discuss the epistemology among several agents. Let I be the population, which is finite. Correspondingly, we associate i with knowledge and belief operators. K_i and

L_r^i[2] are intended to mean "the agent i knows" and "i assigns probability at least r", respectively. The indexes i's in knowledge frames and probability frames also indicate the agents. Note that, in the knowledge-probability frame $\langle \Omega, (\Pi_i)_{i\in I}, (c_i)_{i\in I}, \mathcal{A}, (T_i)_{i\in I}\rangle$, different agents share the same σ-algebra. Since we deal with a group of agents, we can define an additional operator $E : E\phi := \bigwedge_{i\in I} K_i\phi$. It means that "everyone in the group knows". Further we add to our language an "infinitary" operator C. $C\phi$ is intended to be $\bigwedge_{n=1}^{\infty} E^{(n)}\phi$, where $E^{(n)}\phi := E(E^{(n-1)}\phi)$.

LEMMA 5.3.1. *If $A \in \mathcal{A}$, then $E^{-1}(A) := \{w : \Pi_i(w) \subseteq A$ for all $i \in I\} \in \mathcal{A}$ and hence $C^{-1}(A) := \bigcap_{i=1}^{\infty} E^{-n}(A) \in \mathcal{A}$ where $E^{-n}(A)$ is defined inductively as $E^{-n}(A) := E^{-1}(E^{-(n-1)}(A))$.*

PROOF. This follows directly from the facts that c_i are all measurable and $\mathcal{A}$ is closed under countable intersection.

□

A state t is *reachable from s in k steps* if there exist states $s_0, s_1, \cdots, s_k$ such that $s_0 = s, s_k = t$ and for all j with $0 \le j \le j-1$, there exists an agent $i_j \in I$ such that $(s_j, s_{j+1}) \in R_{i_j}$. We say that t is *reachable from s* if t is reachable from s in k steps for some $k \ge 1$.

LEMMA 5.3.2.
(1) *If $(M, s) \models E\phi$ iff $(M, t) \models \phi$ for all t such that sR_it for all $i \in I$;*
(2) *if $(M, s) \models C\phi$ iff $(M, t) \models \phi$ for all t that is reachable from s.*

Another important thing that we need to note is that, if $[[\phi]]$ is measurable, i.e. in $\mathcal{A}$, then $[[E\phi]] \in \mathcal{A}$ and $[[C\phi]] \in \mathcal{A}$. This follows directly from the above two lemmas. There is another more natural interpretation of common knowledge. Let $\mathcal{K}_i$ be the set of unions of atoms of the partition Π_i, which are called *unfields* of agents i. Define $\mathcal{K}_c$ to be the intersection of $\mathcal{K}_i (i \in I)$. Π_c denotes the partition associated with this ufield.

[2]This is the price that we have to pay for the notation from Heifetz and Mongin [17]. Other logicians use L_i^r.

LEMMA 5.3.3. *Let* $\{A_n\}_n$ *is a decreasing sequence of events in* $\mathcal{A}$. *For all* $i \in I$, $K_i^{-1}(\bigcap_n A_n) = \bigcap_n K_i^{-1}(A_n)$.

THEOREM 5.3.4. *If* $A \in \mathcal{A}$, *then* $C^{-1}(A)$ *is the largest event in* $\mathcal{K}_c$ *that is included in* A. *That is,* $C^{-1}(A) = \{w \in \Omega : \Pi_c(w) \subseteq A\}$.

PROOF. Assume that $\Pi_c(w) \subseteq A$. It follows that $\Pi_c(w) = E^{-1}(\Pi_c(w)) \subseteq E^{-1}(A)$. By iterating this argument, we will get:

$$\Pi_c(w) = E_{-n}(\Pi_c(w)) \subseteq E^{-n}(A) \text{ for all } n \geq 1.$$

This implies that $w \in \Pi_c(w) \subseteq \bigcap_n E^{-n}(A) = C^{-1}(A)$. We have shown that $C^{-1}(A) \supseteq \{w \in \Omega : \Pi_c(w) \subseteq A\}$.

Now we show the other direction. But first we prove a claim:

CLAIM 5.3.5. $C^{-1}(A) \in \mathcal{K}_i$ *for all* $i \in I$.

Note that

$$C^{-1}(A) \subseteq \bigcap_{n\geq 2} E^{-1}(E^{-(n-1)}(A)) \subseteq \bigcap_{n\geq 2} K_i^{-1}(E^{-(n-1)}(A)) \subseteq$$
$$K_i^{-1}(\bigcap_{n\geq 2}(E^{-(n-1)}(A))) \subseteq K_i^{-1}(\bigcap_{n\geq 1}(E^{-n)}(A))) \subseteq K_i^{-1}(C^{-1}(A)) \subseteq C^{-1}(A).$$

This implies that $C^{-1}(A) = K_i^{-1}(C^{-1}(A)) \in \mathcal{K}_i$ for all $i \in I$ and hence $C^{-1}(A) \in \mathcal{K}_c$. Moreover, $C^{-1}(A) \subseteq A$. Indeed, $C^{-1}(A) \subseteq K_i^{-1}(A) \subseteq A$. Therefore, $C^{-1}(A) \subseteq \{w \in \Omega : \Pi_c(w) \subseteq A\}$. We have finished the proof of the theorem.

□

COROLLARY 5.3.6. C^{-1} *is a knowledge operator with the associated knowledge ufield* $\mathcal{K}_c$*: for any* $A \in \mathcal{A}$,

(1) $C^{-1}(A) \subseteq A$;
(2) $C^{-2}(A) = C^{-1}(A)$;
(3) $C^{-1}(\Omega - C^{-1}(A)) = \Omega - C^{-1}(A)$.

COROLLARY 5.3.7. $s \in \Pi_c(t)$ *if and only if* t *is reachable from* s.

A c-knowledge-belief frame F_c is a tuple $\langle \Omega, (\Pi)_{i \in I}, \Pi_c, (c_i)_{i \in I}, \mathcal{A}, (T_i)_{i \in I} \rangle$ where

- $\langle \Omega, \Pi_i, c_i, \mathcal{A}, T_i \rangle$ is a knowledge-belief frame for each agent i;
- Π_c is the partition associate with the common knowledge operator.

Now we want to devise a formal system for the above semantics. Although C has an infinitary flavor, we can use a finite machinery to axiomatize it. The deductive system Σ_{ckb} is the system Σ_{kb} plus the following two axioms for common knowledge[3]:

(1) $C\phi \rightarrow E(\phi \wedge C\phi)$;

(2) From $\phi \rightarrow E(\psi \wedge \phi)$ infer $\phi \rightarrow C\psi$.

The rest of this section is devoted to show the following theorem:

THEOREM 5.3.8. *For any formula ϕ, ϕ is valid in the class of c-knowledge-belief frames iff $\vdash_{\Sigma_{ckb}} \phi$.*

PROOF. The proof of the soundness is straightforward. Now we only show the completeness. Assume that ϕ is consistent. We want to prove that it is satisfiable in a c-knowledge-belief frame. Since we have to deal with the common knowledge operator C, our filtration will be a combination of the usual filtration method for knowledge and our filtration method for probabilities. Define $Sub(\phi)$ to the set of all subformulas of ϕ that is closed under simple negation and that contains $E(\psi \wedge C\psi), \psi \wedge C\psi, K_i(\psi \wedge C\psi)(i \in I)$ for each subformula $C\psi$ of ϕ. Let q be the smallest common denominator of the indexes appearing in ϕ and $\mathcal{Q}_q$ is the set of multiples of $\frac{1}{q}$ between 0 and 1. $\Phi(\phi)$ denote the smallest superset of $Sub(\phi)$ that satisfies the following condition:

- if $L_r^i\psi \in \Phi(\phi)$, then $L_t^i\psi \in \Phi(\phi)$ for all $t \in \mathcal{Q}_q$.

First note that $\Phi(\phi)$ is finite. Ω denotes the set of all maximal consistent subsets of $\Phi(\phi)$. For an atom w, define $k_i(w)$ consisting of all formulas in w of the form $K_i\psi$ or $\neg K_i\psi$ and $l_i(w)$ consisting of all formulas in w of the form $L_r^i\psi$ or $\neg L_r^i\psi$. Moreover, we define

- $s \in \Pi_i(t)$ iff $k_i(s) = k_i(t)$ and $l_i(s) = l_i(t)$;
- T_i as in the section for Σ_+;

[3]These axioms are due to Segerberg for PDL.

- $\mathcal{A} := 2^{\Omega}$;
- arbitrary choice function will work for the following completeness proof because the space is finite;
- Π_c is the finest common coarsening of Π_i for $i \in I$;
- $v(p) = \{s \in \Omega : p \in s\}$.

It is easy to check that such a defined frame $F_c := \langle \Omega, (\Pi_i)_{i\in I}, \Pi_c, (c_i)_{i\in I}, \mathcal{A}, (T_i)_{i\in I}\rangle$ is a c-knowledge-belief frame especially,

- $T_i(w)(\Pi_i(w)) = 1$;
- if $s \in \Pi_i(t)$, then $T_i(s) = T_i(t)$.

It remains to show the truth lemma:

(Truth Lemma) for all $\psi \in \Phi(\phi)$, $(F_c, v), s \models \psi$ if and only if $\psi \in s$.

It is trivial for the base case and the boolean cases. For the probability case, the proof is similar to that in the truth lemma for Σ_+. For the knowledge case, the proof is the same as that in the proof for Σ_{kb}. We only need to show the case for common knowledge: $\psi := C\psi'$. The following proof is a generalization of the proof in [14] for common knowledge to our case including probability. If S is a set of formulas, ϕ_S denotes the conjunction of the formulas in S.

Assume that $C\psi' \in s$. We want to show that $(F_c, v), s \models C\psi'$. It suffices to show that for any t reachable from s, $(F_c, v), C\psi', \psi' \in t$. Indeed, since the atoms in $\Pi_c(s)$ are exactly those reachable from s, it follows directly from the induction hypothesis that $(F_c, v), s \models C\psi'$. We prove that by induction on the number n of steps of t from s. When $n = 1$, for any t such that $t \in \Pi_i(s)$, $C\psi' \in t$ and hence $\psi' \in t$. Indeed, $C\psi' \in s$ and hence $K_i(\psi' \wedge C\psi') \in s$. Since $t \in \Pi_i(s)$, $K_i(\psi' \wedge C\psi') \in t$ and hence $C\psi', \psi' \in t$. Note that all these formulas are in $\Phi(\phi)$. Now we assume that it is true for the case $n = n'$. Let t be reachable in $n' + 1$ steps from s. Then there is an atom s' that can be reached n' steps from s. By induction hypothesis, $C\psi', \psi' \in s'$. The proof for the base case also shows that $C\psi', \psi' \in t$. We have finished the induction proof and hence that $(F_c, v), s \models C\psi'$.

Next we show the more difficult direction. Assume that $(F_c, s) \models C\psi'$. We want to show that $C\psi \in s$. Define $\mathcal{C} := \{s \in \Omega : (F_c, s) \models C\psi'\}$. $\phi_{\mathcal{C}} := \bigvee_{s \in \mathcal{C}} \phi_s$.

