SRI International

THE LOGICAL FOUNDATIONS OF EVIDENTIAL REASONING

Technical Note 408

December 20, 1986
Revised April 27, 1987

By: Enrique H. Ruspini
Artificial Intelligence Center
Computer Science and Technology Division

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

This work was supported in part by the United States Army Signal Warfare Center under contract DAAL02-85-C-0082, and in part by the National Science Foundation under contract DCR-85-13139.

SRI International®

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

Abstract

The approach proposed by Carnap for the development of logical bases for probability theory is investigated by using formal structures that are based on epistemic logics. Epistemic logics are modal logics introduced to deal with issues that are relevant to the state of knowledge that rational agents have about the real world. The use of epistemic logics in problems of analysis of evidence is justified by the need to distinguish among such notions as the state of a real system, the state of knowledge possessed by rational agents, and the impact of information on that knowledge.

Carnap's method for generating a universe of possible worlds is followed using an enhanced notion of possible world that encompasses descriptions of knowledge states. Within such generalized or *epistemic* universes, several classes of sets are identified in terms of the truth-values of propositions that describe either the state of the world or the state of knowledge about it. These classes of subsets have the structure of a sigma algebra.

Probabilities defined over one of these sigma algebras, called the *epistemic algebra*, are then shown to have the properties of the belief and basic probability assignment functions of the Dempster-Shafer calculus of evidence.

It is also shown that any extensions of a probability function defined on the epistemic algebra (representing different states of knowledge) to the *truth algebra* (representing true states of the real world) must satisfy the interval probability bounds derived from the Dempster-Shafer theory. These bounds are also shown to correspond to the classical notions of lower and upper probability. Furthermore, these constraints are shown to be the best possible bounds, given a specific state of knowledge.

Finally, the problem of combining the knowledge that several agents have about a real-world system is addressed. Structures representing possible results of the integration of that knowledge are introduced and a general formula for the combination of evidence is derived. From this formula and certain probabilistic independence assumptions, a generalization of the rule of combination of Dempster is readily proved. The meaning of these independence assumptions is made explicit through the insight provided by the formal structures that are used to represent knowledge and truth.

Finally, simple cases of combination of dependent evidence are discussed as an introduction to more general problems of general combination that are examined in a related paper.

ACKNOWLEDGMENTS

The research work described in this paper was sponsored by the U.S. Army Signal Warfare Center under Contract DAAL02-85-C-0082 and by the National Science Foundation under Contract DCR-85-13139.

The author benefited from comments made by, discussions held with, and advice received from numerous colleagues in the Artificial Intelligence Center. Bob Moore, Stan Rosenschein and David Israel suggested the investigation of logic-based formalisms, in general, and of the Carnapian viewpoint, in particular, as prospective avenues that might lead to a better understanding of the relationships among existing approaches to approximate reasoning. John Lowrance, Tom Garvey, Tom Strat, Len Wesley and Stephen Lesh, comprised an abundant source of insight, comments, examples, and counterexamples. I have also profited from conversations with Andy Hanson, Haim Gaifman, Pascal Fua, Leslie Kaelbling, Martin Fischler, Ray Perrault, Kurt Konolige, Ken Laws, Lorna Shinkle, Mike Georgeff, Martha Pollack, Ivan Leclerc, Todd Davies, and Oscar Firschein. To all of them many thanks.

I would also like to thank Professor Lotfi Zadeh of the University of California, Berkeley, for his constant encouragement and helpful criticism. I am particularly thankful for his support during my recent visit as a Visiting Scholar at Berkeley, where this research was begun. I am also indebted to Piero Bonissone of GE Research Labs and Terry Cronin of the Signal Warfare Center, U.S. Army, for their helpful comments and suggestions.

CONTENTS

I. INTRODUCTION . . . 1
I-1. Motivation . . . 1
I-2. Presentation Approach . . . 2
I-3. Organization . . . 3

II. THE EPISTEMIC UNIVERSE . . . 4
II-1. Epistemic Logic . . . 4
II-2. Epistemic Worlds . . . 5
II-3. Necessary Truth . . . 5
II-4. Epistemic States . . . 6
II-5. Most Specific Knowledge . . . 7
II-6. Special Subsets of the Epistemic Universe . . . 8

III. PROBABILITIES IN THE EPISTEMIC UNIVERSE . . . 9
III-1. Probability Measures . . . 9
III-2. Lower and Upper Probabilities . . . 10
III-3. Probabilities in the Epistemic Algebra . . . 14

IV. THE COMBINATION OF KNOWLEDGE . . . 19
IV-1. Multiple Epistemic Agents . . . 19
IV-2. Product Spaces and Universes . . . 19
IV-3. Epistemic Sets in the Product Universe . . . 21
IV-4. Marginal and Product Epistemic Algebras . . . 22
IV-5. Probabilities in the Product Universe . . . 23
IV-6. Probability Masses in the Product Universe . . . 26

V. THE COMBINATION OF INDEPENDENT EVIDENCE . . . 27
V-1. Probabilistic Independence . . . 27

VI. SIMPLE FORMS OF DEPENDENT EVIDENCE . . . 29
VI-1. Compatibility Between Frames of Discernment . . . 29
VI-2. The Combination of Dependent Evidence . . . 32

REFERENCES . . . 33

I
INTRODUCTION

I-1. Motivation

The purpose of this work is to develop of formal theoretical foundations that improve the understanding of the nature and usefulness of evidential reasoning concepts. These foundations have been sought both in mathematical logic, as the major conceptual approach to the study of formal reasoning systems, and in probability theory, as the foremost multivalued logic approach to the analysis of issues related to the *likelihood* of propositions as true descriptors of the state of affairs in the real world.

The approach proposed by Carnap [3] for developing logical bases of probability theory is extended here to information systems that model the state of a real system and the state of knowledge that rational agents have about it. These systems [12] are important in a number of artificial intelligence applications that are concerned with the interpretation and analysis of imprecise and uncertain evidence about the real world.

The methodology followed here is also related in several ways to the probabilistic logic approach of Nilsson [14] in its conception of probabilities as valuations over a family of subsets of a universe of possible worlds that constrain, to different degrees, the probability values over other subsets not in the family. The major differences between this work and the probabilistic logic approach are in the use of epistemic concepts and the derivation of global conditions for probability function extension (i.e., lower and upper probabilities), as opposed to local formulas derived from interval probability theory or from interpolative techniques.

The logical approach to probabilities proposed by Carnap considers the latter as valuations over a universe of *possible worlds*. These valuations allow quantification of the degree of support (called *confirmation* by Carnap) provided by certain propositions (describing evidential knowledge about the world) for the truth of other propositions describing the state of a real-world system of interest [4].

The work presented here departs from Carnap's approach by considering that evidence provides information about the truth of some propositions while failing to give any indication as to the truth of others. For example, the discovery of a lock of hair at the scene of a crime may be helpful in identifying the criminal as belonging to one or more classes of suspects (having similar hair), but may be of no use in pinpointing specific suspects. As shown by this simple example, a piece of evidence enhances the state of *knowledge* about the truth of *certain* propositions, rather than about the truth of *all* propositions. At any single step in the process of acquiring and analyzing evidence it must be recognized that, while the world may be in one of several possible states, the state of knowledge will generally not be sufficient to enable precise definition of that state.

The approach proposed here, while firmly rooted in basic probability theory, recognizes that the acquisition of evidence has the potential of changing the state of knowledge about the real world, but not of altering the actual state of the world itself. To properly model the effect of such changes in knowledge state, formalisms must be developed that make it possible to model both the state of external systems and the state of knowledge about them.

In this work, this foundation is sought in modal logics that augment standard propositional logic by considering an epistemic operator **K** that represents the state of knowledge that a rational agent has about the situation of a real-world system. These logics, originally studied by Hintikka [1], were first applied to artificial intelligence problems by Moore [2]. In epistemic logics, propositions may not only be true or false, as is the case in classical logic, but may also be known to be true or false, or they may not be known to be either true or false. Epistemic logics have been used to study the knowledge properties of artificial intelligence systems, having recently been applied to the design of intelligent robots [12].

Instead of considering a universe of *ontological* possibilities (as was done by Carnap), each describing a different state of the real world, in this approach the corresponding *epistemic universe* consists of descriptions of possible states of the real world as well as of the knowledge rational agents have about it.

Probabilities defined in this universe reflect uncertainty (derived as the result of prior experience or rational considerations) regarding the possible conclusions that may be derived from the evidence. In Carnap's approach, on the other hand, probabilities must be assigned so as to reflect the effect of evidence on any proposition, irrespective of whether or not that proposition can be related experimentally or rationally to the actual contents of such evidence.

In this work, the effect of evidence is represented as changes in the state of knowledge of rational agents. Uncertain knowledge is represented as probability functions defined over a sigma algebra of subsets of the epistemic universe. The elements of this set collection, called the *epistemic algebra*, are subsets of the epistemic universe that are characterized by common epistemic properties (i.e., the same propositions are known to be true). These probabilities are related to similar functions defined over *truth algebras*, that are subsets of possible worlds of the epistemic universe sharing the same *ontological* properties, (i.e., the same propositions about the world are true in each subset).

The major results of this work prove that this generalization of the Carnap approach validates the mathematical theory of evidence of Dempster and Shafer as the proper formalism for the study of evidential problems. These results include all major theorems of Dempster-Shafer theory, including the rule of combination of Dempster. Moreover, the insight provided by the basic conceptual structures introduced here allow a clearer characterization of the semantic aspects of the Dempster-Shafer calculus of evidence.

It is also shown that probabilities defined on the epistemic algebra induce lower and upper probabilities in the truth algebra that are identical to the well-known interval bounds derived in the Dempster-Shafer theory. These results are also consistent with the interpretation of belief measurement advanced by Suppes [5] in his critique of the axiomatic approach of Savage [13].

It is important to note that these results were obtained by simply Carnap's method to epistemic structures. Their derivation appeals only to basic concepts of probability theory and epistemic logic. From this standpoint, Dempster-Shafer theory can be regarded as an enhancement of the classical Bayesian approach to induction which is able to distinguish among different knowledge states.

I-2. Presentation Approach

A major objective of this paper is the development of a comprehensive understanding of epistemic issues that bridges the gap between approaches to knowledge representation and manipulation derived from probability theory, on one hand, and mathematical logic, on the other.

Because of the reliance of different methods on dissimilar types of formalisms — those stressing continuous variable analysis, in probabilistic approaches, and those stressing discrete mathematics, in the case of symbolic logic — it has been difficult to find a common ground for discussion and presentation that will be readily understandable to persons who are familiar only with the concepts and structures of a single methodology. To solve these expository problems we have chosen to present the basic constructs and tenets of each discipline (which are, of course, well known to specialists on both sides) before proceeding to interrelate them within the framework provided by the concept of epistemic universe.

To facilitate the understanding of certain issues, we have also introduced some simplifications, that, if viewed solely from the perspective of the results presented herein, are actually unnecessary. For example, the assumption of finite universes and subset algebras is intended to avoid some of the complexities introduced by the infinite universes and sigma algebras. The results can, however, be readily extended in ways that do not constrain their essential validity.

I-3. Organization

This work is divided into six sections beginning with this introduction.

Section II presents the basic concepts and axioms of epistemic logic that underlie the concept of *epistemic universe*. The important distinction between necessary implication (or entailment) and ordinary implication is also discussed so as to differentiate implications that are valid in every possible world from those that are true only in some possible worlds.

The major result presented in this section states that the epistemic universe may be partitioned into a family of epistemically equivalent disjoint subsets. Each of these sets is characterized by the fact that propositions known to be true in one possible world are the same that are known in any other possible world in the same subset (i.e., both worlds, as far as we know, are equivalent). Furthermore, these subsets can be associated with certain propositions representing the *most detailed* or *specific* knowledge available in each possible world in that epistemic subset.