Claim 5.3.9. *If $s \in \mathcal{C}$, $s' \notin \mathcal{C}$ and $\phi_s \wedge \neg K_i \neg \phi_{s'}$ is consistent, then $k_i(s) \cup l_i(s) \subseteq s'$.*

Assume that $s \in \mathcal{C}$, $s' \notin \mathcal{C}$ and $\phi_s \wedge \neg K_i \neg \phi_{s'}$ is consistent. Suppose that $k_i(s) \cup l_i(s) \nsubseteq s'$. There are two cases that we need to consider. The first case is for the knowledge operators, say, $K_i\psi' \in s$ but $K_i\psi' \notin s'$. It follows that $\vdash K_i\psi' \to \neg\phi_{s'}$ and $\vdash \phi_s \to K_i\psi'$. Reason inside Σ_{kb}:

$$
\begin{array}{rcl}
\phi_s & \to & K_i\psi' \\
K_i\psi' & \to & K_iK_i\psi' \\
K_iK_i\psi' & \to & K_i\neg\phi_{s'} \\
\phi_s & \to & K_i\neg\phi_{s'}
\end{array}
$$

But this contradicts the assumption that $\phi_s \wedge \neg K_i \neg \phi_{s'}$ is consistent. So we have finished the proof for the first case. For the second case, assume that there is formula , say, $\neg L_r^i \psi \in s$ but $\neg L_r^i \notin s'$. Then it follows that $\phi_s \to \neg L_r^i\psi$ and $\neg L_r^i\psi \to \neg\phi_{s'}$ are propositional tautologies. Reason inside Σ_{ckb}:

$$
\begin{array}{rcl}
\phi_s & \to & \neg L_r^i\psi \\
\neg L_r^i\psi & \to & K_i\neg L_r^i\psi \\
K_i\neg L_r^i\psi & \to & K_i\neg\phi_{s'} \\
\phi_s & \to & K_i\neg\phi_{s'}
\end{array}
$$

But this contradicts our assumption that $\phi_s \wedge \neg K_i \neg \phi_{s'}$ is consistent. We have finished the proof for the second case and hence the proof of the claim.

Claim 5.3.10. *if $s \in \mathcal{C}$ and $s' \notin \mathcal{C}$, then $\phi_s \to K_i(\neg\phi_{s'})$ is provable in Σ_{ckb}.*

Suppose that $\phi_s \rightarrow K_i(\neg\phi_{s'})$ is not provable in Σ_{ckb}. It follows that $\phi_s \wedge \neg K_i(\neg\phi_{s'})$ is consistent. By the above claim, we know that $k_i(s) \cup l_i(s) \subseteq s'$. According to our definition, $s \in \Pi_i(s')$. By our assumption that $(F_c, s) \models C\psi'$ and soundness, $(F_c, s) \models K_i(\psi' \wedge C\psi')$ and hence $(F_c, s), s' \models C\psi'$. But this contradicts our assumption that $s' \notin \mathcal{C}$. So we have finished the proof of the above claim.

CLAIM 5.3.11. *If $s \in \mathcal{C}$ and $\phi_s \wedge \neg K_i\psi'$ is consistent, then $k_i(s) \cup l_i(s) \cup \{\neg\psi'\}$ is consistent.*

Assume that $s \in \mathcal{C}$. Suppose that $k_i(s) \cup l_i(s) \cup \{\neg\psi'\}$ is not consistent. Then, reason inside Σ_{ckb}:

$$\begin{array}{rcl}
\bigwedge k_i(s) \wedge \bigwedge l_i(s) & \rightarrow & \psi' \\
K_i(\bigwedge k_i(s) \wedge \bigwedge l_i(s)) & \rightarrow & K_i\psi' \\
\bigwedge k_i(s) \wedge \bigwedge l_i(s) & \rightarrow & K_i(\bigwedge k_i(s) \wedge \bigwedge l_i(s)) \\
\bigwedge k_i(s) \wedge \bigwedge l_i(s) & \rightarrow & K_i\psi'
\end{array}$$

So $\phi_s \wedge \neg K_i\psi'$ is not consistent. Hence we showed the claim.

CLAIM 5.3.12. *If $s \in \mathcal{C}$, then $\phi_s \rightarrow K_i\psi'$ is provable in Σ_{ckb} for all $i \in I$.*

Suppose that $\phi_s \rightarrow K_i\psi'$ is not provable in Σ_{ckb}. It means that $s \in \mathcal{C}$ and $\phi_s \wedge \neg K_i\psi'$ is consistent. By the preceding lemma, $k_i(s) \cup l_i(s) \cup \{\neg\psi'\}$ is consistent. So there is an atom s' which is a superset of $k_i(s) \cup l_i(s) \cup \{\neg\psi'\}$. And $s' \in \Pi_i(s)$. By induction hypothesis, $(F_c, v), s' \models \neg\psi'$. But this contradicts the assumption that $(F_c, v), s \models C\psi'$.

CLAIM 5.3.13. *$\phi_{\mathcal{C}} \rightarrow E(\psi' \wedge \phi_{\mathcal{C}})$ is provable in Σ_{ckb}.*

Reason inside Σ_{ckb}:

$$
\begin{aligned}
\phi_s &\rightarrow K_i(\neg\phi_{s'}), \forall s \in \mathcal{C}, s' \notin \mathcal{C} \\
\phi_s &\rightarrow \bigwedge_{s'\notin\mathcal{C}} K_i(\neg\phi_{s'}) \\
\phi_s &\rightarrow K_i(\bigwedge_{s'\notin\mathcal{C}} (\neg\phi_{s'})) \\
\phi_s &\rightarrow K_i\psi' \\
\phi_s &\rightarrow K_i(\psi' \wedge \bigwedge_{s'\notin\mathcal{C}} (\neg\phi_{s'})) \\
\phi_{\mathcal{C}} &\leftrightarrow \bigwedge_{s'\notin\mathcal{C}} (\neg\phi_{s'}) \\
\phi_s &\rightarrow K_i(\psi' \wedge \phi_{\mathcal{C}}), \forall s \in \mathcal{C} \\
\phi_{\mathcal{C}} &\rightarrow E(\psi' \wedge \phi_{\mathcal{C}})
\end{aligned}
$$

By applying the rule, we know that $\vdash_{ckb} \phi_{\mathcal{C}} \rightarrow C\psi'$. Since $s \in \mathcal{C}$, we have $\vdash_{\Sigma_{ckb}} \phi_s \rightarrow C\psi'$. So $C\psi' \in s$. Now we have finished the other direction of the truth lemma and hence the whole theorem.

□

5.4. Agreeing to Disagree

In Aumann's famous paper [2], he talked about the relationship between knowledge and posteriors on probability spaces under the common-prior assumption. In this section, we show that they can also be translated into the framework of Aumann's knowledge-belief systems. For simplicity, here we assume that the the group consists of two players: player 1 and player 2. $F_P = \langle \Omega, (\Pi_i)_{i=1,2}, \Pi, (c_i)_{i=1,2}, \mathcal{A}, P_1, P_2 \rangle$ is a probability space satisfying the the following conditions:

- P_1 and P_2 are the priors for player 1 and 2, respectively;
- Π_1 and Π_2 are the partitions for players 1 and 2 such that the join $\Pi_1 \vee \Pi_2$ consists of nonnull events, which also implies that both partitions are countable;
- Π is the partition associated with the common knowledge between 1 and 2.

We shall show that F_P is a Aumann's knowledge-belief system. Define

$$T_i(w, A) := \frac{P_i(A \cap \Pi_i(w))}{P_i(\Pi_i(w))}.$$

It is easy to check the following properties:

(1) for any $A \in \mathcal{A}$, $T_i(\cdot, A)$ is an $\mathcal{A}$-measurable function;

(2) for any $w \in \Omega$, $T_i(w, \cdot)$ is a probability measure;

(3) for any $w \in \Omega$, $T_i(w)(\Pi_i(w)) = 1$;

(4) for any $s \in \Pi_i(t)$, $T_i(s) = T_i(t)$.

Indeed F_p is an Aumann's knowledge-belief system with common knowledge.

CHAPTER 6

Probabilistic Bisimulation and Finite Approximation

In modal logic, there are two methods to achieve the finite model property: selection and filtration. The first method is based on one of the three slogans from [6]: modal languages provide an internal, local perspective on relational structure. Filtration is used by taking quotient of Kripke structures over a finite language. In contrast, in our modal logic of probability, the first method cannot go through because our language is not local any longer. But the second method is still available. Actually, we have used this method in our proof of the the weak completeness of our system Σ_+. In this section, we will study this method systematically and use this method to approximate probability models from the perspective of *probabilistic bisimilarity.*

6.1. Bisimulation

DEFINITION 6.1.1. Let $\langle \Omega_1, \mathcal{A}_1, T_1, v_1\rangle$ and $\langle \Omega_2, \mathcal{A}_2, T_2, v_2\rangle$ be two probability models. $f : \Omega_1 \to \Omega_2$ is a *type morphism*[1] if it satisfies the following two conditions:

(1) f is surjective;
(2) for any propositional letter p and any state $s_1 \in \Omega_1$, $s_1 \in v_1(p)$ if and only if $f(s_1) \in v_2(p)$;
(3) for any $A_2 \in \mathcal{A}_2$ and $s_1 \in \Omega_1$, $T(f(s_1))(A_2) = T(s_1)(f^{-1}(A_2))$

The counterpart of type morphism in modal logic is called the back-and-forth condition in bisimulation.

DEFINITION 6.1.2. Two probability models $M_1 = \langle \Omega_1, \mathcal{A}_1, T_1, v_1\rangle$ and $M_2 = \langle \Omega_2, \mathcal{A}_2, T_2, v_2\rangle$ are *probabilistically bisimular* (or, simply, bisimular) to each other if there is a third

[1]It is also a coalgebra morphism.

probability model $M_3 = \langle \Omega_3, \mathcal{A}_3, T_3, v_3 \rangle$ with two type morphisms $f_1 : M_3 \to M_1$ and $f_2 : M_3 \to M_2$:

$$\begin{tikzcd} & M_3 \arrow[dl, "f_1"'] \arrow[dr, "f_2"] & \\ M_1 & \leftrightarrows & M_2 \end{tikzcd}$$

For two states $s_1 \in \Omega_1$ and $s_2 \in \Omega_2$, s_1 *is probabilistically bisimular (or, simply, bisimular) to* s_2(denoted as $s_1 \leftrightarrows s_2$) if there is $s_3 \in \Omega_3$ such that $f_1(s_3) = s_1$ and $f_2(s_3) = s_2$.

The following proposition can be regarded as the Henneessy-Milner theorem for probabilistic logic.

THEOREM 6.1.3. *Let $M_1 = \langle \Omega_1, \mathcal{A}_1, T_1, v_1 \rangle$ and $M_2 = \langle \Omega_2, \mathcal{A}_2, T_2, v_2 \rangle$ be two countable probability models. For any two states $s_1 \in \Omega_1$ and $s_2 \in \Omega_2$, $s_1 \leftrightarrows s_2$ if they satisfy the same formulas.*

PROOF. For the direction from left to right, we leave the proof to the reader. Here we only show the other direction. Assume that $s_1 \in \Omega_1$ and $s_2 \in \Omega_2$ satisfy the same formulas. For $s \in \Omega_1$, $[s]_1$ denotes the set $\{w \in \Omega_1 : w$ and s satisfy the same set of formulas $\}$. Similarly, for any $s' \in \Omega_2$, $[s']_2$ denotes the set $\{w' \in \Omega_2 : w'$ and s' satisfy the same set of formulas $\}$. If $s \in \Omega_1$ and $s' \in \Omega_2$, $s \approx s'$ denotes the relationship that they satisfy the same set of formulas. First we show that Larsen-Skou's theorem about probabilistic bisimulation also works here.