Section III introduces the basic concepts of probability theory. Emphasis is placed on the nature of subset algebras as the domain of definition of probability functions. It is shown that probabilities defined on the epistemic algebra (generated by the epistemic sets discussed in Section II) have the structure of the basic probability assignments of Dempster-Shafer theory. Furthermore, their associated belief functions (i.e., the probability that a given proposition may be known to be true) are related to these mass assignments by the familiar equations of that theory. Also noteworthy are results assuring that, in the epistemic universe, probabilities defined on epistemic algebras induce lower and upper probabilities over truth algebras defining interval bounds that are identical to those derived in the mathematical theory of evidence of Dempster and Shafer.

Section IV deals with problems associated with combining the knowledge of two rational agents, each dealing with a distinct (albeit possibly the same) universe of discourse. Structures are introduced to represent the agents' state of knowledge before and after the two bodies of knowledge have been combined. The principal result presented in this section is a general evidence combination formula that is the basis for the derivation of specific combination formulas, each corresponding to special assumptions about the nature of the evidential bodies being combined and their relations.

Section V introduces the notion of probabilistic independence. The nature of the structures presented in the previous sections enables a clearer understanding of this important concept than that found in the existing literature. A generalization of Dempster's combination formula is derived on the basis of specific assumptions of independence of certain probabilities.

Finally, Section VI deals with certain simple cases of combination of dependent evidence. The purpose of discussing them is the presentation of general problems of combination of dependent and conditional knowledge that are given detailed attention in a related paper [15].

II
THE EPISTEMIC UNIVERSE

This section presents the basic epistemic logic structures used to produce a framework for the representation of uncertainties discussed in Section III. The basic axioms of epistemic logic are introduced to allow description of the *epistemic universe*, i.e., a space of possible states of the real world and of knowledge about it.

The important notion of *necessary implication* is used, together with the epistemic concept of *accessibility relation*, to define important classes of epistemic sets; i.e. subsets of the epistemic universe that are equivalent with respect the extent of their knowledge. Similarly, subsets of the epistemic universe, called *truth sets*, are defined on the basis of their common ontological properties. It will also be proved below that every possible world in the epistemic universe is associated with a unique sentence that describes the most specific knowledge of the modeled system that is represented in that possible world.

II-1. Epistemic Logic

II-1-1. Symbols

(1) The special characters Θ and φ are symbols.

(2) Elements of a finite alphabet $\mathcal{A}$ are symbols.

II-1-2. Sentences

Sentences will be denoted by script letters $\mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}$ (sometimes with a subscript). Sentences are defined recursively by the following axioms:

(S1) If σ is a symbol, then it is a sentence.

(S2) If $\mathcal{E}, \mathcal{F}$ are sentences, so are $\mathcal{E} \wedge \mathcal{F}$ and $\mathcal{E} \vee \mathcal{F}$.

(S3) If $\mathcal{E}$ is a sentence, so is $\neg\mathcal{E}$.

(S4) If $\mathcal{E}$ is a sentence, so is $\mathbf{K}\mathcal{E}$.

(S5) If $\mathcal{E}$ and $\mathcal{F}$ are sentences, then $\mathcal{E} \rightarrow \mathcal{F}$ is a sentence.

II-1-3. Objective Sentences

If $\mathcal{E}$ is a sentence that does not include the unary operator $\mathbf{K}$, then $\mathcal{E}$ is said to be an *objective* sentence.

II-1-4. Truth Values

The truth values are $\mathbf{T}$ and $\mathbf{F}$, denoting *true* and *false*, respectively.

II-1-5. Sentence Space

The set of all well-formed sentences (henceforth called the *sentence space*) will be denoted $\mathcal{S}$.

II-2. Epistemic Worlds

II-2-1. Interpretations

An interpretation $\mathcal{W}$ for $\mathcal{S}$ is a mapping from the sentence space $\mathcal{S}$ into the set of possible truth values $\{\mathbf{T}, \mathbf{F}\}$.

A sentence $\mathcal{E}$ is *true* in $\mathcal{W}$ if and only if $\mathcal{W}$ maps $\mathcal{E}$ into the truth value $\mathbf{T}$. Otherwise the sentence is said to be *false*.

II-2-2. Possible Worlds

An interpretation $\mathcal{W}$ is a *possible epistemic world* (or simply a *possible world*) for the sentence space $\mathcal{S}$ if and only if it satisfies the following axioms :

(M1) Axioms of ordinary propositional logic with Θ denoting a sentence that is always true and φ denoting a sentence that is always false.

(M2) If $\mathbf{K}\mathcal{E}$ is true, then $\mathcal{E}$ is true.

(M3) If $\mathbf{K}\mathcal{E}$ is true, then $\mathbf{KK}\mathcal{E}$ is true.

(M4) If $\mathbf{K}(\mathcal{E} \to \mathcal{F})$ is true, then $\mathbf{K}\mathcal{E} \to \mathbf{K}\mathcal{F}$ is true.

(M5) If $\mathcal{E}$ is an axiom, then $\mathbf{K}\mathcal{E}$ is an axiom.

(M6) If $\neg\mathbf{K}\mathcal{E}$ is true, then $\mathbf{K}\neg\mathbf{K}\mathcal{E}$ is true.

This axiom schemata is an enhancement (by addition of **(M6)**) of that originally proposed by Moore [2]. The resulting logical system is equivalent to the modal logic system **S5**.

II-2-3. Remark

As is well known, the truth of objective sentences in a possible world $\mathcal{W}$ is determined by the truth of sentences consisting of a single symbol by application of the laws of logic (embodied in the above schemata).

The truth of nonobjective sentences, however, while derivable from the truth of other sentences in $\mathcal{S}$, will not be, in general, a function of the state of knowledge about the truth of simpler propositions (e.g., $\mathbf{K}\mathcal{E}$ and $\mathbf{K}\mathcal{F}$ may be false, but $\mathbf{K}(\mathcal{E} \vee \mathcal{F})$ may be true).

II-2-4. Universes

The space of all possible worlds for $\mathcal{S}$, called the *universe*, will be denoted by $\mathcal{U}(\mathcal{S})$.

II-3. Necessary Truth

II-3-1. Logical Implication

A sentence $\mathcal{E}$ is said to *logically imply* a sentence $\mathcal{F}$, denoted $\mathcal{E} \Rightarrow \mathcal{F}$, if on the basis the axioms of epistemic logic, without regard to the truth value of other sentences that could be *possibly true* or *possibly false* (i.e, not necessarily true or false), and by the use of *rules of deduction*, (e.g., modus ponens), $\mathcal{F}$ can be shown to be *true* whenever $\mathcal{E}$ is true.

Clearly, if $\mathcal{E}$ is true in a possible world $\mathcal{W}$, then $\mathcal{F}$ is also true in $\mathcal{W}$. In addition, if $\mathcal{E} \Rightarrow \mathcal{F}$, then $\mathcal{E} \to \mathcal{F}$ in every possible world $\mathcal{W}$, and conversely.

II-3-2. Semantic Aspects of Logical Implication

The notion of logical implication, introduced above, corresponds to Carnap's notion of *necessary* or *L-implication* [3, pp. 11 ff.].

Informally speaking, the validity of $\mathcal{E} \Rightarrow \mathcal{F}$ means that, solely on the basis of the semantical rules of the system and without reference to extralinguistic facts, the assumed truth of the sentence $\mathcal{E}$ is sufficient to assure the truth of the sentence $\mathcal{F}$.

If, as Carnap points out, $\mathcal{E} \not\Rightarrow \mathcal{F}$, then, by definition, there exists a possible world where $\mathcal{E}$ is true and where $\mathcal{F}$ is false, The truth of $\mathcal{F}$, therefore, cannot be derived solely from the truth of $\mathcal{E}$ and the semantics of the system. To establish such truth *extralinguistic facts* are required (i.e., the truth of other sentences that are themselves not *necessarily true*).

The concept of necessary implication is intended to capture the notion of *semantic entailment* of one sentence by the other as an inevitable consequence of axiomatic considerations, rather than resulting from the mere fact that, in every possible world, whenever the entailing sentence is true, so is the entailed sentence.

It is valid, therefore (as will be done below for problems of knowledge combination), to restrict the scope of a universe by eliminating from consideration possible worlds in which certain pairs of sentences do not satisfy some specific entailment conditions. The semantic interpretation of entailment allows the use of this restriction, which would otherwise be improper (i.e., a possible world is defined in terms of the scope of the class of possible worlds).

II-3-3. Necessary Implications as Axioms

Throughout this work it will be assumed that, if $\mathcal{E} \Rightarrow \mathcal{F}$, then $\mathcal{E} \rightarrow \mathcal{F}$ is an axiom that is true in every possible world $\mathcal{W}$ and, as required by **(M5)**, it is always known. This requirement is very natural; it restricts our consideration of possible worlds to those in which both general laws of epistemic logic and specific truths applicable to certain systems are both assumed to be known.

In what follows, it will be assumed that, if $\mathcal{E} \Rightarrow \mathcal{F}$, then $K(\mathcal{E} \rightarrow \mathcal{F})$ is true for every possible world $\mathcal{W}$ (hence $\mathbf{K}(\mathcal{E} \Rightarrow \mathcal{F})$, and $\mathbf{K}\mathcal{E} \Rightarrow \mathbf{K}\mathcal{F}$).

II-3-4. Logical Equivalence

Two sentences $\mathcal{E}$ and $\mathcal{F}$ are said to be logically equivalent, denoted by $\mathcal{E} \Leftrightarrow \mathcal{F}$, if and only if $\mathcal{E} \Rightarrow \mathcal{F}$ and $\mathcal{F} \Rightarrow \mathcal{E}$.

II-3-5. Frames of Discernment

The quotient space of the set of objective sentences by the equivalence relation $\Leftrightarrow$, will be called a *frame of discernment* and will be denoted by $\Phi(\mathcal{S})$.

II-4. Epistemic States

II-4-1. Simple Epistemic Equivalence

Two possible worlds $\mathcal{W}_1, \mathcal{W}_2$ for the sentence space $\mathcal{S}$ are said to be *epistemically equivalent* (denoted $\mathcal{W}_1 \sim \mathcal{W}_2$) if, for any objective sentence $\mathcal{E}$, the sentence $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}_1$ if and only if it is also true in $\mathcal{W}_2$.

It can be readily verified that the relation $\sim$ is an equivalence relation among possible worlds for $\mathcal{S}$. This relation is called the *accessibility* relation [2].[1]

II-4-2. Epistemic Space

The quotient space of the set $\mathcal{U}(\mathcal{S})$ of all possible worlds for $\mathcal{S}$ by the equivalence relation $\sim$ will be called the *epistemic space* of $\mathcal{S}$ (denoted by $\Sigma(\mathcal{S})$). Its members are called *epistemic states.*

[1] This relation is identical to the accessibility relation of modal logic only in the case of equivalence relations. In our case, this assumption is valid since the underlying modal logic system is equivalent to **S5**.

II-5. Most Specific Knowledge

II-5-1. Existence and Uniqueness of Most Specific Knowledge

Theorem (Most Specific Knowledge): *There exists an injective mapping* **M** *between the space* $\Sigma(\mathcal{S})$ *of all epistemic states and the frame of discernment* $\Phi(\mathcal{S})$. *This mapping assigns a unique (except for logical equivalences) objective sentence* $\mathcal{E} = \mathbf{M}(\mathbf{e})$ *to every epistemic state* **e**, *so that, for every possible world* $\mathcal{W}$ *in* **e**, *and for every objective sentence* $\mathcal{F}$, *the sentence* $\mathbf{K}\mathcal{F}$ *is true in* $\mathcal{W}$ *if and only if* $\mathcal{E} \Rightarrow \mathcal{F}$.