CLAIM 6.1.4. *Let $s_1, t_1 \in \Omega_1$ and $s_2, t_2 \in \Omega_2$. If $s_1 \approx s_2$ and $t_1 \approx t_2$, then $T_1(s_1)([t_1]_1) = T_2(s_2)([t_2]_2)$.*

Let Σ_1 be the sigma-algebra generated by the collection of sets of the form $[[\phi]]_1 (= \{w \in \Omega_1 : M_1, s \models \phi\}$. And Σ_2 is defined similarly. Note that $[t_1]_1 \in \Sigma_1$ because $[t_1]_1 = \bigcap_{t_1 \in [[\phi]]_1} [[\phi]]_1$. Similarly, $[t_2]_2 \in \Sigma_2$. Since $s_1 \approx s_2$, $T(s_1)([[\phi]]_1) = T(s_2)([[\phi]]_2)$ for all formulas. Moreover, according to Dynkin's π-λ theorem, we know that

- $T_1(s_1)([t_1]_1) = inf\{\sum_i T_1(s_1)([[\phi^i]]_1) : [t_1]_1 \subseteq \bigcup_i [[\phi^i]]_1\}$;
- $T_2(s_2)([t_2]_2) = inf\{\sum_i T_2(s_2)([[\phi^i]]_2) : [t_2]_2 \subseteq \bigcup_i [[\phi^i]]_2\}$.

It is easy to see that $T_1(s_1)([t_1]_1) = T_2(s_2)([t_2]_2)$ because, for any sequence of formulas ϕ^i, $[t_1]_1 \subseteq \bigcup_i [[\phi^i]]_1\}$ iff $[t_2]_2 \subseteq \bigcup_i [[\phi^i]]_2\}$. So we finished the proof of the claim and hence s_1 and s_2 are bisimular in the sense of Larsen-Skou [25].

Define $\Omega_3 := \{(w_1, w_2) \in \Omega_1 \times \Omega_2 : w_1 \approx w_2\}$. For any $(s_1, s_2), (t_1, t_2) \in \Omega_3$, define

$$T_3((s_1, s_2))((t_1, t_2)) = \tfrac{T_1(s_1)(t_1) T_2(s_2)(t_2)}{T_2(s_2, [t_2]_2)}$$

Note that, for any $(t_1, t_2) \in \Omega_3$,

$$\begin{aligned}
\sum_{(t_1,t_2)\in\Omega_3} T_3((s_1, s_2))((t_1, t_2)) &= \sum_{t_1\in\Omega_1} \sum_{t_2\approx t_1} T_3((s_1, s_2))((t_1, t_2)) \\
&= \sum_{t_1\in\Omega_1} \sum_{t_2\approx t_1} \frac{T_1(s_1)(t_1) T_2(s_2)(t_2)}{T_2(s_2, [t_2]_2)} \\
&= \sum_{t_1\in\Omega_1} T_1(s_1)(t_1) \\
&= 1
\end{aligned}$$

Now define $v_3(p) = \{(w_1, w_2) \in \Omega_3 : w_1 \in v_1(p)\}$ for any propositional letter p. It is easy to see that v_3 is well-defined. $\mathcal{A}_3$ denotes the sigma-algebra generated by the sets of $[[\phi]]_3$ for all formulas ϕ. So $M_3 := \langle \Omega_3, \mathcal{A}_3, T_3, v_3 \rangle$ is a probability model. π_1 and π_2 are the two projections from Ω_3 to Ω_1 and Ω_2, respectively. In order to show that $s_1 \leftrightarrows s_2$, it suffices to show that both π_1 and π_2 are type morphisms. We have assumed that $s_1 \approx s_2$. We have

$$
\begin{aligned}
T_3((w_1,w_2),\pi_2^{-1}(s_2)) &= T_3((w_1,w_2),\{s_1\in\Omega_1 : s_2\approx s_1\}\times s_2)\\
&= \sum_{s_1\in\Omega_2 : s_1\approx s_2}\frac{T_1(w_1,s_1)T_2(w_2,s_2)}{T_2(w_2,[s_2]_2)}\\
&= \sum_{s_1\in\Omega_2 : s_1\approx s_2}\frac{T_1(w_1,s_1)T_2(w_2,s_2)}{T_1(w_1,[s_1]_1)}\\
&= T_2(w_2,s_2)\sum_{s_1\in\Omega_2 : s_1\approx s_2}\frac{T_1(w_1,s_1)}{T_1(w_1,[s_1]_1)}\\
&= T_2(w_2,s_2)
\end{aligned}
$$

It is easier to show that π_2 is a type morphism from Ω_3 to Ω_2. So $s_1 \leftrightarrows s_2$.

□

It is well-known in modal logic that the class of m-saturated Krike models has the above so-called *Hennesy-Milner property.* In probabilistic logic, there is also an analogous theorem. A *Polish* space is the topological space underlying a complete, separable metric space. An *analytic space* is the image of a Polish space under a continuous function.

THEOREM 6.1.5. *Let $M_1 = \langle \Omega_1, \mathcal{A}_1, T_1, v_1\rangle$ and $M_2 = \langle \Omega_2, \mathcal{A}_2, T_2, v_2\rangle$ both be probability models under analytic spaces. For $s_1 \in \Omega_1$ and $s_2 \in \Omega_2$, s_1 is bisimular to s_2 if and only if s_1 and s_2 satisfy the same set of formulas.*

PROOF. For the detailed proof, the reader is referred to [8].

□

Here we owe the reader a counterexample that s_1 and s_2 satisfy the same formulas but they are not bisimular. Let R be a binary relation on the set S. A set $X \subseteq S$ is *R-closed* if $R(X) := \{t : \exists s \in X(sRt)\} \subseteq X$. Now we define a functional F on the set of binary relations on $(S \times S, \subseteq)$:

$$sF(R)t \text{ iff for all } R\text{-closed } C \in \mathcal{A}, T(s)(C) = T(t)(C), \text{ and } sRt.$$

Note that F is monotonic.

THEOREM 6.1.6. *Let $M = \langle \Omega, \mathcal{A}, T, v\rangle$ be a probability model under an analytic space. Define a family of relations $R_i \subseteq S \times S$ on Ω as follows:*

- *sR_0t if s and t satisfy the same propositional letters i,e, for any propositional letter p, $s \models p \Leftrightarrow t \models p$;*
- *$R_{i+1} = F(R_i)$;*
- *$R = \bigcap R_i$.*

Then sRt if and only if s is bisimular to t.

This theorem can be used to guide our following finite approximation.

6.2. Filtration

Alternatively, we can view bisimulation as the maximum fixed point of the functional F. Also the above theorem tells us that, sR_it iff s and t satisfy the same formulas of depth up to i. So we call the relation R_i i-bisimulation. Assume that we can enumerate all the propositional letters in the language: $p_1, p_2, \cdots, p_n, \cdots$. For any natural numbers q, d and n, we define the set $\Phi(q, d, w)$ of formula to be the smallest set of formulas (by taking the quotient of propositional reasoning) satisfying the following conditions:

- it contains propositional letters from p_1 to p_w;
- the indexes should be multiples of the unit $1/q$;
- depthes of formulas should be less than or equal to d;
- closed under propositional tautologies.

The above defined set $\Phi(q, d, w)$ gives rise to a set $\Omega(q, d, w)$ of maximal consistent sets of formulas in $\Phi(q, d, w)$. Given a probability model $M = \langle S, \mathcal{A}, T, v\rangle$, we define a filtered probability model through $\Phi(q, d, w)$. Define

(1) $S(q, d, w) := \{[\wedge\Gamma]_M \neq \emptyset : \Gamma \in \Omega(q, d, n)\}$;
(2) $\mathcal{A}(q, d, w)$ be the power set of $S(q, d, w)$.
(3) $v(p) := \{[\wedge\Gamma] \in S(q, d, w) : p \in \Gamma\}$;

Equivalently, we can get the above structure through an equivalence relation. Define

for any points $s, s' \in S$, $s \sim s'$ if, for all formulas ϕ in $\Phi(q, d, w)$, $M, s \models \phi \Leftrightarrow M, s' \models \phi$.

It is easy to see that the above defined $\sim$ is an equivalent relation. So the equivalence class $|s|$ is actually $[\wedge\Gamma]_M$ for some atom $\Gamma \in \Omega(q, d, w)$. For any $s \in S$, define $\Gamma_s := \{\psi \in$

$\Phi(q,d,w) : M, s \models \psi\}$. Note that Γ_s is an atom in $\Omega(q,d,w)$ and $|s| = [\Gamma_s]_M$. Moreover, for any atom $\Gamma \in \Omega(q,d,w)$ such that $[\wedge\Gamma]_M \neq \emptyset$, there is a $s \in [\wedge\Gamma]_m$ such that $\Gamma_s = \Gamma$. In other words, there is a one-to-one correspondence between the set of equivalence classes and the set of atoms Γ such that $[\wedge\Gamma]_M \neq \emptyset$. In order to define a probability structure on $S(q,d,w)$, it remains to define the transition probability function $T(q,d,w)$ (or Markov kernel). Our definition is based on T. For any $[\wedge\Gamma] \in S(q,d,n)$, since $[\wedge\Gamma] \neq \emptyset$, choose a point $s_\Gamma \in S$ such that $M, s_\Gamma \models \wedge\Gamma$. We define

$$\text{for any } |s| \in S(q,d,w),\ T(q,d,w)(|s|)(|s'|) := T(s_{\Gamma_s})([\wedge\Gamma_{s'}]_M)$$

$T(q,d,w)(|s|)$ is additive on $S(q,d,w)$. It is easy to see that $T(q,d,w)(S(q,d,w)) = 1$. So $T(q,d,n)$ is a probability measure on $S(q,d,n)$. Note that this definition depends on the choice function from the nonempty $[\wedge\Gamma]_M$ to s_Γ. But, the following filtration theorem does not depend on the choice of the function. Let $M(q,d,w) := \langle S(q,d,w), \mathcal{A}(q,d,w), T(q,d,w), v\rangle$ be the filtered probability model.

LEMMA 6.2.1. *For any $s \in S$ and $\phi \in \Phi(q,d,w)$,*

$$M(q,d,w), |s| \models \phi \textit{ iff } \phi \in \Gamma_s.$$

PROOF. We prove this by induction on the formula ϕ.

(1) Base case: ϕ is a propositional letter. This follows directly from the definition of $v(p)$ on the filtered model.
(2) Boolean case. This is straightforward from the fact that $|s|$ is a maximal and consistent set of formulas in the above local language.
(3) Crucial case: $\phi = L_r\phi'$ where r is a multiple of the unit $1/q$.

$$\begin{aligned}
M(q,d,w),|s| \models L_r\phi' &\Leftrightarrow T(q,d,w)(|s|)([\phi']_{M(q,d,w)}) \geq r \\
&\Leftrightarrow T(q,d,w)(|s|)(\{|s'| : M(q,d,w),|s'| \models \phi'\}) \geq r \\
&\Leftrightarrow T(q,d,w)(|s|)(\{|s'| : \phi' \in |s'|\}) \geq r (I.H.) \\
&\Leftrightarrow T(s_{\Gamma_s})(\bigcup\{|s'| : \phi' \in |s'|\}) \geq r \\
&\Leftrightarrow T(s_{\Gamma_s})([\phi']_M) \geq r \\
&\Leftrightarrow L_r\phi' \in \Gamma_{s_{\Gamma_s}} \\
&\Leftrightarrow L_r\phi' \in \Gamma_s
\end{aligned}$$

□

Corollary 6.2.2. *For any formula ϕ and $s \in S$, $T(q,d,w)(|s|)([\phi]_{M(q,d,w)}) = T(s_{\Gamma_s})$ $([\phi]_M)$.*

Theorem 6.2.3. *(Filtration Theorem) Let $M := \langle S, \mathcal{A}, T, v\rangle$ be a probability model and $M(q,d,w) := \langle S(q,d,w), \mathcal{A}(q,d,w), T(q,d,w), v\rangle$ be its filtered probability model through $\Phi(q,d,w)$. Then, for all formulas $\phi \in \Phi(q,d,w)$ and all points $s \in S$,*

$$M, s \models \phi \text{ iff } M(q,d,w), |s| \models \phi.$$

Proof. The proof follows immediately from the above lemma.

$$\begin{aligned}
M(q,d,w),|s| \models \phi &\Leftrightarrow \phi \in \Gamma_s \\
&\Leftrightarrow M, s \models \phi
\end{aligned}$$

.

□

COROLLARY 6.2.4. *Assume that* $s \sim_{(n,n,n)} s'$ *and* ϕ *is a formula in the language of* $M(n-1, n-1, n-1)$. *Then* $|T(s)([\phi]_M) - T(s')([\phi]_M)| \leq 1/n$.

6.2.1. Finite Approximation. This part is motivated by [9]. Their approximation is similar to the approximation that is used in measure theory to show that the set of simple functions is dense in that of measurable functions. But we use the above filtration method to approximate any labeled Markov process. Now we are going to give a sequence of finite approximation of the probability model M. As above, our local languages consists of the following three aspects: depth of formulas, the number of propositional letters and the accuracy of the indexes. First we set some conventions that we will use later:

(1) probability models are invariant under permutation of propositional letters, i.e. for two probability models $M_1 = \langle S_1, \mathcal{A}_1, T_1, v_1 \rangle$ and $M_2 = \langle S_2, \mathcal{A}_2, T_2, v_2 \rangle$, if $S_1 = S_2, \mathcal{A}_1 = \mathcal{A}_2$, $T_1 = T_2$ and, moreover, $v_1(p) = v(\tau(p))$ for some permutation of propositional letters p, then M_1 and M_2 are regarded as the same. In other words, probability models are a kind of equivalence class.
(2) if the number of propositional letters in the language of M is n (finite or countably infinite), we always choose $p_1, p_2, \cdots, p_n$ as a representative for the above corresponding equivalence class.

In order to define probability measures at all points in the filtered models, we need to define a choice function mapping an atom Γ in a local language to an element of $[\wedge\Gamma]_M$. For any Γ in some $\Omega(q, d, w)$ such that $[\wedge\Gamma]_M \neq \emptyset$, $s_- : \Gamma \mapsto s_\Gamma \in [\wedge\Gamma]$. The following theorem follows immediately from the above theorems:

THEOREM 6.2.5. *For any formula* ϕ, *if* ϕ *is a formula in the language of* $M(q, d, w)$, *then* $M(q, d, w), |s|_{(q,d,w)} \models \phi$ *iff* $M, s \models \phi$ *(where* $|s|_{(q,d,w)}$ *is the equivalence class of* s *by taking the quotient of the equivalence relation* $\sim_{(q,d,w)}$*). Moreover,* $\bigcup[\phi]_{M(q,p,w)} = [\phi]_M$.

Now we show that one can reconstruct the original process (actually a bisimulation equivalent of the original process) from the approximants $M(n, n, n)$. We don't reconstruct the original state space, but we reconstruct all the transition probability information, i.e., the dynamic aspects of the process.

THEOREM 6.2.6. *Assume that $M = \langle S, \mathcal{A}, T, v\rangle$ is a minimal probability model in the sense that $M = M'/\approx$ for some probability model M' where $\approx$ is a bisimulaion on M'. If we are given all finite state approximants $M(n,n,n)$, then we can recover M.*

PROOF. We can recover the state space by just taking the union of states at any level of any approximants $M(n,n,n)$. Since M is a minimal probability model, $\mathcal{A}$ is generated by $\mathcal{A}_\Phi := \{[\phi]_M : \phi$ is a formula in the language of $M\}$. It is easy to check that $\mathcal{A}_\Phi$ is a field. For any $s \in S$ and any formula ϕ (in the language of $M(n,n,n)$ for $n > k$ for some k), define

$$\mu(s)([\phi]_M) = \lim_{n\to\infty} T(n,n,n)(|s_{\Gamma_s^n}|)([\phi]_{M(n,n,n)})$$

Note that Γ_s^n is the atom in $\Omega(n,n,n)$ such that $s \in [\wedge\Gamma_s]_M$. According to the definition of the choice function s_-, we know that $s \sim_{(n,n,n)} s_{\Gamma_s}$.

CLAIM 6.2.7. $\lim_{n\to\infty} T(n,n,n)(|s_{\Gamma_s^n}|)([\phi]_{M(n,n,n)})$ *exists.*

In order to show the claim, it suffices to show that, for any ε, there is an N such that, if $n \geq N$, $|T(n,n,n)(|s_{\Gamma_s^n}|)([\phi]_{M(n,n,n)}) - T(n+p,n+p,n+p)(|s_{\Gamma_s^{n+p}}|)([\phi]_{M(n+p,n+p,n+p)})| \leq \varepsilon$ for any natural number p. Given any ε, there is an N such that $1/N < \varepsilon$. Also for any natural number p,

$|T(n,n,n)(|s_{\Gamma_s^n}|)([\phi]_{M(n,n,n)}) - T(n+p,n+p,n+p)(|s_{\Gamma_s^{n+p}}|)([\phi]_{M(n+p,n+p,n+p)})|$
$= |T(s_{\Gamma_s^n})([\phi]_M) - T(s_{\Gamma_s^{n+p}})([\phi]_M)|$
$\leq 1/n$ ($s_{\Gamma_s^{n+p}} \in [s_{\Gamma_s^n}]_{M(n,n,n)}$)
$\leq 1/N$ for $n \geq N$
$\leq \varepsilon$.

By Cauchy's criterion for sequence convergence, we know that the above limit exists. This is to say, $\mu(s)([\phi]_M$ is well-defined.

CLAIM 6.2.8. *For any $s \in S$ and for any formula ϕ, $\mu(s)([\phi]_M) = T(s)([\phi]_M)$.*

The proof of this claim is much simpler. Note that

$$
\begin{aligned}
&|T(s)([\phi]_M) - T(n,n,n)(|s_{\Gamma_s^n}|)([\phi]_{M(n,n,n)})| \\
&= |T(s)([\phi]_M) - T(s_{\Gamma_s^n})([\phi]_M)| \\
&\leq 1/n \to 0.
\end{aligned}
$$

The last inequality follows from the fact that s and $s_{\Gamma_s^n}$ are in the same equivalence class of $\sim_{(n,n,n)}$. So we have shown that, for any $s \in S$, the above defined $\mu(s)$ and $T(s)$ matches on the algebra $\mathcal{A}_\Phi$. This implies that the probability measure $\mu^*(s)$ on the σ-algebra $\mathcal{A}$ generated by $\mu(s)$ is the same as $T(s)$. So we have recovered the original probability model. In other words, the probability model M is uniquely determined by its finite approximants $M(n,n,n)$.

□

6.3. A Countable Basis for Probability Models

Just like the Dedekind's cut for reals, we want to use finite rational probability models (in the sense that all transition probabilities are rationals) to approximate any probability model. In the above finite approximation, our definition of probability measures of the filtered models at equivalence classes are based on some probability model of the original models at some representatives. This also means that, if we want to get finite rational approximations, we cannot get the probability measures on the filtered models from those of the original models for free. We have to define new probability measures. Actually, this is already achieved in our proof of Moss' conjecture by Fourier-Motzkin's elimination method. Given any filtered model $M(q,d,w)$, $M^r(q,d,w)$ is defined the same as $M(q,d,r)$ except the definition of transition probability function T^r: $T^r(|s|)(|s'|)$ is obtained by the elimination method. What is the difference between these two transition probability functions?

Lemma 6.3.1. *Assume that ϕ is a formula in the language of $M(n-1,n-1,n-1)$ and $|s|$ is a point in the model $M(n,n,n)$. Then,*

$$|T(|s|)([\phi]_{M(n,n,n)}) - T^r(|s|)([\phi]_{M(n,n,n)})| \leq 1/n.$$

PROOF. This proof is actually implicit in the proof of the completeness of Σ_+. For ϕ, either both $L_r\phi \in |s|$ and $M_r\phi \in |s|$ for some r or $L_r\phi \wedge \neg L_{r+1/n}\phi \in |s|$. For the first case, $T(|s|)([\phi]_{M(n,n,n)}) = T^r(|s|)([\phi]_{M(n,n,n)}) = r$. For the second case, $r \leq T(|s|)([\phi]_{M(n,n,n)})$, $T^r(|s|)([\phi]_{M(n,n,n)}) < r+1/n$. This implies that $|T(|s|)([\phi]_{M(n,n,n)}) - T^r(|s|)([\phi]_{M(n,n,n)})| \leq 1/n$.

□

THEOREM 6.3.2. *$M^r(n,n,n)$ is also a filtered probability model of M through $\Omega(n,n,n)$, i.e., for any formula ϕ in the language of $M^r(n,n,n)$ and $s \in S$,*

$$M, s \models \phi \text{ iff } M^r(n,n,n), |s| \models \phi.$$

Moreover, $\bigcup[\phi]_{M(n,n,n)} = [\phi]_M$.

THEOREM 6.3.3. *Assume that M is a minimal probability model. Then M is uniquely determined by its rational approximation.*

PROOF. The proof is very similar to that of the above theorem. We define

$$\mu^r(s)([\phi]_M) = \lim_{n\to\infty} T^r(n,n,n)(|s|)([\phi]_{M(n,n,n)})$$

The proof of the existence of the limit is similar to that of the existence above except replacing $T(|s|)$ by $T^r(n,n,n)(|s|)$ because of the above lemma 6.3.1. Using a similar argument as above, we can show that

$$\mu^r(s)([\phi]_M) = T(s)([\phi]_M).$$

□

Since the collection of finite rational probability models is countable and any probability models is an approximations of finite rational probability models in a certain metric space, it is reasonable to expect that the collection of probability models can be endowed with a Polish space. We can also simulate the work in [10] to define a metric on this space. But it is an open problem whether such a defined metric space is complete or not.

CHAPTER 7

Semantics of Probabilistic and Quantum Programs

There is a concurrence between game theory and semantics for programs. In game theory with incomplete information, type functions provide a satisfactory semantics for belief grammar. In addition, they are also used to model primitive probabilistic programs. However, there are fundamental differences. In probability logics for type spaces, we focused on formulas while in dynamic logic of probabilistic programs, the subjects are programs. Probabilistic programs are constructed according to a well-known syntax. As Markovian kernels(or Markov transitions), type functions can be interpreted both as measure-transformers (in forward semantics) and as predicate transformers (backward semantics). Kozen interpreted *while* probabilistic programs as positive linear operators with norms ≤ 1 on the space of measures of some measurable spaces. These two interpretations are dual to each other.

Just as probabilistic programs are used to reason about uncertainty in the classical world, quantum programs are developed to reason about uncertainty in quantum computation [30]. Quantum computation has recently become an important topic in theoretical computer science. Traditionally it is studied at the hardware level. However, in a recent paper [4], Peter Selinger explores quantum computing from the perspective of programming languages. He studied the subject by dealing with both data flow and control flow but not depending on any particular hardware model. His approach to quantum computing can be summarized in the slogan: “quantum data and classical control”. The programs may take quantum data, which are in the form of quantum superpositions while the controls are always classical.