Proof: First note that, by definition, epistemically equivalent possible worlds share a common set of objective sentences $\mathcal{F}$ such that $\mathbf{K}\mathcal{F}$ is true. Proof of the existence of a sentence satisfying the theorem thesis for a possible world $\mathcal{W}$ in the epistemic set **e** is therefore equivalent to its proof for any other possible world $\mathcal{W}'$ in **e** such that $\mathcal{W}' \sim \mathcal{W}$. Furthermore, such a sentence should be the same for all epistemically equivalent possible worlds.

To see that such a sentence exists, consider any possible world $\mathcal{W}$ in **e** and let

$$\{\mathcal{F}_1, \mathcal{F}_2, \ldots\}$$

be the set of all objective sentences $\mathcal{F}$ in $\mathcal{W}$ such that $\mathbf{K}\mathcal{F}$ is true. Let

$$\mathcal{E} = \mathcal{F}_1 \wedge \mathcal{F}_2 \wedge \ldots$$

Then $\mathcal{E}$ is also objective and, by virtue of the axioms of propositional logic (relating truth values of conjuncts with that of their conjunction), if $\mathcal{E}$ is true in any possible world $\mathcal{W}$, then so is $\mathcal{F}_i$, $i = 1, 2, \ldots$. Therefore, by construction, if $\mathbf{K}\mathcal{F}$ is true in $\mathcal{W}$, then $\mathcal{E} \Rightarrow \mathcal{F}$. In particular, since $\mathcal{E} \Rightarrow \mathcal{E}$, $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$.

Conversely, if $\mathcal{E} \Rightarrow \mathcal{F}$, then, since $\mathbf{K}(\mathcal{E} \rightarrow \mathcal{F})$, it follows from (**M4**) that $\mathbf{K}\mathcal{F}$ is true in $\mathcal{W}$.

In addition, if **e** and **e**$'$ are two epistemic spaces such that $\mathbf{M}(\mathbf{e}) = \mathbf{M}(\mathbf{e}') = \mathcal{E}$, then any two possible worlds in **e** or **e**$'$ share the same set of true sentences of the form $\mathbf{K}\mathcal{F}$ where $\mathcal{F}$ is objective (i.e. those that satisfy $\mathcal{E} \Rightarrow \mathcal{F}$). Therefore those two worlds belong to the same epistemic space.

To prove that the mapping is unique, assume that $\mathcal{E}_1$ and $\mathcal{E}_2$ are two sentences that satisfy (for a possible world $\mathcal{W}$) the thesis of the theorem. Then $\mathcal{E}_1 \wedge \mathcal{E}_2$ is objective and true and, by virtue of the assumed properties of $\mathcal{E}_1$ and $\mathcal{E}_2$, it follows that $\mathcal{E}_1 \Rightarrow (\mathcal{E}_1 \wedge \mathcal{E}_2)$ and $\mathcal{E}_2 \Rightarrow (\mathcal{E}_1 \wedge \mathcal{E}_2)$ are both true. But then, since $(\mathcal{E}_1 \wedge \mathcal{E}_2) \Rightarrow \mathcal{E}_1$ and $(\mathcal{E}_1 \wedge \mathcal{E}_2) \Rightarrow \mathcal{E}_2$, it follows that $\mathcal{E}_1$ and $\mathcal{E}_2$ are logically equivalent.

This completes the proof of the theorem.∎

II-5-2. Epistemic Mapping

If $\mathcal{E} = \mathbf{M}(\mathbf{e})$, then $\mathcal{E}$ is said to be the *most specific objective sentence* known in **e**. By abuse of language we shall also use the notation **M** to associate every possible world with the most specific objective sentence known in that possible world.

Conversely, $\mathbf{e}(\mathcal{E})$ will denote the epistemic set containing all the possible worlds that have as their most specific knowledge the objective sentence $\mathcal{E}$. Clearly, **e** is the inverse of **M**, mapping classes of logically equivalent objective sentences (i.e., $\Phi(\mathcal{S})$) into classes of epistemically equivalent possible worlds (i.e., $\Sigma(\mathcal{S})$).

It is important to note, however, that some epistemic sets may be empty. For example, since the sentence φ is false in every possible world, the epistemic set $\mathbf{e}(\varphi)$ does not contain a possible world.

II-5-3. Generalizations of the Most Specific Knowledge Theorem

Note that, although the theorem proved above was concerned with the existence of most specific *objective sentences*, the arguments used are applicable to any class of sentences that is closed under logical conjunction. Moreover, in any possible world there always exists a most specific known sentence (i.e., the conjunction of all sentences, whether objective or not, that are known to be true). Similar generalizations will be invoked in the rest of this paper involving, in each case, collections of subsets of the universe that are closed under set intersection.

II-6. Special Subsets of the Epistemic Universe

II-6-1. Truth and Support Sets

The set of all possible worlds in $\mathcal{U}(S)$ that map the objective sentence $\mathcal{E}$ into the truth value **T**, called the *truth set* of $\mathcal{E}$, will be denoted by $\mathrm{t}(\mathcal{E})$.

The set of all possible worlds in $\mathcal{U}(S)$ that map the sentence $\mathbf{K}\mathcal{E}$, for $\mathcal{E}$ in $\Phi(S)$, into the truth value **T** will be denoted by $\mathrm{k}(\mathcal{E})$. This set will be called the *support set* of $\mathcal{E}$.

II-6-2. Relations between Epistemic and Support Sets

Theorem: *Let $\mathcal{E}$ be a sentence in the frame of discernment $\Phi(S)$. Then*

$$\mathrm{k}(\mathcal{E}) = \bigcup_{\mathcal{F} \Rightarrow \mathcal{E}} \mathrm{e}(\mathcal{F}),$$

where the union is over sentences $\mathcal{F}$ in $\Phi(S)$ such that $\mathcal{F} \Rightarrow \mathcal{E}$.

Proof: By virtue of the theorem proved above, if $\mathcal{W}$ is a possible world and $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$, then there exists an objective sentence $\mathcal{F}$, $\mathcal{F} \Rightarrow \mathcal{E}$, such that $\mathcal{W}$ belongs to $\mathrm{e}(\mathcal{F})$.

Conversely, if $\mathcal{F}$ is objective and $\mathcal{F} \Rightarrow \mathcal{E}$ (treatment of necessary implications as axioms), then $\mathbf{K}\mathcal{E}$ is true.∎

Theorem: *Let $\mathcal{F}$ be a sentence in the frame of discernment $\Phi(S)$. Then*

$$\mathrm{e}(\mathcal{F}) = \mathrm{k}(\mathcal{F}) \cap \left[\bigcup_{\substack{\mathcal{E} \Rightarrow \mathcal{F} \\ \mathcal{E} \not\Leftarrow \mathcal{F}}} \overline{\mathrm{k}(\mathcal{E})} \right],$$

$$\overline{\mathrm{k}(\mathcal{F})} = \bigcup_{\mathcal{E} \not\Rightarrow \mathcal{F}} \mathrm{e}(\mathcal{E}),$$

$$\overline{\mathrm{k}(\mathcal{F})} = \bigcup_{\mathcal{E} \not\Rightarrow \mathcal{F}} \mathrm{k}(\mathcal{E}).$$

Proof: The first relation follows at once from the definition of epistemic set and the most specific knowledge theorem.

The second relation is a direct consequence of the theorem proved immediately above.

Finally, if $\mathbf{K}\mathcal{F}$ is true, then by the most specific knowledge theorem, $\mathbf{K}\mathcal{E}$ is true for some objective sentence $\mathcal{E} \Rightarrow \mathcal{F}$. The converse of this statement is obviously true. It follows, therefore, that

$$\mathrm{k}(\mathcal{F}) = \bigcup_{\mathcal{E} \Rightarrow \mathcal{F}} \mathrm{k}(\mathcal{E}).$$

Taking the complement on both sides, the third relation follows at once.∎

III
PROBABILITIES IN THE EPISTEMIC UNIVERSE

This section introduces probabilities as functions defined over certain special families of subsets of the epistemic universe. These collections have certain important algebraic properties that allow derivation of the probabilities of certain sets as a function of probabilities of other sets in the family. The collection of epistemic sets introduced in the preceding section is shown to have the structure of a subset algebra and probabilities over such a collection are shown to possess the properties of the basic functions of the Dempster-Shafer theory: *belief functions* and *basic probability assignments.*

To study problems associated with the extension of a probabilities defined on a subset algebra to another subset algebra containing it (i.e., a richer subset collection), the conventional probability concepts of lower and upper probabilities are introduced. Belief functions are shown to be the lower probabilities induced in the truth algebra (consisting of sets with similar ontological properties) by a probability defined in the epistemic algebra. These bounds are shown to be attained and therefore to be the best possible. Lower probabilities are further discussed in terms of their more general role as solutions to a wide variety of evidential problems. These results confirm the interpretation of belief measurements advanced by Suppes in 1974 [5][2].

The important concept of conditional probability is also introduced in this section.

III-1. Probability Measures

III-1-1. Subset Algebras

A subset[3] algebra Ω in the set $\mathbf{X}$ is a collection of subsets of $\mathbf{X}$ that includes its empty set, $\mathbf{X}$ itself, and that is closed under set complementation, union, and intersection.

A subset algebra Ω_1 is said to be a *subalgebra* of another subset algebra Ω_2 if and only if $\Omega_1 \subseteq \Omega_2$ as collections of subsets of $\mathbf{X}$.

III-1-2. Special Algebras in the Epistemic Universe

Four subset algebras in $\mathcal{U}(\mathcal{S})$ are of particular interest:

(1) The *trivial* subset algebra of $\mathcal{U}(\mathcal{S})$, consisting solely of $\mathcal{U}(\mathcal{S})$ and its empty set.

(2) The subset algebra $\wp(\mathcal{U}(\mathcal{S}))$ [4], identical to the power set of $\mathcal{U}(\mathcal{S})$, called the *possible worlds* algebra of $\mathcal{U}(\mathcal{S})$.

(3) The subset algebra $\Omega_{\mathbf{E}}$, called the *epistemic* algebra of $\mathcal{U}(\mathcal{S})$, equal to the smallest subset algebra containing the class of the support sets $\mathrm{k}(\mathcal{E})$, where $\mathcal{E}$ is in the frame of discernment $\Phi(\mathcal{S})$ (Note, that by virtue of the relations between epistemic and support sets proved above, the nonempty members of $\Omega_{\mathbf{E}}$ are unions of collections of epistemic sets $\mathrm{e}(\mathcal{F})$, where $\mathcal{F}$ is in the frame of discernment $\Phi(\mathcal{S})$.).

---

[2] It is important to note that these results support Suppes' interpretation of lower and upper probabilities as the major constraints on probabilities over richer subset algebras, i.e., as opposed to the arbitrary interval constraints of interval probability theory. The resulting bounds agree with those derived in the Dempster-Shafer theory only when the inducing probabilities are defined on certain subset algebras, e.g., the epistemic algebra.

[3] For simplicity, we shall restrict our attention to finite sets. Extension to arbitrary spaces requires introduction of the concept of sigma algebra (see, for example, Halmos [10]).

[4] In general, the notation $\wp(\mathbf{X})$ (or simply $\wp$) will be used to denote the power set of a set $\mathbf{X}$.

(4) The subset algebra Ω_T, called the *truth* algebra of $\mathcal{U}(\mathcal{S})$, consisting of the class of the truth sets $t(\mathcal{E})$, where $\mathcal{E}$ is in the frame of discernment $\Phi(\mathcal{S})$.

III-1-3. Probability Functions

A *probability* defined on the subset algebra Ω of a set $\mathbf{X}$ is a mapping $\mathbf{P}$ from Ω into the $[0,1]$ interval of the real line, that satisfies the following conditions

(1) $\mathbf{P}(\emptyset) = 0$,

(2) $\mathbf{P}(\mathbf{X}) = 1$,

(3) $\mathbf{P}(\omega_1 \cup \omega_2) + \mathbf{P}(\omega_1 \cap \omega_2) = \mathbf{P}(\omega_1) + \mathbf{P}(\omega_2)$, whenever ω_1, ω_2 are in Ω.