His quantum programming language is *functional*. In other words, each program transforms a set of inputs to outputs. The syntax of this language is in the form of quantum flow charts. They are constructed from a set of basic flow charts by context extension, sequential composition, parallel composition and loops. Each edge of any flow chart can be associated with a pair of typing context and its annotation. The typing context is actually a formal expression of signature for some Hilbert space. And the annotation tells us the density matrix at this state. According to this interpretation, each quantum flow chart maps a density matrix on one Hilbert space to a density matrix on another Hilbert space, which is a superoperator. More importantly, this semantics for quantum flow charts is *full* in the sense that any superoperator can be realized by a quantum flow chart. The crucial step to show this proposition is the normal form theorem for superoperators. It says that each superoperator can be factored as a product of $M \circ E \circ U$ where M is a measurement operator, E an eraser operator and U a subunitary.

There is another backward semantics for quantum programs which is dual to the above program-as-superoperator forward one. In [1], D'Hondt *et.al* provided a quantum weakest precondition semantics for quantum programs. Each program takes an input predicate and output another predicate. Actually this backward semantics is dual to the forward semantics in the sense that the trace of each predicate at the transformed density matrix is equal to the trace of the transformed predicate at that density matrix. This duality is proved through the well-known Kraus representation theorem.

In this paper, we "streamline" the semantics for both probabilistic and quantum programs. In the first part, we present the predecessor of Selinger's semantics, i.e., Kozen's semantics for probabilistic programs. In the second part, the dual semantics for quantum programs are organized in a similar way to Kozen's approach to probabilistic programs. In the last part, we are trying to answer one question raised by Professor Sabry: how to construct unitary matrices from arbitrary matrix with complex entries.

7.1. Semantics for Probabilistic Programs

In this section, we summarize some important results on the semantics for probabilistic programs ([1]).

7.1.1. Background in Probability Theory. In the following, $\mathcal{R}$ and $\mathcal{R}^+$ denote the real numbers and the nonnegative real numbers, respectively. Recall that a *measurable space* is a pair $(S, \mathcal{A})$ where S is a nonempty set of states and $\mathcal{A}$ is a σ-algebra on the power set of S. Elements of $\mathcal{A}$ are called *events*. Given two measurable spaces $(S_1, \mathcal{A}_1)$ and $(S_2, \mathcal{A}_2)$, a function f is *measurable* if $f^{-1}(A) \in \mathcal{A}_1$ whenever $A \in \mathcal{A}_2$. A measure μ on $(S, \mathcal{A})$ is a function $\mathcal{A} \to \mathcal{R}$ that is countably additive and $\mu(\emptyset) = 0$. It is a *probability measure* if it is positive and $\mu(S) = 1$, and is a subprobability measure if it is positive and $\mu(S) \leq 1$. If μ is a measure and $B \in \mathcal{A}$, let μ_B denotes the measure $\mu(A) = \mu(A \cap B)$. Then $\mu_B/\mu(B)$ is called the *conditional probability* relative to B.

Every measure can be decomposed into two positive measures μ^+ and μ^- such that $\mu^+ = \mu_B$ and $\mu^- = -\mu_{\neg B}$ for some $B \in \mathcal{A}$. This is the so-called *Jordan decomposition* of μ. μ^+ and μ^- are called the *positive* and *negative* parts of μ. Define $|\mu| = \mu^+ + \mu^-$. And it is called the *total variation* or *absolute value* of μ. Let $\mathcal{M}$ denote the set of measures on $(S, \mathcal{A})$. Then the above defined total variation introduces a norm on $\mathcal{M}$ as follows: for each μ, $||\mu|| = |\mu|(S)$. A random measure $X : (S_1, \mathcal{A}_1, \mu) \to (S_2, \mathcal{A}_2)$ induces a subprobability measure $\mu \cdot X^{-1}$ on $(S_2, \mathcal{A}_2)$:

$$\mu \cdot X^{-1}(A) = \mu(X^{-1}(A)).$$

If X is total, then $\mu \cdot X^{-1}$ is a probability measure.

Definition 7.1.1. Let $(\mathcal{B}, ||\cdot||)$ and $(\mathcal{C}, ||\cdot||)$ be two normed vector spaces and $T : \mathcal{B} \to \mathcal{C}$ be a linear transformation. T is $||\cdot||$-*bounded* if $\sup||T(x)|| < \infty$ for all $x \in \mathcal{S} = \{x : ||x|| \leq 1\}$. And the space of all $||\cdot||$-bounded linear transformations from $\mathcal{B}$ to $\mathcal{C}$ is a normed vector space under pointwise addition and scalar multiplication with the uniform norm

$||T|| = sup_s||T(x)||$. A *positive cone* $\mathcal{P}$ of $\mathcal{B}$ is a subset of $\mathcal{B}$ satisfying the following two conditions:

(1) $ax + by \in \mathcal{P}$ whenever $x, y \in \mathcal{P}$ and $a, b \geq 0$;
(2) only one of $x, \neg x$ is in $\mathcal{P}$.

It is easy to see that the positive cone $\mathcal{P}$ induces a partial order on $\mathcal{B}$: $x \leq y$ if $x - y \in \mathcal{P}$. $(\mathcal{B}, \mathcal{P})$ is a *vector lattice* if each pair $x, y \in \mathcal{B}$ has a $\leq$ least upper bound. $(\mathcal{B}, \mathcal{P}, ||\cdot||)$ is a *Banach lattice* if it is a Banach space and a vector lattice such that

(1) $\||x|\| = \|x\|$;
(2) if $0 \leq x \leq y$, then $\|x\| \leq \|y\|$.

7.1.2. Probabilistic *while* Programs. The syntax of probabilistic programs is similar to that of classical programs.

7.1.2.1. *Syntax.* We consider the while programs over the variables $x_1, \cdots, x_n$. Basic programs are:

- simple assignment, $x_i := f_i(x_1, \cdots, x_n)$;
- random assignment. $x_i := random$.

As usual, there are three types of program constructs:

- sequential composition, $S; T$;
- conditional, if B then S else T;
- loop, while B do S.

7.1.2.2. *Forward Semantics.* Let $(\Omega, \mathcal{A})$ be a measure space and $\mathcal{B}$ be the vector space of the measures on $(S^n, \mathcal{A}^n)$. $\mathcal{P}$ denotes the set of positive measures on $(\Omega, \mathcal{A})$ and define a norm $\|\cdot\|$ as follows: $\|\mu\| := |\mu|(S^n)$. Then we can show that $(\mathcal{B}, \mathcal{P}, \|\cdot\|)$ is a Banach lattice. B' is the Banach space of linear operators T on $\mathcal{B}$ with the uniform norm $\|T\| := sup_{\mu \in S}\|T(\mu\|)$ where $S := \{\mu : \|\mu\| = 1\}$. Let $\mathcal{P}'$ be the set of linear operators that preserve $\mathcal{P}$, S' be the set of linear operators on $\mathcal{B}$ that preserve S. It is easy to see that S' is the closed unit ball of $\mathcal{B}'$. Since every program will map a probability distribution to a subprobability distribution, the linear operators described by programs will preserve both

S and $\mathcal{P}$, and then are in $\mathcal{S}' \cap \mathcal{P}'$.

In the following, we will use $p, q, \cdots$ for programs and $T, U, V, \cdots$ for their interpretations as linear operators. The following are the semantics for basic probabilistic programs and program constructs:

(1) **simple assignment** Let p be the basic program: $x_i := f_i(x_1, \cdots, x_n)$ where f_i is a measurable function from Ω^n to Ω. Define $F(x_1, \cdots, x_n) = (x_1, \cdots, x_{i-1}, f_i(x_1, \cdots, x_n), x_{i+1}, \cdots, x_n)$. Then the linear operator T for the program p is: $T(\mu) = \mu \cdot F^{-1}$.

(2) **random assignment** Let p be the random assignment program $x_i := \mathit{random}$ and ρ is the distribution of a random generator. Then the linear operator T for p is uniquely determined by the following definition: for any $B_i \in \mathcal{A}_i$, $T(\mu)(B_1 \times \cdots \times B_n) = \mu(B_1 \times B_{i-1} \times S \times B_{i+1} \times \cdots \times B_n) \cdot \rho(B_i)$.

(3) **sequential composition** Let p_1, p_2 be two programs and T_1, T_2 be their semantics as linear operators, respectively. Then the linear operator for $p_1; p_2$ is $T_1; T_2$.

(4) **conditional** Let p be the program $B; p_1 + \neg B; p_2$ and T_1, T_2 be the interpretations of p_1, p_2 as linear operators. Then the linear operator T for p is $T_1 \cdot e_B + T_2 \cdot e_{\neg B}$ where $e_B(\mu) = \mu_B$ and $e_{\neg B}(\mu) = \mu_{\neg B}$. This interpretation is justified by the following equality:

$$(T_1 \cdot e_B + T_2 \cdot e_{\neg B})(A) = \mu(B) T_1(\mu_B / \mu(B))(A) + \mu(\neg B) T_1(\mu_{\neg B} / \mu(\neg B))(A)$$

(5) **loops** Let p be the program while B do p' and T' be its interpretation as linear operator. Then the linear operator T for p is the unique solution to the following affine equation scheme of T:

$$T = e_{\neg B} + T \circ T' \circ e_B.$$

This is justified by the equivalence between the programs p and if B then p' else; p.

By summarizing the above results, we can show the following main theorem:

THEOREM 7.1.2. *Let p be any while program over the program variables $x_1, \cdots, x_n$ and let $\mathcal{B}$ be the Banach lattice of measures on $(\Omega^n, \mathcal{A}^n)$. Then p can be interpreted as a linear operator in $\mathcal{S}' \cap \mathcal{P}'$.*

7.1.2.3. *Backward Semantics.* In this section, we provide another semantics for probabilistic programs, which is most analogous to the binary relations in PDL. Each program is interpreted as a Markov transition on a measurable space $(S, \mathcal{A})$. A *Markov transition* p is a function from $\Omega \times \mathcal{A}$ to $[0, 1]$ such that

- for each $s \in S$, $p(s, \cdot)$ is a probability measure on $\mathcal{A}$;
- for each $A \in \mathcal{A}$, $p(\cdot, A)$ is a measurable function.

Now we provide a semantics for all probabilistic programs:

(1) **basic programs** are interpreted as Markov transitions;
(2) **sequential composition** $(p; q)(s, A) = \int_{t \in S} q(t, A)p(s, dt)$;
(3) **conditional** (if B then p else q)(s, A) $= (B?; p + \neg B?; q)(s, A)$ where $(B?)(s, A) = \chi_{A \cap B}(s)$;
(4) **loops** (while B do $p)(s, A) = (\neg B? + (B?; p)^*)(s, A)$ where $p^* = \sum_{i=0}^{\infty} p^i$.

Each Markov transition can be extended to a unique predicate transformer which is a linear transformation on the set of bounded measurable functions on $(\Omega, \mathcal{A})$. For any bounded measurable function f, $(\langle p \rangle f)(s) = \int f(t)p(s, dt)$.

7.1.2.4. *Duality.* Actually the above two semantics for probabilistic programs are equivalent in the following sense. Each point s can be regarded as a point mass. And since discrete measures are dense in the set of probability measures, each Markov transition can be extended to a unique measure-transformer which maps $\mathcal{M}$, the set of probability measures on $(\Omega, \mathcal{A})$, to $\mathcal{M}$. In other words, $(\mu\langle p \rangle)(A) = \int p(s, A)\mu(ds)$.

THEOREM 7.1.3. $(\mu\langle p \rangle, A) = (\mu, \langle p \rangle A)$*; hence* $(\mu\langle p \rangle, f) = (\mu, \langle p \rangle f)$.

PROOF. The proof is actually an application of the Fubili's theorem.

$$\begin{aligned}
(\mu\langle p\rangle, A) &= \int p(s, A)\mu(ds) \\
&= \int\int_A p(s,t)dt\mu(ds) \\
&= \int_A\int p(s,t)\mu(ds)dt \\
&= (\mu, \langle p\rangle A)
\end{aligned}$$

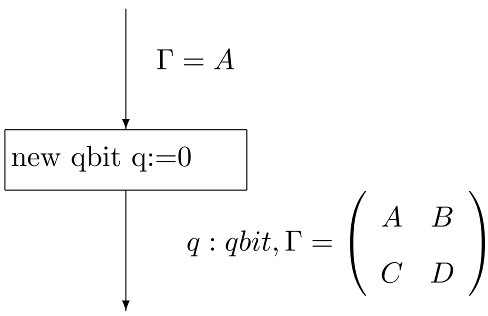
Since simple functions are dense in the set of bounded measurable functions, the second part follows immediately.

□

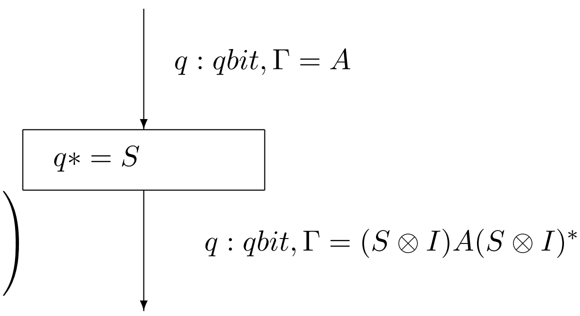
7.2. Quantum Flow Charts

In this section, we present Peter Selinger's quantum programming languages. The slogan for this language is "quantum data and classical control". Its syntax is in the form of quantum flow charts. We attach a pair to each edge in the following charts. The first coordinate is its label, which is a typing context, and the second is its annotation, which is a density matrix. The following are 6 basic flow charts.

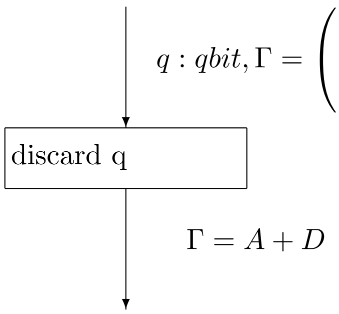
Allocate qbit



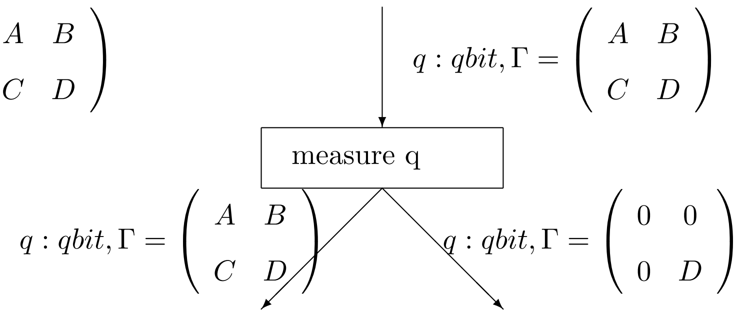
Unitary Transformation



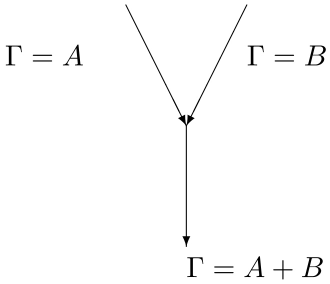
Discard qbit



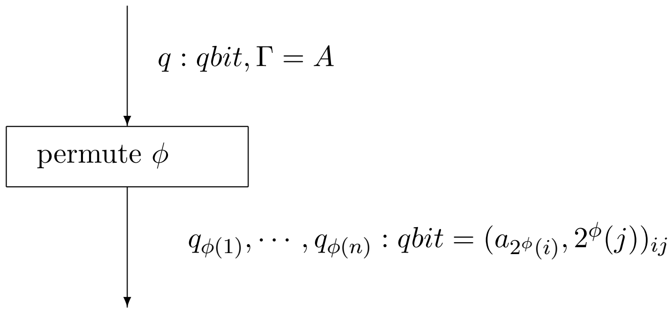
Measurement



Merge



Permutation

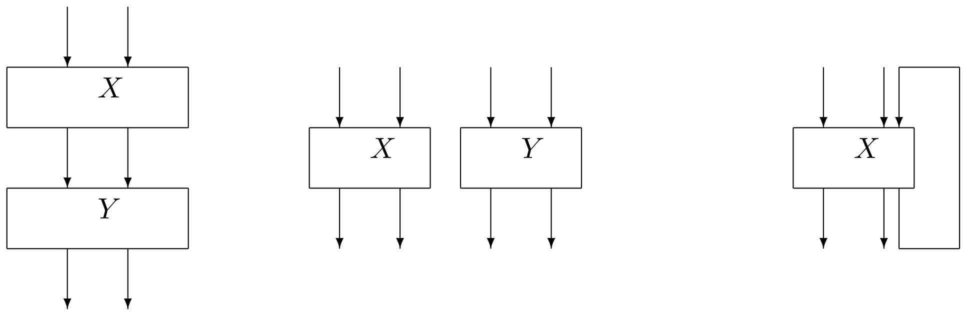


There are four program constructs.

Context Extension

$$\boxed{\quad X \quad} \otimes \boxed{\Gamma} = \boxed{\quad Y \quad}$$

Vertical Composition **Horizontal Composition** **Loops**



7.3. Semantics for Quantum Programs

7.3.1. Forward Semantics. In this section, we present a systematic and formal treatment of the semantics of quantum flow charts.

DEFINITION 7.3.1. A *signature* σ is a list of non-zero natural numbers $(n_1, \cdots, n_s)$. The vector space V_σ with the signature σ is $V_\sigma = C^{n_1 \times n_1} \times \cdots \times C^{n_s \times n_s}$. So every element of V_σ is a list of matrixes $(A_1, \cdots, A_s)$ where A_i is an $n_i \times n_i$ matrix. And $tr(A) = \sum_i tr(A_i)$. Define $D_\sigma = \{A \in V_\sigma | A \text{ is positve and } tr(A) \leq 1\}$.

DEFINITION 7.3.2. Let F be a linear operator from V_σ to $V_{\sigma'}$. F is positive if F maps positive matrixes to positive ones. F is *completely positive* if $id_\tau \otimes F : V_{\tau\otimes\sigma} \to V_{\tau\otimes\sigma'}$ is positive for any signature τ. F is a *superoperator* if it is completely positive and satisfies the following condition: for any $A \in V_\sigma, tr(F(A)) \leq tr(A)$.

THEOREM 7.3.3. *(Kraus representation theorem) Assume that F is a completely positive linear operator from $C^{n\times n}$ to $C^{m\times m}$. The following two statements are equivalent:*

- *F is a superoperator;*
- *$F(A) = \sum_i U_i A U_i^*$ for some matrices $U_i's$ such that $\sum_i U_i U_i^* \sqsubseteq Id$.*

Note that we can interpret any typing context Γ as a vector space $V^{2^{n_1}\times 2^{n_1}} \times \cdots V^{2^{n_s}\times 2^{n_s}}$ for some $n_1, \cdots, n_s$ where s depends on the classical control and $n_1, \cdots, n_s$ depend on quantum data. And we denote this meaning as $[[\Gamma]]$. So any quantum program S is interpreted as a linear operator, which is denoted as $[[S]]$.

THEOREM 7.3.4. *Under this semantics, for any quantum program S, $[[S]]$ is a superoperator.*

PROOF. According to the definition of quantum programs, any quantum program is constructed from basic quantum flow charts by context extension, sequential composition, horizontal composition and loops. Here we just take the basic quantum flow chart: unitary transformation and the context extension as an illustration of the proof.

We know that the unitary transformation S maps density matrix A to $(U \otimes I)A(U \otimes I)^*$ where U is a unitary matrix. Since $U \otimes I$ is still a unitary, S is completely positive. Moreover, by Kraus representation theorem, we know that it is also a superoperator because $(U \otimes I)(U \otimes I)^* = I$.

It is easy to see that context extension means a linear mapping $I \otimes Z$ where Z is a zero mapping. $I \otimes Z$ is a completely positive operator and also a superoperator.

□

More importantly, the converse to this proposition also holds. Here we give a proof sketch.

DEFINITION 7.3.5. A matrix U is *subunitary* if it is a submatrix of some unitary matrix in the following sense that there are U_1, U_2, U_3 such that

$$\left(\begin{array}{c|c} U & U_1 \\ \hline U_2 & U_3 \end{array}\right)$$

is a unitary. Let $\sigma = (N_1, \cdots, n_s)$ and $\sigma' = n_1 + \cdots + n_s$. S is a *measurement operator* if

$$S\begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & \ddots & \vdots \\ A_{s1} & \cdots & A_{ss} \end{pmatrix} = (A_{11}, A_{22}, \cdots, A_{ss})$$

S is an *eraser operator* or a *partial trace operator* if $S = tr_\sigma \otimes id_\tau : V_{\sigma\otimes\tau} \to V_\tau$.

THEOREM 7.3.6. *(Normal form theorem) Every superoperator F can be factored as $F = M \circ E \circ S$ where M is a measurement operator, E is an eraser operator and S is subunitary.*

THEOREM 7.3.7. *(Fullness of interpretation) Given typing contexts* Γ, Γ', *if* $F : [[\Gamma]] \to [[\Gamma']]$ *is a superoperator, then there is a quantum program* T *such that* $[[T]] = F$.

PROOF. Here we give a proof sketch through the above normal form theorem. Assume that $F = M \circ E \circ S$. First note that any subunitary S can be realized by allocating new qbits, measuring it and then discarding qbits (intuitively this is clear). Moreover, E is definable by discarding and merging, and M can be realized by several measurements. So, indeed any superoperator can be realized by quantum flow charts. $\square$

7.3.2. Backward Semantics. We can also give a backward semantics for quantum programs as predicate transformers. Its connection with forward semantics is provided through quantum weakest preconditions.

DEFINITION 7.3.8. A *quantum predicate* is a positive operator with eigenvalues bounded by 1. A *healthy predicate transformer* α is a linear mapping from $\mathcal{P}(\mathcal{H})$ to $\mathcal{P}(\mathcal{H}')$ that is linear and completely positive. We denote the set of healthy predicate transformers by $\mathcal{PT}(\mathcal{H}, \mathcal{H}')$.

The reason is that the expectation $tr(M\rho)$ of one predicate M at one state ρ should be less than 1. Let $\mathcal{P}(\mathcal{H})$ denote the set of all predicates on the Hilbert space $\mathcal{H}$, and $\mathcal{DM}(\mathcal{H})$ the set of density matrices on $\mathcal{H}$. From the last subsection, we already know that the interpretation of quantum programs as superoperators is full. So, in this subsection, we don't distinguish quantum programs and superoperators. Define a partial order on predicates $M \sqsubseteq N$ iff $N - M$ is positive.

DEFINITION 7.3.9. The predicate M is called a *precondition* for the predicate N with respect to the quantum program T if

$$\text{for any } \rho \in \mathcal{DM}(\mathcal{H}), (tr(M))(\rho) \leq tr(N(T(\rho))).$$

It is denoted by $M\{T\}N$. M is a *weakest precondition* for N with respect to T if, for any M' such that $M'\{T\}N$, $M' \sqsubseteq M$. This weakest precondition is denoted by $wp(T)(N)$.