The triple $(\mathbf{X}, \Omega, \mathbf{P})$ is called a *probability space*.

III-1-4. Consistent Probabilities

Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be probabilities defined on the subset algebras Ω_1 and Ω_2, respectively, of a set $\mathbf{X}$. Then $\mathbf{P}_1$ and $\mathbf{P}_2$ are said to be *consistent* if there exists a probability function $\mathbf{P}$, defined in the smallest subset algebra Ω that includes both Ω_1 and Ω_2, such that

$$\mathbf{P}(\omega) = \begin{cases} \mathbf{P}_1(\omega), & \text{if } \omega \text{ is in } \Omega_1; \\ \mathbf{P}_2(\omega), & \text{if } \omega \text{ is in } \Omega_2. \end{cases}$$

III-2. Lower and Upper Probabilities

III-2-1. Kernel Sets for a Subset Algebra

Let Ω be a subset algebra of the set $\mathbf{X}$. Then the *kernel* of $\wp(\mathbf{X})$ in Ω is the mapping

$$\mathcal{K}_\Omega : \wp(\mathbf{X}) \mapsto \Omega$$

that assigns to every subset ω of $\mathbf{X}$ the largest member of Ω contained in ω (the proof of the existence of such a set is well known and will be omitted here).

For any subset ω in $\wp$, the subset $\mathcal{K}_\Omega(\omega)$ will be called the *kernel* of ω in Ω.

III-2-2. Kernel of a Truth Set in the Epistemic Algebra

Theorem: *The kernel in Ω_E of the truth set $t(\mathcal{E})$ in $\mathcal{U}(\mathcal{S})$ is the support set $k(\mathcal{E})$.*

Proof: If $\mathbf{K}\mathcal{E}$ ($\mathcal{E}$ objective) is true in $\mathcal{W}$ then, by virtue of **(M2)**, $\mathcal{E}$ is true and $k(\mathcal{E}) \subseteq t(\mathcal{E})$. Now let $\mathcal{F}$ be an objective sentence such that

$$e(\mathcal{F}) \subseteq t(\mathcal{E}).$$

Since the sentence

$$\mathbf{K}\mathcal{F} \wedge \Big(\bigwedge_{\substack{\mathcal{G} \Rightarrow \mathcal{F} \\ \mathcal{G} \not\Leftarrow \mathcal{F}}} \neg \mathbf{K}\mathcal{G} \Big),$$

is true in $\mathcal{W}$ if and only if $\mathcal{W}$ is in the epistemic set $e(\mathcal{F})$, it follows that

$$\mathbf{K}\mathcal{F} \wedge \Big(\bigwedge_{\substack{\mathcal{G} \Rightarrow \mathcal{F} \\ \mathcal{G} \not\Leftarrow \mathcal{F}}} \neg \mathbf{K}\mathcal{G} \Big) \Rightarrow \mathcal{E}.$$

But then, by virtue of the axiom (**M4**), it follows that

$$\mathbf{K}\left(\mathbf{K}\mathcal{F} \wedge \left(\bigwedge_{\substack{\mathcal{G}\Rightarrow\mathcal{F}\\ \mathcal{G}\not\Leftarrow\mathcal{F}}} \neg\mathbf{K}\mathcal{G}\right)\right) \Rightarrow \mathbf{K}\mathcal{E}\,.$$

Distributing the epistemic operator **K** in the left-hand-side by application of the axioms (**M3**), (**M5**), and (**M6**), it follows that

$$\mathbf{K}\mathcal{F} \wedge \left(\bigwedge_{\substack{\mathcal{G}\Rightarrow\mathcal{F}\\ \mathcal{G}\not\Leftarrow\mathcal{F}}} \neg\mathbf{K}\mathcal{G}\right) \Rightarrow \mathbf{K}\mathcal{E}\,,$$

or that

$$\mathrm{e}(\mathcal{F}) \subseteq \mathbf{k}(\mathcal{E})\,.$$

Since each epistemic subset that is a subset of $\mathrm{t}(\mathcal{E})$ is also a subset of $\mathbf{k}(\mathcal{E})$, and since any member of the epistemic algebra is a union of (disjoint) epistemic sets, it follows that $\mathbf{k}(\mathcal{E})$ is the largest member of the epistemic algebra that is a subset of $\mathrm{t}(\mathcal{E})$.∎

III-2-3. Lower Probabilities

Given a probability function **P** defined on the subset algebra Ω of a set **X**, the function

$$\mathbf{P}_* : \wp(\mathbf{X}) \mapsto [0,1] : \omega \mapsto \mathbf{P}\big(\mathcal{K}_\Omega(\omega)\big)$$

is called the *lower probability induced by* **P** *in* $\wp(\mathbf{X})$.

III-2-4. Upper Probabilities

Given a probability **P** defined on the subset algebra Ω of the set **X**, the function

$$\mathbf{P}^* : \wp(\mathbf{X}) \mapsto [0,1] : \omega \mapsto 1 - \mathbf{P}_*(\overline{\omega})$$

is called the *upper probability induced by* **P** *in* $\wp(\mathbf{X})$.

The notion of upper probability is related to the concept of a *cover set* through a relationship that is the dual of that between lower probabilities and kernel sets.

Let Ω be a subset algebra of the set **X**. The *cover of* $\wp(\mathbf{x})$ *in* Ω is the mapping

$$\mathcal{C} : \wp(\mathbf{X}) \mapsto \Omega$$

that assigns to every subset ω of **X** the smallest member of Ω that contains ω. The subset $\mathcal{C}(\omega)$ is called the *cover of* ω in Ω.

The proof of the existence of a cover and the proof that

$$\mathbf{P}^*(\omega) = \mathbf{P}\big(\mathcal{C}(\omega)\big)$$

are well known.

III-2-5. Relations Between Probabilities and Lower Probabilities

Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be two consistent probabilities defined on the subset algebras Ω_1 and Ω_2, respectively, of a set **X**. Further, let $\Omega_1 \subseteq \Omega_2$. Also, let $\mathcal{K}_{\Omega_1}$ and $\mathbf{P}_{1*}$ be the kernel of $\wp(\mathbf{X})$ in Ω_1 and the lower probability induced by $\mathbf{P}_1$ respectively. Then, since $\mathcal{K}_{\Omega_1}(\omega) \subseteq \omega$ for any ω in $\wp(\mathbf{X})$ (hence, for any ω in Ω_2), it is

$$\mathbf{P}_{1*}(\omega) = \mathbf{P}_1\big(\mathcal{K}_{\Omega_1}(\omega)\big) \leq \mathbf{P}_1(\omega) = \mathbf{P}_2(\omega)\,.$$

III-2-6. Lower Probabilities as Best Bounds

The following results show that the bounds provided by the relations between lower probabilities and consistent probabilities cannot be improved.

Theorem: *Let* $\mathbf{P}$ *be a probability defined on the subset algebra* Ω *of the universe* $\mathcal{U}(\mathcal{S})$.

If $\Upsilon = \{\omega_1, \omega_2, \ldots, \omega_n\}$ *is a finite collection of subsets of* $\mathcal{U}(\mathcal{S})$, *then there exists a probability* $\widetilde{\mathbf{P}}$ *in* $\widetilde{\Omega}$, *the smallest subset algebra containing both* Ω *and* Υ, *such that*

$$\widetilde{\mathbf{P}}(\omega_1) = \mathbf{P}_*(\omega_1),$$
$$\widetilde{\mathbf{P}}(\omega_i) \geq \mathbf{P}_*(\omega_i), \quad i = 2, \ldots, n,$$
$$\widetilde{\mathbf{P}}(\omega) \geq \mathbf{P}(\omega), \quad \text{for all } \omega \text{ in } \Omega.$$

Proof: Let Ω_1 be the smallest subset algebra containing ω_1 and Ω.

Define the subsets κ and υ of $\mathcal{U}(\mathcal{S})$ as follows:

$$\kappa = K_\Omega(\omega_1),$$
$$\upsilon = \omega_1 \cap \overline{\kappa},$$

and define a set function $\widetilde{\mathbf{P}}_1$ over the following members of Ω_1 as follows:

$$\widetilde{\mathbf{P}}_1(\omega) = 0, \text{ if } \omega \subseteq \upsilon,$$
$$\widetilde{\mathbf{P}}_1(\omega) = \mathbf{P}(\omega \cap \kappa), \text{ if } \omega \subseteq \kappa,$$
$$\widetilde{\mathbf{P}}_1(\omega) = \max_{\Delta(\omega)} \mathbf{P}(\sigma), \text{ if } \omega \cap \omega_1 = \emptyset,$$

where, in the last line, the maximum is defined over the collection of sets

$$\Delta(\omega) = \{\sigma \text{ in } \Omega : \sigma \cap \kappa = \emptyset;\ \sigma \cap \overline{\omega_1} = \omega \cap \overline{\omega_1}\}.$$

The function $\widetilde{\mathbf{P}}_1$ is then defined over other subsets in Ω_1 by additivity (note that the above equations define $\widetilde{\mathbf{P}}_1$ separately over the sets κ, υ, and $\overline{\omega_1}$).

Then, clearly, $\widetilde{\mathbf{P}}_1(\emptyset) = 0$, while

$$\widetilde{\mathbf{P}}_1(\mathcal{U}(\mathcal{S})) = \mathbf{P}(\kappa) + \max_{\Delta(\omega)} \mathbf{P}(\sigma) = \mathbf{P}(\kappa) + \mathbf{P}(\overline{\kappa}) = 1.$$

Note also that $\widetilde{\mathbf{P}}_1$ is additive over subsets of κ in Ω_1, as these are also in Ω, and $\widetilde{\mathbf{P}}_1 \equiv \mathbf{P}$ for such subsets. Further, $\widetilde{\mathbf{P}}_1$ is additive over subsets of υ in Ω_1 since it is equal to zero for all such subsets.

If now χ_1 and χ_2 are disjoint members of Ω_1, both fully contained in $\overline{\omega_1}$, then

$$\widetilde{\mathbf{P}}_1(\chi_1) + \widetilde{\mathbf{P}}_1(\chi_2) = \max_{\Delta(\chi_1)} \mathbf{P}(\sigma) + \max_{\Delta(\chi_2)} \mathbf{P}(\sigma). \tag{III.1}$$

Assume that the maximum values on the right-hand side of the above equation are attained for σ_1 and σ_2, respectively, both nonintersecting with κ.

Then it must be $\sigma_1 \cap \sigma_2 = \emptyset$. Otherwise, since

$$\sigma_1 \cap \sigma_2 \cap \overline{\omega_1} = \chi_1 \cap \chi_2 \cap \overline{\omega_1} = \emptyset,$$

it must be

$$\sigma_1 \cap \sigma_2 \cap \upsilon = \sigma_1 \cap \sigma_2 \neq \emptyset .$$

However, since both σ_1 and σ_2 are in Ω, then $\sigma_1 \cap \sigma_2$, a subset of υ and therefore of ω_1, is also in Ω. Then, by the definition of kernel, it must be

$$\sigma_1 \cap \sigma_2 \subseteq \kappa ,$$

contradicting the assumptions made about both σ_1 and σ_2.