From the following proposition, we can see that $wp(T)$ induces a predicate transformer.

THEOREM 7.3.10. *For any quantum program T and any predicate $M \in \mathcal{P}(\mathcal{H})$, $wp(T)(M) \in \mathcal{P}(\mathcal{H})$ and is unique. Moreover,*

$$\text{for any } \rho \in \mathcal{DM}(\mathcal{H}), tr(M(T(\rho))) = tr(wp(T)(M)).$$

Hence $wp(T)$ is a healthy predicate transformer.

PROOF. Since T is a superoperator, $T(A) = \sum_i U_i A U_i^*$ for some U_i such that $\sum_i U_i U_i^* \sqsubseteq Id$. Define:

$$N := \sum_i U_i M U_i^*.$$

It is easy to check that $N \in \mathcal{P}(\mathcal{H})$ and such defined N is a weakest precondition for M with respect to T.

□

We give meaning to measurements only as an illustration of the backward semantics for quantum programs. Consider the following measurement:

$$[[\text{measure q}]]\text{: } qbit \to qbit \oplus qbit : \rho \to P_0 \rho P_0 + P_1 \rho P_1.$$

where P_0 and P_1 are the corresponding projections to the measurement with results 0 and 1, respectively. Dually, the backward semantics for the quantum program measure q should be:

$$\text{for any predicate } M, \text{ wp(measure q)}(M) = P_0 M P_0 + P_1 M P_1.$$

Actually the backward semantics for quantum programs is dual to its forward semantics in the following sense. Like Kozen's notation for probabilistic semantics, $(\rho, \langle T \rangle M)$ denote $tr((wp(T)(M))(\rho))$ and $(M, \rho\langle T \rangle)$ denote $tr(M(T(\rho)))$. By the above theorem, we know that:

$$(\rho, \langle T \rangle M) = (M, \rho\langle T \rangle)$$

This is to say, these two semantics for quantum programs are equivalent to each other.

7.4. Constructing Unitary Matrix

In this section, we want to explore the possibility of constructing unitary matrices from arbitrary matrices with complex entries. Before that, we review the construction of reversible matrices from arbitrary matrices with entries from Z_2, where $(Z_2, \cdot, \oplus)$ is the finite field with only two elements. Given any function $f : Z_2^n \to Z_2^m$, can we find a function f^R that is a bijection from $Z_2^{n+m} \to Z_2^{n+m}$ and subsumes f as a subfunction in the following sense:

$$f^R(x, \overrightarrow{0}) = (x, f(x)) \text{ for all } x \in Z_2^n?$$

Define $f^R(x, z) = (x, f(x) \oplus z)$ for $x \in Z_2^n$ and $z \in Z_2^m$ where $\oplus$ is the addition defined on Z_2^m.

CLAIM 7.4.1. *f is a subfunction of f^R.*

Note that $f^R(x, \overrightarrow{0}) = (x, f(x) \oplus \overrightarrow{0}) = (x, f(x))$. So, indeed f^R subsumes f as its subfunction.

CLAIM 7.4.2. *f^R is a bijection from $Z_2^{n+m} \to Z_2^{n+m}$.*

It suffices to show that f^R is one-to-one. Assume that $(x_1, z_1) \neq (x_2, z_2)$. If $x_1 \neq x_2$, then $f^R(x_1, z_1) = (x_1, f(x_1) \oplus z_1) \neq (x_2, f(x_2) \oplus z_2) = f(x_2, z_2)$. If $x_1 = x_2$ and $z_1 \neq z_2$, then $f^R(x_1, z_1) \neq f^R(x_2, z_2)$ because $z_1 \oplus f(x_1) \neq z_2 \oplus f(x_2)$.

Next we want to explore the extent to which the above result can be generalized to the quantum case. Here we consider a simple case. Let f be a linear function from $\mathcal{C}^2$ to $\mathcal{C}^{2^2}$. We can also use a matrix to represent this transformation:

$$f(|0\rangle, |1\rangle) = (|00\rangle, |01\rangle, |10\rangle, |11\rangle) \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{pmatrix}$$

Define

$$B = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{pmatrix}$$

By abuse of notation, we define $\oplus : |ij\rangle \oplus \imath' j'\rangle = |i \oplus_2 i', j \oplus_2 j'\rangle$ where $\oplus_2$ is the addition in Z_2 and then extend it linearly to the vector space $\mathcal{C}^{2^2}$. Define:

$$f^R(x, z) = (x, z \oplus f(x)) \text{ for } x \in \mathcal{C}^2, z \in \mathcal{C}^{2^2}.$$

If we use the matrix representation, the function f^R can be written as:

$$f(|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle) = (|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle)A$$

where

$$A = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & 0 & 0 & 0 & 0 \\ x_2 & x_1 & x_4 & x_3 & 0 & 0 & 0 & 0 \\ x_3 & x_4 & x_1 & x_2 & 0 & 0 & 0 & 0 \\ x_4 & x_3 & x_2 & x_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y_1 & y_2 & y_3 & y_4 \\ 0 & 0 & 0 & 0 & y_2 & y_1 & y_4 & y_3 \\ 0 & 0 & 0 & 0 & y_3 & y_4 & y_1 & y_2 \\ 0 & 0 & 0 & 0 & y_4 & y_3 & y_2 & y_1 \end{pmatrix}$$

Denote

$$A_1 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_4 & x_3 \\ x_3 & x_4 & x_1 & x_2 \\ x_4 & x_3 & x_2 & x_1 \end{pmatrix}, A_2 = \begin{pmatrix} y_1 & y_2 & y_3 & y_4 \\ y_2 & y_1 & y_4 & y_3 \\ y_3 & y_4 & y_1 & y_2 \\ y_4 & y_3 & y_2 & y_1 \end{pmatrix}$$

Note that A is unitary iff both A_1 and A_2 are unitary.

THEOREM 7.4.3. *Assume that all entries in A_1 are real. A_1 is unitary iff (x_1, x_2, x_3, x_4) is*

- *either a permutation of 0, 0, 0 and 1;*

- *or a permutation of 0, 0, 0 and -1;*
- *or a permutation of 1/2, 1/2, 1/2 and -1/2;*
- *or a permutation of -1/2, -1/2, -1/2 and 1/2.*

PROOF. Assume that all entries in A_1 are real and A_1 is unitary. Then we can get the following group of equations:

$$\begin{cases} x_1^2 + x_2^2 + x_3^2 + x_4^2 & = & 1 & (1) \\ x_1x_2 + x_3x_4 & = & 0 & (2) \\ x_1x_3 + x_2x_4 & = & 0 & (3) \\ x_1x_4 + x_2x_3 & = & 0 & (4) \end{cases}$$

It follows immediately that

$$\begin{cases} (x_1 + x_2)(x_3 + x_4) & = & 0 & (5) \\ (x_1 + x_3)(x_2 + x_4) & = & 0 & (6) \\ (x_1 + x_4)(x_2 + x_3) & = & 0 & (7) \end{cases}$$

First we assume that at least one of x_i's is zero. Since x_1, x_2, x_3 and x_4 are symmetric, we assume that $x_1 = 0$ without loss of generality. From (2), we know that at least one of x_3 and x_4 is zero. Without loss of generality, assume that $x_3 = 0$. From (3), we conclude that at least one of x_2 and x_4 is zero. But since $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$, at least one of x_2 and x_4 is ± 1. In other words, if at least one of x_i's is zero, then (x_1, x_2, x_3, x_4) is a permutation of either $(1, 0, 0, 0)$ or of $(-1, 0, 0, 0)$.

Next assume that none of them is zero. From (5), we know that $x_1 = -x_2$ or $x_3 = -x_4$. Without loss of generality, we assume that $x_1 = -x_2$. From (6), we know that $x_1 = -x_3$ or $x_2 = -x_4$. Since x_1 is symmetric to x_2, without loss of generality, $x_1 = -x_3$. It follows that $x_2 = x_3$. Since we assume that none of them is zero, $x_2 + x_3 \neq 0$. This implies that $x_1 = -x_4$ or $x_4 = x_2 = x_3$. Since $x_1^2+x_2^2+x_3^2+x_4^2 = 1$, $x_1 = \pm 1/2$. In other words, if none of x_1, x_2, x_3 and x_4 is zero, then (x_1, x_2, x_3, x_4) is a permutation of either $(1/2, 1/2, 1/2, -1/2)$ or of $(-1/2, -1/2, -1/2, 1/2)$. This also finishes the proof of the left-to-right direction of the theorem.

The proof of the other direction is straightforward.

□

THEOREM 7.4.4. *Assume that all entries in A_1 are complex. If A_1 is unitary, then*

- $|x_1|^2 + |x_2|^2 + |x_3|^2 + |x_4|^2 = 1$;
- $|x_1 + x_2 + x_3 + x_4| = 1$

PROOF. Assume that A_1 is unitary. In other words,

$$\begin{cases} |x_1|^2 + |x_2|^2 + |x_3|^2 + |x_4|^2 & = & 1 \\ (x_1 x_2^* + x_1^* x_2) + (x_3 x_4^* + x_3^* x_4) & = & 0 \\ (x_1 x_3^* + x_1^* x_3) + (x_2 x_4^* + x_2^* x_4) & = & 0 \\ (x_1 x_4^* + x_1^* x_4) + (x_3 x_2^* + x_2^* x_4) & = & 0 \end{cases}$$

If we add all these equalities, we have:

$$(x_1 + x_2 + x_3 + x_4)(x_1 + x_2 + x_3 + x_4)^* = 1.$$

That is to say, $|x_1 + x_2 + x_3 + x_4| = 1$.

□

But, if we don't require that the extended function has the form of f^R, then almost all f's have a unitary extension.

THEOREM 7.4.5. *For the above $f : \mathcal{C}^2 \to \mathcal{C}^{2^2}$, if both $x_1x_2x_3x_4 \neq 0$ and $y_1y_2y_3y_4 \neq 0$, then there is a linear function f_R such that*

- $f_R(x \otimes |00\rangle) = (x \otimes f(x))$ *for all* $x \in \mathcal{C}^2$;
- *the matrix associated with f_R is unitary.*

PROOF. Consider the column vector $(x_1, x_2, x_3, x_4)^t$. Since the vector space $V_4 := \{(a_1, a_2, a_3, a_4)^t | a_i \in \mathcal{C}(1 \leq i \leq 4)\}$ is of dimension 4 and $v_0 := (x_1, x_2, x_3, x_4) \neq 0$, there are three vectors:

$$v_1 = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{pmatrix}, v_2 = \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \end{pmatrix}, v_3 = \begin{pmatrix} a_{31} \\ a_{32} \\ a_{33} \\ a_{34} \end{pmatrix}$$

such that v_0, v_1, v_2 and v_3 are linearly independent. By using Gram-Schmit normalization method, we *construct* from v_1, v_2 and v_3 the following three unit vectors

$$v'_1 = \begin{pmatrix} a'_{11} \\ a'_{12} \\ a'_{13} \\ a'_{14} \end{pmatrix}, v'_2 = \begin{pmatrix} a'_{21} \\ a'_{22} \\ a'_{23} \\ a'_{24} \end{pmatrix}, v'_3 = \begin{pmatrix} a'_{31} \\ a'_{32} \\ a'_{33} \\ a'_{34} \end{pmatrix}$$

such that

$$A'_1 = \begin{pmatrix} x_1 & a'_{11} & a'_{21} & a'_{31} \\ x_2 & a'_{12} & a'_{22} & a'_{32} \\ x_3 & a'_{13} & a'_{23} & a'_{33} \\ x_4 & a'_{14} & a'_{24} & a'_{34} \end{pmatrix}$$

is unitary. Similarly we can find three unit vectors

$$w'_1 = \begin{pmatrix} b'_{11} \\ b'_{12} \\ b'_{13} \\ b'_{14} \end{pmatrix}, w'_2 = \begin{pmatrix} b'_{21} \\ b'_{22} \\ b'_{23} \\ b'_{24} \end{pmatrix}, w'_3 = \begin{pmatrix} b'_{31} \\ b'_{32} \\ b'_{33} \\ b'_{34} \end{pmatrix}$$

such that

$$B'_1 = \begin{pmatrix} y_1 & b'_{11} & b'_{21} & b'_{31} \\ y_2 & b'_{12} & b'_{22} & b'_{32} \\ y_3 & b'_{13} & b'_{23} & b'_{33} \\ y_4 & b'_{14} & b'_{24} & b'_{34} \end{pmatrix}$$

is unitary. Denote

$$B' = \begin{pmatrix} x_1 & a'_{11} & a'_{21} & a'_{31} & 0 & 0 & 0 & 0 \\ x_2 & a'_{12} & a'_{22} & a'_{32} & 0 & 0 & 0 & 0 \\ x_3 & a'_{13} & a'_{23} & a'_{33} & 0 & 0 & 0 & 0 \\ x_4 & a'_{14} & a'_{24} & a'_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y_1 & b'_{11} & b'_{21} & b'_{31} \\ 0 & 0 & 0 & 0 & y_2 & b'_{12} & b'_{22} & b'_{32} \\ 0 & 0 & 0 & 0 & y_3 & b'_{13} & b'_{23} & b'_{33} \\ 0 & 0 & 0 & 0 & y_4 & b'_{14} & b'_{24} & b'_{34} \end{pmatrix}$$

It is easy to see that B' is unitary. Define f_R to be the linear mapping: $\mathcal{C}^{2^3} \to \mathcal{C}^{2^3}$ whose associated matrix is B'.

CLAIM 7.4.6. *$f_R(x \otimes |00\rangle) = (x \otimes f(x))$ for all $x \in \mathcal{C}^2$.*

Note that $f(|000\rangle) = x_1|000\rangle + x_2|001\rangle + x_3|010\rangle + x_4|011\rangle = |0\rangle \otimes f(|0\rangle)$ and $f(|100\rangle) = y_1|100\rangle + y_2|101\rangle + y_3|110\rangle + y_4|111\rangle = |1\rangle \otimes f(|1\rangle)$.

$\square$

APPENDIX A

Future Research

The perspective of my dissertation is mainly *logical*; that is, I focused on such important issues in traditional modal logic as (weak and strong) completeness, finite model property and decidability. In contrast, the perspective on probability logic of theoretical computer scientists is mainly *coalgebraic*. After finishing the above foundational work, my future research will combine these two perspectives and further develop the model theory of probability *logics* with the approach of *algebras* and *co-algebras* [36].

A measurable function from the measurable space (M, Σ) to the measurable space (M', Σ') is a *morphism* of these spaces. This gives rise a category which is often called **Meas**. **Meas** has products, coproducts and the following important endofunctor $\Delta :$ **Meas** $\rightarrow$ **Meas**. It is a morphism of measurable spaces $\Delta : (M, \Sigma) \rightarrow (\Delta(M), \Sigma_M)$ where $\Delta(M)$ is the set of probability measures on (M, Σ) and Σ_M is a σ-algebra generated by the sets $\{\mu \in \Delta(M) : \mu(E) \geq p\}$ for all $E \in \Sigma$ and rational p between 0 and 1. The functors that we are interested in are called *measurable polynomial functors*. They are the functions constructed from the identity functor Id and the constant functors, and are closed under products, coproducts and Δ. Given measurable spaces M_1 and M_2, $F : X \rightarrow M_1 \times \Delta(M_2, X)$ is a measurable polynomial functor. A *coalgebra* of a measurable polynomial function T is a pair (A, f) consisting of a measurable space A and a measurable map $f : A \rightarrow T(A)$.

My first task is to provide a general theory of probabilistic bisimulation, the most discussed probabilistic process equivalence. Up to now, the most general characterization of probabilistic bisimulations is [8] and [13]. The main theorem there is: given a cospan of stochastic systems over analytic spaces with Borel measurable transition functions, there exists a span based on a stochastic system over a Polish space with a Borel measurable transition function.

- (Project 1) I would like to investigate weak morphisms, weak logical equivalence and bisimilarities along the paths indicated in [11]. More generally, I want to explore the coalgebraic logical languages like $\mathcal{L}(T)$ presented in [29] to characterize bisimulations for all measure polynomial functors T.

The idea that the functor of a coalgebra determines a certain modal logic was first put forward by Moss [28]. [21] developed a many-sorted modal logical system that captures the natural notion of validity for the restricted, inductively defined class of (Kripke) polynomial functors. Similarly, [29] generated a rather expressive coalgebraic language, of which syntax and semantics are directly and uniformly derived from measurable polynomial functors.

- (Project 2) My ongoing work with Professor Moss is to use modular construction in [7] to construct complete deductive systems for the languages $\mathcal{L}(T)$ [29] based on the system Σ_+ [38]. These logics thus obtained should inherit soundness, completeness and expressiveness properties from their building blocks.

The coalgebraic investigation of probabilistic dynamic logic [23] with the above approach would be an interesting object of study. This research is a further application of Venema's coalgebra automata for fixed point logic [35].

- (Project 3) I will examine *simulation* and *safety* for probabilistic bisimulation: what operations on probability models preserve bisimulation? I am especially interested in some operation which is the counterpart of ultrafilter extension in modal logic [6]. Its correlation with duality for labeled Markov processes [27] is expected. Such a question is well answered in propositional dynamic logic [34].

Just as PDL is a special instance of modal μ-calculus, probabilistic dynamic logic is included in a much richer area of probability logic with mu-operators.

- (Project 4) My most ambitious goal is to generalize the coalgebraic perspective of automata and fixed point logic by Venema [35] to the probabilistic setting, for example, probability logic with common belief [14] and stochastic logic with fixed point operator [12].

The last project is isolated from the above four. It came from the logical investigation in my dissertation of the transition from knowledge to belief.

- (Project 5) I would like to explore the connection between ergodic theory and higher order probabilities. This connection is needed in the study of the correspondence between semantics and syntax of probabilistic logics. In [31], Samet used the ergodic nature of type functions to show the equivalence of different axioms. I hope that the research in this area will solve the question (positively or negatively) of completeness of higher-order probability logics.

Bibliography

[1] C. Aliprantis and K. Border. *Infinite Dimension Analysis.* Springer, third edition, 2006. A Hitchhiker's Guide.

[2] R. Aumann. Agreeing to disagree. *Ann. Statist*, 4:1236–1239, 1976.

[3] R. Aumann. Interactive epistemology: Knowledge. *Int. J. Game Theory*, 28:263–300, 1999.

[4] R. Aumann. Interactive epistemology: Probability. *Int. J. Game Theory*, 28:301–314, 1999.

[5] A. Baltag, P. Van Ditmarsch, and L. Moss. Epsitemic logic and information update. *Handbook of Philosophy of Information*, to appear.

[6] M. Blackburn P., de Rijke and Y. Venema. *Modal Logic.* Cambridge Tracts in Theoretical Computer Science, 53. Oxford University Press, 2000.

[7] C. Cîrstea and D. Pattinson. Modular construction of modal logics. In *CONCUR*, pages 258–275, 2004.

[8] J Desharnais, A. Edalat, and Panangaden P. Bisimulation for labeled markov processes. *Inf. and Compt.*, 179:163–193, 2002.

[9] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labeled markov processes. *Inf. and Comp.*, 184:160–200, 2003.

[10] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labeled markov processes. *Theoretical Computer Science*, 318:323–353, 2004.

[11] E. E Doberkat. Weak logical equivalence for markov transition systems. draft available at http://ls10-www.cs.uni-dortmund.de.

[12] E. E Doberkat. The hennessymilner equivalence for continuous time stochastic logic with mu-operator. *J. Applied Logic*, 5(3):519–544, 2006.

[13] E. E Doberkat. Stochastic relations: congruence, bisimulations and the hennessy-milner theorem. *SIAM J. Computing*, 35(3):590–626, 2006.

[14] R Fagin and J. Halpern. Reasoning about knowledge and probability. *J. ACM*, 41:340–367, 1994.

[15] J. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54:319–379, 1992.

[16] J. Harsanyi. Games with incomplete information played by bayesian players, part one. *Management Sci.*, 14:159–182, 1967.

[17] A. Heifetz and P. Mongin. Probability logic for type spaces. *Games Econom. Behav.*, 35:31–53, 2001.

[18] A. Heifetz and D. Samet. Topology-free typology of beliefs. *J. Econom. Theory*, 82:324–241, 1998.

[19] A. Heifetz and D. Samet. Coherent beliefs are not always types. *J. Math. Econom.*, 32:475–488, 1999.

[20] J. Hintikka. *Knowledge and Belief.* Cornell University Press, Ithaca, NY, 1962.

[21] B. Jacobs. Many-sorted coalgebraic modal logic: a model-theoretic study. *ITA*, 35(1):31–59, 2001.

[22] D. Kozen. Semantics of probabilistic programs. *J. Computer. System Sci.*, 22:328–350, 1981.

[23] D. Kozen. A probabilistic pdl. *J. Computer. System Sci.*, 30:162–178, 1985.

[24] S. Kripke. A completeness theorem in modal logic. *J. Symbolic Logic*, 24:1–14, 1959.

[25] K. G. Larsen and Skou A. Bisimulation through probabilistic testing. *Inf. and Comp.*, 94:1–28, 1991.

[26] J. Mertens and Zamir S. Formulation of bayesian analysis for games with incomplete information. *Int. J. Game Theory*, 14:1–29, 1985.

[27] M. Mislove, J. Ouaknine, D. Pavlovic, and J. Worrell. Duality for labelled markov processes. In *FoSSaCS*, pages 393–407, 2004.

[28] L. Moss. Coalgebraic logic. *Ann. Pure Appl. Logic*, 96(1-3):277–317, 1999.

[29] L. Moss and I. Viglizzo. Final coalgebras for functors on measurable spaces. *Inf. and Comp.*, 204:610–636, 2006.

[30] M. Nielsen and I. Chung. *Quantum Information and Quantum Computation.* Cambridge University Press, 2000.

[31] D. Samet. Quantified beliefs and believed quantities. *J. Economic Theory*, 95:169–185, 2000.

[32] A. Schrijver. *Theory of Linear and Integer Progamming.* Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, 1986.

[33] P. Selinger. Towards a quantum programming language. *Math. Structures Comput. Sci.*, 14:527–586, 2004.

[34] J. van Benthem. *Exploring Logical Dynamics.* Studies in Logic, Language and Information. CSLI Publications, 1996.

[35] Y. Venema. Automata and fixed point logic: A coalgebraic perspective. *Inf. Comput.*, 204(4):637–678, 2006.

[36] Y. Venema. Algebras and co-algebras. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, Studies in Logic and Practical Reasoning, pages 331–426. 2007.

[37] I. Viglizzo. *Coalgebras on Measurable Spaces.* PhD thesis, Indiana University, Bloomington, 2005.

[38] C Zhou. *Complete Deductive Systems for Probability Logic with Application in Harsanyi Type spaces.* PhD thesis, Indiana University, Bloomington, 2007.

[39] W. Ziemer. Modern real analysis. Online Lecture Notes.