Returning now to Equation III.1, it follows that

$$\begin{aligned} \tilde{\mathbf{P}}_1(\chi_1) + \tilde{\mathbf{P}}_1(\chi_2) &= \mathbf{P}(\sigma_1) + \mathbf{P}(\sigma_2) \\ &= \mathbf{P}(\sigma_1 \cup \sigma_2) \\ &\leq \max_{\Delta(\chi_1 \cup \chi_2)} \mathbf{P}(\sigma) \\ &= \mathbf{P}(\chi_1 \cup \chi_2) . \end{aligned} \tag{III.2}$$

Since, on the other hand, it is obvious that

$$\max_{\Delta(\chi_1 \cup \chi_2)} \mathbf{P}(\sigma) \leq \max_{\Delta(\chi_1)} \mathbf{P}(\sigma) + \max_{\Delta(\chi_2)} \mathbf{P}(\sigma) , \tag{III.3}$$

then combination of the inequalities III.2 and III.3 yields the equation

$$\tilde{\mathbf{P}}_1(\chi_1) + \tilde{\mathbf{P}}_1(\chi_2) = \tilde{\mathbf{P}}_1(\chi_1 \cup \chi_2) ,$$

completing the proof that the set function $\tilde{\mathbf{P}}_1$ is a probability on Ω_1.

Furthermore,

$$\tilde{\mathbf{P}}_1(\omega_1) = \tilde{\mathbf{P}}_1(\omega_1 \cap \kappa) + \tilde{\mathbf{P}}_1(\upsilon) = \tilde{\mathbf{P}}_1(\kappa) = \mathbf{P}_*(\omega_1) ,$$

by the definition of lower probability.

If now $\omega \subseteq \kappa$, then, by definition, $\tilde{\mathbf{P}}_1(\omega) = \mathbf{P}(\omega)$. Otherwise, if ω is in Ω, $\omega \not\subseteq \kappa$, then

$$\tilde{\mathbf{P}}_1(\omega) = \tilde{\mathbf{P}}_1(\omega \cap \overline{\omega_1}) + \tilde{\mathbf{P}}_1(\omega \cap \upsilon) + \tilde{\mathbf{P}}_1(\omega \cap \kappa) \geq \mathbf{P}(\omega \cap \overline{\kappa}) + \mathbf{P}(\omega \cap \kappa) = \mathbf{P}(\omega) .$$

It is also clear from this relation that

$$\tilde{\mathbf{P}}_{1*}(\omega) \geq \mathbf{P}_*(\omega), \quad \text{for all } \omega \text{ in } \wp(\mathcal{U}(\mathcal{S})).$$

If the above probability extension process is repeated inductively for $i = 2, 3, \ldots, n$ replacing $\mathbf{P}$ by $\tilde{\mathbf{P}}_{i-1}$ and letting Ω_i be the smallest subset algebra containing both Ω_{i-1} and ω_i, then $\mathbf{P}_n \equiv \tilde{\mathbf{P}}$ is a probability defined on $\tilde{\Omega}$, the smallest subset algebra containing both Ω and Υ. Further, $\tilde{\mathbf{P}}(\omega_1) = \mathbf{P}_*(\omega_1)$.∎

Corollary: *Let* $\mathbf{P}$, Ω, $\tilde{\Omega}$, *and* Υ *be defined as before. Then there exists a probability* $\mathbf{P}'$ *such that*

$$\begin{aligned} \mathbf{P}'(\omega_1) &= \alpha , \\ \mathbf{P}'(\omega) &\geq \mathbf{P}_*(\omega) , \quad \text{for all } \omega \text{ in } \tilde{\Omega} , \end{aligned}$$

where α *is an arbitrary value between* $\mathbf{P}_*(\omega_1)$ *and* $1 - \mathbf{P}_*(\overline{\omega_1})$.

Proof: Through the same construction process used in the preceding theorem, it is possible to extend the probability function $\mathbf{P}$ to a probability $\hat{\mathbf{P}}$ defined over $\tilde{\Omega}$ such that

$$\hat{\mathbf{P}}(\overline{\omega_1}) = \mathbf{P}_*(\overline{\omega_1}) .$$

The corollary follows immediately by consideration of convex combinations of the probability functions $\tilde{\mathbf{P}}$ and $\hat{\mathbf{P}}$, both defined on the subset algebra $\tilde{\Omega}$.∎

III-2-7. Conditional Probabilities

Let $\mathbf{P}$ be a probability defined on the subset algebra Ω of the set $\mathbf{X}$. Let σ be an arbitrary subset of $\mathbf{X}$ such that $\mathbf{P}(\sigma) > 0$. The function

$$\mathbf{P}(\omega/\sigma) = \frac{\mathbf{P}(\omega \cap \sigma)}{\mathbf{P}(\sigma)}$$

is called the *conditional probability of* $\mathbf{P}$ *with respect to* σ.

III-3. Probabilities on the Epistemic Algebra

III-3-1. Basic Probability Assignments

Let $\mathbf{P}$ be a probability function defined on the epistemic algebra $\Omega_{\mathbf{E}}$ of the universe $\mathcal{U}(\mathcal{S})$. Then $\mathbf{P}$ is said to be a *basic probability assignment on* $\mathcal{U}(\mathcal{S})$. By abuse of language we shall also say that $\mathbf{P}$ is a probability defined on the associated epistemic space $\Sigma(\mathcal{S})$.

From the definitions of probability on $\mathcal{U}(\mathcal{S})$, the subset algebra $\Omega_{\mathbf{E}}$, and the disjointness of epistemic states, it is clear that $\sum_{\mathbf{e}\in\Sigma} \mathbf{P}(\mathbf{e}) = 1$.

III-3-2. Probability Masses

The function $m : \Phi(\mathcal{S}) \mapsto [0,1]$, defined by $m(\mathcal{E}) = \mathbf{P}\big(\mathbf{e}(\mathcal{E})\big)$, is called the *probability mass associated with* $\mathbf{P}$ *in* $\Phi(\mathcal{S})$.

It is clear that $\sum_{\mathcal{E}\in\Phi} m(\mathcal{E}) = 1$.

III-3-3. Support and Plausibility

Let S be a function mapping objective sentences in the frame of discernment $\Phi(\mathcal{S})$ to real numbers in the $[0,1]$ interval, defined by

$$S(\mathcal{E}) = \mathbf{P}\big(\{\mathcal{W} \in \mathcal{U}(\mathcal{S}) : \mathbf{K}\mathcal{E} \text{ true in } \mathcal{W}\}\big) = \mathbf{P}\big(\mathbf{k}(\mathcal{E})\big)\,.$$

The real function S is said to be the *support function*[5] *in* $\Phi(\mathcal{S})$ *associated with the basic probability assignment* $\mathbf{P}$.

Correspondingly, the function Pl defined by

$$Pl : \Phi(\mathcal{S}) \mapsto [0,1] : \mathcal{E} \mapsto 1 - S(\neg\mathcal{E})$$

is said to be the *plausibility function in* $\Phi(\mathcal{S})$ *associated with the basic probability assignment* $\mathbf{P}$. Clearly,

$$Pl(\mathcal{E}) = 1 - \mathbf{P}\big(\neg\mathbf{k}(\mathcal{E})\big)$$

is the upper probability function in $\wp(\mathcal{U}(\mathcal{S}))$ associated with $\mathbf{P}$.

III-3-4. Relation between Supports and Probability Masses

From the definitions of *support* and *plausibility* functions, it follows at once that

$$S(\mathcal{E}) = \sum_{\mathcal{F}\Rightarrow\mathcal{E}} m(\mathcal{F})\,,$$
$$Pl(\mathcal{E}) = \sum_{\mathcal{F}\not\Rightarrow\neg\mathcal{E}} m(\mathcal{F})\,.$$

[5] These functions are called *belief functions* in [Shafer, 7]. The term *support* is used throughout this work as it expresses the meaning of the function S more adequately.

III-3-5. Supports as Lower Probabilities

It was seen earlier that the kernel in the subset algebra $\Omega_{\mathbf{E}}$ of the truth set $\mathbf{t}(\mathcal{E})$ is the set

$$\mathbf{k}(\mathcal{E}) = \bigcup_{\mathcal{F} \Rightarrow \mathcal{E}} \mathbf{e}(\mathcal{F}) .$$

By virtue of the relation between lower probabilities and probabilities, it is clear that, if $\mathbf{P}$ is a probability defined on the subset algebra $\Omega_{\mathbf{E}}$ of the universe $\mathcal{U}(\mathcal{S})$, then

$$S(\mathcal{E}) = \sum_{\mathcal{F} \Rightarrow \mathcal{E}} m(\mathcal{F}) = \sum_{\mathcal{F} \Rightarrow \mathcal{E}} \mathbf{P}(\mathbf{e}(\mathcal{F})) = \mathbf{P}(\mathbf{k}(\mathcal{E})) = \mathbf{P}_*(\mathbf{t}(\mathcal{E})) \leq \mathbf{P}(\mathbf{t}(\mathcal{E})) .$$

Supports are, therefore, lower bounds for the probabilities of certain sets of possible worlds in which a given objective sentence is true. Consistency between probabilities defined on the epistemic algebra $\Omega_{\mathbf{E}}$ and on the truth algebra $\Omega_{\mathbf{T}}$ requires that probabilities over the latter and support functions over the former satisfy certain inequality constraints.

Similarly, it may be seen that plausibility functions are upper bounds for probabilities defined on the truth set algebra $\Omega_{\mathbf{T}}$.

By using the theorem proved above regarding the nature of lower probabilities as sharp bounds for consistent probabilities, it can be seen that the bounds for $\mathbf{P}$ in terms of the support function S cannot be improved (simply consider $\Omega = \Omega_{\mathbf{E}}$ and $\Upsilon = \Omega_{\mathbf{T}}$ in the statement of the theorem of section III-2-6). It follows, therefore, that it is possible to construct a probability $\mathbf{P}$ that for the truth set $\mathbf{t}(\mathcal{E})$, may attain any value satisfying

$$S(\mathcal{E}) \leq \mathbf{P}(\mathbf{t}(\mathcal{E})) \leq Pl(\mathcal{E}) .$$

III-3-6. On the Role of Lower Probabilities

The results just derived show the importance of the concepts of lower and upper probabilities in evidential reasoning. Beyond the validation of support and plausibility functions as bounds on the values of the probability of a truth set, lower and upper probabilities play a central role in most evidential reasoning problems.

The treatment of these problems follows a general two-step scheme. The first step is the translation of evidential observations into a number of consistent probabilities defined over some epistemic universe. The second step consists of the extension of these probabilities to some subset algebra of interest, usually the smallest subset algebra containing those subset algebras identified in the first step. This extension identifies possible values of any probability function defined over the richer subset algebra. The best characterization of possible values, as shown above, is provided by the upper and lower probability functions.

III-3-7. Perfect Probabilistic Information

Note that, in conditions of perfect probabilistic information, i.e. $\mathcal{E} \rightarrow \mathbf{K}\mathcal{E}$ in $\mathcal{U}(\mathcal{S})$, support sets are identical to the truth sets; moreover, the epistemic universe is the same as the Carnapian universe. In these cases, the interval $[S(\mathcal{E}), Pl(\mathcal{E})]$ collapses to a single point and, as is well known, the results provided by the Dempster-Shafer theory are identical to those obtained by the direct application of probability theory to probability functions defined over the truth algebra (which are now known due to the assumed characteristics of the information).

III-3-8. Certain Support Functions

If a support function S in $\Phi(\mathcal{S})$ is associated with an epistemic probability $\mathbf{P}$ such that

$$m(\mathcal{E}) = \mathbf{P}\big(\mathrm{e}(\mathcal{E})\big) = 1\,,$$

then S is said to be a *certain support function focused on* $\mathcal{E}$. It is clear that

$$S(\mathcal{F}) = \begin{cases} 1, & \text{if } \mathcal{E} \Rightarrow \mathcal{F}; \\ 0, & \text{otherwise.} \end{cases}$$

III-3-9. Möbius Inversion Formula

In the ensuing discussion there will be several opportunities to use the following theorem [7, 8], which is a special form of a basic result of combinatorial theory (Möbius inversions), which was studied in detail by G.C. Rota [18].

Theorem: *Let Ω be a finite subset algebra of a set $\mathcal{U}$. Assume that there exist real functions f and g, defined on Ω, such that*

$$g(x) = \sum_{y \subseteq x} f(y), \qquad \text{for all } x \text{ of } \Omega\,.$$

Then it is

$$f(x) = \sum_{y \subseteq x} (-1)^{|y-x|} g(y)\,, \qquad \text{for all } x \text{ of } \Omega\,,$$

where $|y - x|$ is the number of elements z of Ω such that $x \subseteq z \subseteq y$.

III-3-10. Shafer Axioms and Relations

Applying the basic result on Möbius inversions to the epistemic algebra $\Omega_{\mathbf{E}}$, the basic formula expressing probability masses in terms of support functions [7] can be readily derived:

$$m(\mathcal{E}) = \sum_{\mathcal{F} \Rightarrow \mathcal{E}} (-1)^{|\mathcal{E}-\mathcal{F}|} S(\mathcal{F})\,, \tag{III.4}$$

where $|\mathcal{E} - \mathcal{F}|$ is the number of elements $\mathcal{G}$ in $\Phi(\mathcal{S})$ such that $\mathcal{E} \Rightarrow \mathcal{G} \Rightarrow \mathcal{F}$.

By utilizing related results from combinatorial theory, it is possible to derive the following inequality, which Shafer uses as an axiom for support functions:

$$S(\mathcal{E}_1 \vee \cdots \vee \mathcal{E}_n) \geq \sum_{\substack{I \subseteq \{1,\cdots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1}\, S\Big(\bigwedge_{i \in I} \mathcal{E}_i\Big)\,, \tag{III.5}$$

where $|I|$ is the cardinality of the index subset I.

III-3-11. Evidence as Conditional Probabilities

If the sentence space is rich enough to include propositions that describe not only the results of the analysis of evidential observations and measurements, but the actual observations and measurements themselves, then the probabilistic representation of a body of evidence may be thought of as the conditionalization of an epistemic probability with respect to a support set $\mathbf{k}(\xi)$. The proposition ξ represents the actual observations and measurements which are assumed known with certainty.

This interpretation assumes that, through appropriate extension of the universe of discourse, it is possible to represent conditioning with respect to *uncertain evidence* as classical conditional probabilities with respect to sets that represent certain evidence. The validity of this assumption has been discussed by Kyburg [19].

In the rest of this work, the simplified notation $\mathbf{P}(\cdot/\xi)$ will be used as shorthand for $\mathbf{P}(\cdot/\mathbf{k}(\xi))$. In addition, in the rest of this paper, the Greek letter ξ will be used to denote evidential bodies and their associated sets.

The formal impact of evidential observations or measurements may be thought of, therefore, as the replacement of an epistemic probability $\mathbf{P}$ by the result

$$\mathbf{P}(\cdot/\xi) = \mathbf{P}(\cdot/\mathbf{k}(\xi)),$$

of its conditioning with respect to the support set $\mathbf{k}(\xi)$. The result of the integration of successive bodies of evidence $\xi_1, \xi_2, \ldots$ results in the computation (through *knowledge combination* formulas) of the values of the probability functions $\mathbf{P}(\cdot/\xi_1)$, $\mathbf{P}(\cdot/\xi_1, \xi_2), \ldots \mathbf{P}(\cdot/\xi_1, \xi_2, \ldots)$.

III-3-12. Semantics of Evidential Reasoning

The epistemological basis of evidential reasoning lies in its interpretation of observations of the real world as equivalent to certain probabilities that are defined on the epistemic algebra $\Omega_\mathbf{E}$.

According to this view, the informational content of evidence permits more than simple discrimination between possible and impossible states of the world as dependent upon their logical or physical consistency with available observations. Either from past experience, which furnishes the rationale for objective estimates, or from rational considerations, resulting in a consistent belief system leading to subjective estimates, the information provided by evidence allows qualification or quantification of the likelihood of certain propositions given the observed facts.

The evidential approach equates knowledge of the truth of certain propositions with availability of observations supporting that validity. This relation between the concepts of knowledge and confirming evidence is supported by the etymology of words in the epistemological lore. The word *evidence* is related to the Latin verb *vidēre*, "to see", itself cognate with the Greek *ideîn* (from the earlier Greek *wideîn*), and the Sanskrit *vid*, "to know" [16]. The term *evident* literally means "making itself seen" [17]. Asserting that a proposition is known can be thought therefore as declaring the existence of supporting evidence. Whenever such evidence, however, is uncertain or inconclusive, such assertion may be qualified by means of a probabilistic statement. This statement measures the extent by which the observer relies in his measurements and observations as true grounds supporting the knowledge of propositional truth.

Unlike certain classical approaches that assume that experimental data or rational considerations always enable relative quantification of the degree of support accorded by evidence for *every* relevant proposition, the viewpoint expressed here is that evidence, by its very nature, provides information only about the truth of *certain* propositions, while failing to furnish any indication, either relative or absolute, about the truth or falsity of others. In this approach, valuations (i.e., measures of relative evidential support) can be regarded both as assessments of the relative likelihood of certain statements as true descriptors of the real world and as measures of the *resolving power* of evidential bodies. Consequently, an assignment of a null value to the probability of an epistemic set simply states that the evidence does not support the truth of a proposition, rather than indicating explicitly that the proposition is false.

If, for example, a purely objective and frequentist interpretation of probability is considered, the value of a probability over an epistemic set measures how often, in past experience, acquisition of this type of evidence has resulted in *knowledge* that a proposition (but not any other proposition that implies it) was true.

For example, if, at the scene of a crime, a lock of hair is found, having a certain color and other physical properties, such evidence may indicate that the criminal is in a class of individuals specified by hair attributes. Furthermore, errors involved in collecting evidence, analyzing it, or interpreting the analysis may introduce uncertainties about the nature of this class, resulting in the specification of likelihoods for different classes of suspects. This distribution, however, should reflect the fact that the original evidence failed to resolve the identity of specific suspects — an informational deficiency that some approaches erroneously interpret as equal likelihood. These interpretations of evidence assume incorrectly that bodies of evidence are capable of defining a probability distribution over the universe of *possible*, *objective* worlds; i.e., over alternative states of nature that are otherwise incapable of being discriminated by the informational content. In the evidential interpretation, such capability is restricted to certain subset algebras, i.e. those having as *atoms* subsets which may be distinguished by the evidence.

Finally, this interpretation properly recognizes that evidence changes not the state of the world but rather the state of knowledge about it. The acquisition of evidence should, therefore, lead to changes in the formal structures that are used to represent knowledge states.

IV
THE COMBINATION OF KNOWLEDGE

This section is devoted to the discussion of the problems associated with integrating the knowledge of two distinct, mutually trusting agents.

Two important spaces are introduced: the first is simply the Cartesian product of the two universes being combined, representing possible states of knowledge that each agent may possess solely by virtue of the evidence available to him; the second, called the *logical product universe*, represents the possible results of this combination.

Epistemic algebras and probabilities defined on them are then related to arrive at the additive combination theorem, a general result that serves as a basis for deriving a variety of evidence combination formulas.

IV-1. Multiple Epistemic Agents

IV-1-1. Multiple Worlds

Throughout the rest of this work we shall be considering more than one universe $\mathcal{U}(\mathcal{S})$, each satisfying the axioms of epistemic logic. Each universe may be based on different symbol alphabets $\mathcal{A}_1, \mathcal{A}_2, \ldots$. In general, therefore, different universes will contain possible worlds that map different sentence spaces into the set of truth values $\{\mathbf{T}, \mathbf{F}\}$.

Different universes and their associated structures (i.e., epistemic spaces, most specific mappings, frames of discernment, etc.) will be differentiated by the use of numeric subscripts, e.g. $\mathcal{U}(\mathcal{S}_1)$, and $\Phi(\mathcal{S}_2)$.

IV-1-2. Multiple Agents

Subscripts will also be used, whenever necessary, to differentiate unary epistemic operators, e.g. $\mathbf{K}_1, \mathbf{K}_2$.

Different operators will usually be interpreted as representing the knowledge of different, rational, and mutually trusting agents that are combining their knowledge to arrive at a consensual agreement (usually represented by an unsubscripted epistemic operator $\mathbf{K}$).

Throughout this section we consider the problems that arise when of the knowledge of two mutually trusting rational agents is combined.

IV-2. Product Spaces and Universes

IV-2-1. Product of Sentence Spaces

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two sentence spaces. The *logical product* $\mathcal{S}_1 \otimes \mathcal{S}_2$ of $\mathcal{S}_1$ and $\mathcal{S}_2$ (or *product*, for short) is the space of all sentences defined by the axioms:

(PS1) If $\mathcal{E}$ is an objective sentence in $\mathcal{S}_i$, for $i = 1$ or 2, then $\mathcal{E}$ is a sentence in the product space $\mathcal{S}_1 \otimes \mathcal{S}_2$.

(PS2) Axioms **(S2)** – **(S5)**, given in section II-1-2, defining well-formed epistemic sentences.

IV-2-2. Possible Worlds in Product Space

A possible world in the product sentence space $S_1 \otimes S_2$ is a mapping from that space into the truth set $\{\mathbf{T}, \mathbf{F}\}$ that satisfies the following axioms:

(PU1) The possible world $\mathcal{W}$ satisfies the axiom schemata **(M)**.

(PU2) If $\mathcal{E}$ is a sentence in $S_1 \otimes S_2$, then the sentence $\mathbf{K}\mathcal{E}$ is true in the possible world $\mathcal{W}$ if and only if there exist sentences $\mathcal{E}_1$ and $\mathcal{E}_2$ in S_1 and S_2, respectively, such that $\mathbf{K}\mathcal{E}_i$ is true in $\mathcal{W}$ and $\mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{E}$.

IV-2-3. Remarks

It is important to note that the notion of possible worlds in a product space is not only well defined, but also captures the notion of combination of the knowledge of two evidential bodies.

Axiom **(PU2)** restricts the scope of possible worlds by requiring that propositions known to be true in a possible world in the product space be either sentences in the spaces being combined that are known to be true in that world; or logical consequences (i.e., necessary or semantic entailments) of the truth of those propositions.

The correctness of this type of axiom was noted above in section II-3-2 when discussing the semantic aspects of logical implication.

IV-2-4. Product of Universes

The set of all possible worlds over the product sentence $S_1 \otimes S_2$ that satisfy the axiom schemata **(PU)** will be called the *logical product universe* over $S_1 \otimes S_2$ (or *product universe*, for short), denoted $\mathcal{U}_\otimes(S_1 \otimes S_2)$.

IV-2-5. Product of Frames of Discernment

The *product* of two frames of discernment $\Phi(S_1)$ and $\Phi(S_2)$ is the frame of discernment $\Phi(S_1 \otimes S_2)$, equal to the quotient set of the objective sentences of $S_1 \otimes S_2$ by the equivalence relation $\Leftrightarrow$ in $\mathcal{U}_\otimes(S_1 \otimes S_2)$.

IV-2-6. Remarks

Note that, by construction, the symbols in the product space $S_1 \otimes S_2$ are those that represent sentences that are in either S_1 in S_2 with the unary operators $\mathbf{K}_1$ and $\mathbf{K}_2$ replaced by $\mathbf{K}$.

Furthermore, also by construction, the truth of objective sentences in a possible world $\mathcal{W}$ in $\mathcal{U}_\otimes(S_1 \otimes S_2)$ is determined by the truth of sentences that are either in $\Phi(S_1)$ or in $\Phi(S_2)$. If $\mathcal{E}$ is a sentence that does not include the epistemic operator $\mathbf{K}$, then $\mathcal{E}$ is true in $\mathcal{W}$ if and only there exist sentences $\mathcal{E}_1$ in S_1 and $\mathcal{E}_2$ in S_2 such that $\mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{E}$.

IV-2-7. Cartesian Projection

The mapping $\mathbf{\Pi}$, defined by the expression

$$\mathbf{\Pi} : \mathcal{U}_\otimes(S_1 \otimes S_2) \mapsto \mathcal{U}(S_1) \times \mathcal{U}(S_2) : \mathcal{W} \mapsto (\mathcal{W}_1, \mathcal{W}_2),$$

where $\mathcal{W}_i$ $(i = 1, 2)$ is the unique possible world in $\mathcal{U}(S_i)$ defined by the conditions

(1) The sentence $\mathcal{E}$ in $\Phi(S_i)$ is true in $\mathcal{W}_i$ if and only if $\mathcal{E}$ is true in $\mathcal{W}$,

(2) The sentence $\mathbf{K}_i\mathcal{E}$, with $\mathcal{E}$ in $\Phi(S_i)$, is true in $\mathcal{W}_i$ if and only if the sentence $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$,

is called the *Cartesian projection* of $\mathcal{U}_\otimes(S_1 \otimes S_2)$.

Proposition: *The Cartesian projection of $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ is an injective mapping.*

Proof: As noted above, the truth of objective sentences in $\mathcal{W}$ is determined by the truth of objective sentences in $\mathcal{E}_1$ or in $\mathcal{E}_2$.

Furthermore, if a sentence is of the form $\mathbf{K}\mathcal{E}$, where $\mathcal{E}$ is objective, its truth is also determined (because of the axiom (**PU2**)) by the truth of sentences of the form $\mathbf{K}\mathcal{F}$, where $\mathcal{F}$ is in either $\Phi(S_1)$ or $\Phi(S_2)$.

Finally, the truth of every other sentence in $S_1 \otimes S_2$ is determined by the truth of sentences in the above classes.

Therefore, if two possible worlds in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ assign the same truth values to sentences in S_1 and S_2, then they assign the same truth values to all sentences; in other words, they are identical.∎

IV-3. Epistemic Sets in the Product Universe

IV-3-1. Marginal Support Sets

The *marginal support set* in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ for the sentence $\mathcal{E}_i$ in $\Phi(S_i)$, denoted by $\hat{\mathrm{k}}_i(\mathcal{E}_i)$, is the set of all possible worlds $\mathcal{W}$ in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ such that $\mathbf{K}\mathcal{E}_i$ is true in $\mathcal{W}$, $(i = 1, 2)$.

IV-3-2. Marginal Epistemic Sets

The *marginal epistemic set* in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ for the sentence $\mathcal{E}_i$ in $\Phi(S_i)$, $(i = 1, 2)$, denoted by $\hat{\mathrm{e}}_i(\mathcal{E}_i)$; is the set of all possible worlds $\mathcal{W}$ in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ such that $\mathcal{E}_i$ is the most specific sentence in $\Phi(S_i)$ that is known true in $\mathcal{W}$. The existence of this sentence is proved in the same manner as before (section II-5).

It is clear from the definitions and the basic theorem relating epistemic and support sets that

$$\hat{\mathrm{k}}_i(\mathcal{E}) = \bigcup_{\mathcal{F} \Rightarrow \mathcal{E}} \hat{\mathrm{e}}_i(\mathcal{F}), \quad i = 1, 2,$$

where the union in the above formula is over marginal epistemic sets $\hat{\mathrm{e}}_i(\mathcal{F})$ such that $\mathcal{F}$ is in $\Phi(S_i)$ and $\mathcal{F} \Rightarrow \mathcal{E}$.

Note also that, since there is always a sentence $\mathcal{E}_i$ in $\Phi(S_i)$, $i = 1, 2$, such that $\mathbf{K}\mathcal{E}_i$ is true (e.g., $\mathcal{E}_i = \Theta_i$), then every possible world in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$ belongs to some marginal epistemic set $\hat{\mathrm{e}}_i(\mathcal{E})$, $i = 1, 2$.

IV-3-3. The Basic Combination Theorem

Lemma : *Let $\mathcal{E}$ be a sentence in $\Phi(S_1 \otimes S_2)$. If $\mathrm{e}(\mathcal{E})$ is nonvoid, then there exist sentences $\mathcal{E}_1$ and $\mathcal{E}_2$ in $\Phi(S_1)$ and $\Phi(S_2)$, respectively, such that $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2$.*

Proof: Let $\mathcal{W}$ be in the nonvoid epistemic set $\mathrm{e}(\mathcal{E})$. Since $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$, then, by virtue of (**PU2**) there exist sentences $\mathcal{E}_1$ and $\mathcal{E}_2$ in S_1 and S_2, respectively, such that $\mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{E}$ and $\mathbf{K}(\mathcal{E}_1 \wedge \mathcal{E}_2)$ is true in $\mathcal{W}$.

But then, by the definition of epistemic set, it must be $\mathcal{E} \Rightarrow \mathcal{E}_1 \wedge \mathcal{E}_2$ and, since also $\mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{E}$, it follows that $\mathcal{E}_1 \wedge \mathcal{E}_2 \Leftrightarrow \mathcal{E}$.∎

Corollary: *There exists a mapping*

$$\Gamma : \Phi(S_1 \otimes S_2) \mapsto \wp(\Phi(S_1) \times \Phi(S_2))$$

that assigns to every sentence $\mathcal{E}$ in $\Phi(S_1 \otimes S_2)$ a subset of sentence pairs $(\mathcal{E}_1, \mathcal{E}_2)$ with $\mathcal{E}_i$ in $\Phi(S_i)$, $i = 1, 2$, such that $\mathcal{E}_1 \wedge \mathcal{E}_2 \Leftrightarrow \mathcal{E}$ in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$.

Further, if $\mathcal{E} \not\Leftrightarrow \mathcal{E}'$, then

$$\Gamma(\mathcal{E}) \cap \Gamma(\mathcal{E}') = \emptyset .$$

Proof: The existence of the mapping Γ for sentences $\mathcal{E}$ such that $\mathrm{e}(\mathcal{E})$ is nonvoid follows directly from the lemma. The definition of Γ is completed by defining $\Gamma(\mathcal{E}) = \emptyset$ whenever $\mathrm{e}(\mathcal{E}) = \emptyset$.

If $\Gamma(\mathcal{E}) \cap \Gamma(\mathcal{E}') \neq \emptyset$, then there exist, by virtue of the lemma, a possible world $\mathcal{W}$ in $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$ and sentences $\mathcal{E}_1$ and $\mathcal{E}_2$ in $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, such that $\mathbf{K}(\mathcal{E}_1 \wedge \mathcal{E}_2)$ is true in $\mathcal{W}$ and

$$\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2 \Leftrightarrow \mathcal{E}' .$$

It is thus clear that $\Gamma(\mathcal{E}) \equiv \Gamma(\mathcal{E}')$.∎

The mapping $\Gamma(\mathcal{E})$ will be called the *compatibility relation* of $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$.

Theorem (Basic Combination Theorem): *Let $\mathcal{E}$ be an objective sentence in the product frame of discernment $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$. Then*

$$\mathrm{e}(\mathcal{E}) = \bigcup_{\Gamma(\mathcal{E})} [\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)] , \tag{IV.1}$$

where the union is over all pairs $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Phi(\mathcal{S}_1) \times \Phi(\mathcal{S}_2)$ such that $(\mathcal{E}_1, \mathcal{E}_2)$ is in $\Gamma(\mathcal{E})$.

Proof: If $\mathrm{e}(\mathcal{E})$ is empty, then $\Gamma(\mathcal{E})$ is empty and the theorem follows at once.

Otherwise, let $\mathcal{W}$ be in $\mathrm{e}(\mathcal{E})$. Then since every possible world is in some marginal epistemic set $\hat{\mathrm{e}}_i(\mathcal{E}_i)$, $i = 1, 2$, it follows that there exist sentences $\mathcal{E}_1, \mathcal{E}_2$ in $\Phi(\mathcal{S}_1)$ and $\Phi(\mathcal{S}_2)$, respectively, such that $\mathcal{W}$ is in $\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)$.

Now, by virtue of the lemma, there exist sentences $\mathcal{F}_1$ and $\mathcal{F}_2$ in $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, such that $\mathcal{E} \Leftrightarrow \mathcal{F}_1 \wedge \mathcal{F}_2$. Further, both $\mathbf{K}\mathcal{E}_1$ and $\mathbf{K}\mathcal{E}_2$ are true in $\mathcal{W}$.

By the definition of marginal epistemic set it follows that $\mathcal{E}_i \Rightarrow \mathcal{F}_i$, $i = 1, 2$, and therefore $\mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{F}_1 \wedge \mathcal{F}_2 \Leftrightarrow \mathcal{E}$. Since $\mathcal{W}$, on the other hand, is in $\mathrm{e}(\mathcal{E})$, and since $\mathbf{K}(\mathcal{E}_1 \wedge \mathcal{E}_2)$ is true in $\mathcal{W}$, it follows that $\mathcal{E} \Rightarrow \mathcal{E}_1 \wedge \mathcal{E}_2$ and therefore

$$\mathcal{E}_1 \wedge \mathcal{E}_2 \Leftrightarrow \mathcal{E} .$$

Now let $\mathcal{W}$ be in $\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)$, such that $\mathcal{E}_1$ and $\mathcal{E}_2$ are in $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, and $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2$.

Then, since $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$, it follows that $\mathbf{K}\mathcal{F}$ is true whenever $\mathcal{E} \Rightarrow \mathcal{F}$.

If, on the other hand, $\mathbf{K}\mathcal{F}$ is true in $\mathcal{W}$, then by virtue of the axiom (**PU2**), there exist sentences $\mathcal{F}_1$ and $\mathcal{F}_2$ in $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, such that $\mathbf{K}\mathcal{F}_1$ and $\mathbf{K}\mathcal{F}_2$ are both true in $\mathcal{W}$ and $\mathcal{F}_1 \wedge \mathcal{F}_2 \Rightarrow \mathcal{F}$. However, since $\mathcal{W}$ is in $\hat{\mathrm{e}}_1(\mathcal{E}_1)$ and in $\hat{\mathrm{e}}_2(\mathcal{E}_2)$, it follows that

$$\mathcal{E}_1 \Rightarrow \mathcal{F}_1 \text{ and } \mathcal{E}_2 \Rightarrow \mathcal{F}_2 .$$

Therefore, it is $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2 \Rightarrow \mathcal{F}_1 \wedge \mathcal{F}_2 \Rightarrow \mathcal{F}$.∎

IV-4. Marginal and Product Epistemic Algebras

IV-4-1. Marginal Epistemic Algebras

The smallest algebra in $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$ that contains the marginal epistemic sets $\hat{\mathrm{e}}_i(\mathcal{E})$ for $\mathcal{E}$ in $\Phi(\mathcal{S}_i)$, for $i =$ 1 or 2, will be called the *marginal epistemic algebra* for $\Phi(\mathcal{S}_i)$, denoted by $\Omega^i_{\mathrm{E}}(\mathcal{U}_{\otimes})$.

IV-4-2. Product Epistemic Algebra in the Product Universe

The smallest subset algebra of $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$ that contains the marginal epistemic algebras $\Omega^1_{\mathrm{E}}(\mathcal{U}_{\otimes})$ and $\Omega^2_{\mathrm{E}}(\mathcal{U}_{\otimes})$ is called the *product epistemic algebra* of $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$, denoted by $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_{\otimes})$.

Clearly, this subset algebra contains all the sets of the form $\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)$, and therefore, by virtue of the basic combination theorem, all epistemic sets $\mathrm{e}(\mathcal{E})$.

Note, however, that the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_{\otimes})$ contains, in general, the epistemic algebra $\Omega_{\mathrm{E}}(\mathcal{U}_{\otimes})$ of $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$. The atoms of the latter, as shown by the basic combination theorem, are unions of atoms of the former.

IV-4-3. Special Subset Algebras in $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$

Let $\tilde{e}_1(\mathcal{E}_1)$ and $\tilde{e}_2(\mathcal{E}_2)$ denote the subsets of $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$ defined by the expressions

$$\tilde{e}_1(\mathcal{E}_1) = e_1(\mathcal{E}_1) \times \mathcal{U}(S_2),$$
$$\tilde{e}_2(\mathcal{E}_2) = \mathcal{U}(S_1) \times e_2(\mathcal{E}_2),$$

for $\mathcal{E}_1$ in $\Phi(S_1)$ and $\mathcal{E}_2$ in $\Phi(S_2)$.

The *marginal epistemic algebras* $\Omega^1_E(\mathcal{U}_\times)$ and $\Omega^2_E(\mathcal{U}_\times)$ are the smallest subset algebras of $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$ that contain the collections of subsets

$$\{\tilde{e}_1(\mathcal{E}_1) : \mathcal{E}_1 \text{ is in } \Phi(S_1)\}, \quad \{\tilde{e}_2(\mathcal{E}_2) : \mathcal{E}_2 \text{ is in } \Phi(S_2)\},$$

respectively.

The *product epistemic algebra* $\widehat{\Omega}_E(\mathcal{U}_\times)$ is the smallest subset algebra of $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$ that contains the marginal epistemic algebras Ω^1_E and Ω^2_E. Once again, we must note that this algebra contains, in general, the algebra generated by the projections by the mapping $\mathbf{\Pi}$ of the epistemic sets $e(\mathcal{E})$ in $\mathcal{U}_\otimes(S_1 \otimes S_2)$.

IV-4-4. Epistemic Sets in the Cartesian Product $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$

By abuse of language, the term *marginal epistemic sets* (in $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$) will be used to describe sets of the type $\tilde{e}_1(\mathcal{E}_1)$ and $\tilde{e}_2(\mathcal{E}_2)$, with $\mathcal{E}_1$ in $\Phi(S_1)$ and $\mathcal{E}_2$ in $\Phi(S_2)$. The corresponding support sets will be similarly referred to as *marginal support sets.*

The term *product epistemic sets* (in $\mathcal{U}(S_1)\times\mathcal{U}(S_2)$) will be used to describe those subsets of $\mathcal{U}(S_1)\times\mathcal{U}(S_2)$ that are in the subset algebra generated by sets of the form

$$\tilde{e}_1(\mathcal{E}_1) \cap \tilde{e}_2(\mathcal{E}_2) = e_1(\mathcal{E}_1) \times e_2(\mathcal{E}_2), \quad \mathcal{E}_1 \text{ in } \Phi(S_1),\ \mathcal{E}_2 \text{ in } \Phi(S_2).$$

IV-5. Probabilities in the Product Universe

IV-5-1. Relations between Compatibility Mappings and Cartesian Projections

Theorem: *Let $\mathcal{E}$ be an objective sentence in S such that the epistemic set $e(\mathcal{E})$ is nonvoid in $\mathcal{U}_\otimes(S_1\otimes S_2)$.*

If $\mathcal{W}$ is a possible world in $\mathcal{U}_\otimes(S_1 \otimes S_2)$, then there exists a pair $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Gamma(\mathcal{E})$ such that the Cartesian projection $(\mathcal{W}_1, \mathcal{W}_2)$ of $\mathcal{W}$ is in the subset $e_1(\mathcal{E}_1) \times e_2(\mathcal{E}_2)$ of $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$

Further, if $(\mathcal{W}_1, \mathcal{W}_2)$ is in $e_1(\mathcal{E}_1) \times e_2(\mathcal{E}_2)$, $(\mathcal{E}_1, \mathcal{E}_2)$ in $\mathbf{\Pi}(\mathcal{E})$, then there exists a unique world $\mathcal{W}$ in $\mathcal{U}_\otimes(S_1 \otimes S_2)$ such that $\mathbf{\Pi}(\mathcal{W}) = (\mathcal{W}_1, \mathcal{W}_2)$.

Proof: By the definition of Cartesian projection, a sentence $\mathbf{K}_i\mathcal{E}$, with $\mathcal{E}$ in $\Phi(S_i)$ $i = 1, 2$, is true in the world $\mathcal{W}_i$ in $\mathcal{U}(S_i)$ if and only if $\mathbf{K}\mathcal{E}$ is true in $\mathcal{W}$.

Therefore $\mathbf{\Pi}$ maps worlds in $\hat{e}_i(\mathcal{E}_i)$ into pairs of worlds in $\tilde{e}_i(\mathcal{E}_i)$, $i = 1, 2$.

Since every world $\mathcal{W}$ in $e(\mathcal{E})$ is a set intersection of the form $\hat{e}_1(\mathcal{E}_1) \cap \hat{e}_2(\mathcal{E}_2)$ for some pair $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Gamma(\mathcal{E})$, it follows at once that $\mathbf{\Pi}$ maps $\mathcal{W}$ into the Cartesian product

$$e_1(\mathcal{E}_1) \times e(\mathcal{E}_2) = \tilde{e}_1(\mathcal{E}_1) \cap \tilde{e}_2(\mathcal{E}_2).$$

Conversely, if $(\mathcal{W}_1, \mathcal{W}_2)$ is in $e_1(\mathcal{E}_1) \times e_2(\mathcal{E}_2)$, $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Gamma(\mathcal{E})$, then, as proved above, there exists a unique possible world $\mathcal{W}$ in $\mathcal{U}_\otimes(S_1 \otimes S_2)$, mapped by $\mathbf{\Pi}$ into $(\mathcal{W}_1, \mathcal{W}_2)$. Further, $\mathcal{W}$ is in $e(\mathcal{E}_1 \wedge \mathcal{E}_2) \equiv e(\mathcal{E})$.∎

Corollary: *The mapping* Π *defines a one-to-one transformation between* $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ *and a subset* $\Pi_\otimes$ *of the Cartesian product* $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$ *such that*

(i) The images of the marginal epistemic algebras $\Omega^i_{\mathrm{E}}(\mathcal{U}_\otimes)$ *of* $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ *by* Π *are subalgebras of the marginal epistemic algebras* $\Omega^i_{\mathrm{E}}(\mathcal{U}_\times)$ *of* $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$.

(ii) The image of the epistemic algebra $\Omega_{\mathrm{E}}(\mathcal{U}_\otimes)$ *of* $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ *is a subalgebra of the product epistemic algebra* $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$ *of* $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$.

(iii) The image $\Pi_\otimes$ *of* $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ *by* Π *is the union of sets of the type* $\mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2)$ *with* $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *in* $\Phi(\mathcal{S}_1)$ *and* $\Phi(\mathcal{S}_2)$, *respectively.*

Proof: The first and second parts of the corollary follow immediately from the theorem.

To prove the third part, it is sufficient to observe that, if $\Pi_\otimes$ intersects a set of the type $\mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2)$, then it includes such a set.∎

IV-5-2. Probabilities Induced in the Product Universe

Let $\mathbf{P}$ be a probability defined on the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$ of $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$. Then the probability function $\widehat{\mathbf{P}}$ defined on the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\otimes)$ by the expression

$$\widehat{\mathbf{P}}(\omega) = \mathbf{P}\big(\Pi(\omega)/\Pi_\otimes\big)$$

will be called the *epistemic probability* induced by Π and $\mathbf{P}$ in $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$.

The proof that $\widehat{\mathbf{P}}$ is a probability is straightforward and therefore it will be omitted here.

IV-5-3. The Nature of Probabilities in Product Universes

The injective nature of the Cartesian projection Π indicates that $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ may be regarded as a set embedded in the Cartesian product $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$. The epistemic probability is thus the result of constraining probabilistic knowledge in $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$ to those worlds that are possible after the knowledge of two agents has been combined.

The foregoing developments have resulted in the introduction of two probability functions, $\mathbf{P}$ and $\widehat{\mathbf{P}}$, in universes that, respectively, represent possible states of knowledge of two rational agents, on one hand, and the results of the combining that knowledge, on the other. Through the Cartesian projection mapping Π, these probabilities can be considered as defined on the same Cartesian space $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$, with the function $\widehat{\mathbf{P}}$ being the conditional probability of $\mathbf{P}$ with respect to the subset $\Pi_\otimes$ of $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$.

The unconditioned probability $\mathbf{P}$ represents possible states of knowledge of two agents, A_1 and A_2, *prior* to their combination. In such a state, it is conceivable for one of the agents, for example A_1, to regard as possible consideration by A_2 of certain worlds that are logically inconsistent with the evidence available to A_1. The probability $\mathbf{P}\big(\mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2)\big)$ is a measure of the joint degree of support given by A_1 and by A_2 to the possibility that the evidence known to A_1 supports $\mathcal{E}_1$ and that the evidence known to A_2 supports the truth of $\mathcal{E}_2$. It is explicitly assumed that neither A_1 nor A_2 know the evidence available to the other. It is then possible for one of them to consider that a proposition may be true while the other (with different evidence at his disposal) may know such proposition to be false.

The probability $\widehat{\mathbf{P}}$, on the other hand, represents the consensus of both agents after consideration of the evidence available to *both*. Unlike $\mathbf{P}$, which represents states of separate knowledge, the probability $\widehat{\mathbf{P}}$ represents the result of their integration. Clearly, logical or physical impossibilities, represented by pairs $(\mathcal{E}_1, \mathcal{E}_2)$ in the complement $\overline{\Pi_\otimes}$, cannot be assigned positive probabilities. The remaining possibilities, represented by the set intersections $\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)$, are given probabilistic weights that are consistent with the support values assigned prior to evidential combination and with the known relations between $\hat{\mathrm{e}}_1(\mathcal{E}_1)$ and $\hat{\mathrm{e}}_2(\mathcal{E}_2)$.

IV-5-4. Relations Between Inducing and Induced Probabilities

Proposition: *Let $\widehat{\mathbf{P}}$ be the epistemic probability induced by $\mathbf{P}$ and $\mathbf{\Pi}$ in $\mathcal{U}_{\otimes}(S_1 \otimes S_2)$. Further, assume that $\mathbf{P}(\mathbf{\Pi}_{\otimes})$ is positive.*

If ω in the subset algebra $\Omega_{\mathrm{E}}(\mathcal{U}_{\times})$, is such that $\omega = \mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2)$ for some sentences $\mathcal{E}_1$ in $\Phi(S_1)$ and $\mathcal{E}_2$ in $\Phi(S_2)$, then

$$\widehat{\mathbf{P}}(\omega) = \begin{cases} \kappa\,\mathbf{P}\big(\mathbf{\Pi}(\omega)\big), & \text{if } \omega \subseteq \mathbf{\Pi}_{\otimes}; \\ 0, & \text{otherwise.} \end{cases}$$

where $\kappa = \mathbf{P}(\mathbf{\Pi}_{\otimes})^{-1}$ is a constant independent of ω.

Proof: If

$$\omega = \mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2) \subseteq \mathbf{\Pi}_{\otimes}\,,$$

for some $\mathcal{E}_1$ in S_1 and $\mathcal{E}_2$ in S_2, then $\omega \cap \mathbf{\Pi}_{\otimes} = \omega$ and, clearly,

$$\widehat{\mathbf{P}}(\omega) = \kappa\,\mathbf{P}(\omega)\,.$$

Otherwise $\omega \cap \mathbf{\Pi}_{\otimes} = \emptyset$ and

$$\widehat{\mathbf{P}}(\omega) = \mathbf{P}(\omega/\mathbf{\Pi}_{\otimes}) = 0. \blacksquare$$

IV-5-5. Knowledge Combination as Conditionalization

It has been remarked in Section III-3-11 that epistemic probabilities may be interpreted as the result of conditioning an epistemic probability $\mathbf{P}$ with respect to a support set $\mathbf{k}(\xi)$ that represents possible worlds consistent with evidential observations. Furthermore, combination of two evidential bodies may be regarded as the derivation of a conditional epistemic probability $\mathbf{P}(\cdot/\xi_1, \xi_2)$ from the probability functions $\mathbf{P}(\cdot/\xi_1)$ and $\mathbf{P}(\cdot/\xi_2)$.

Since, by virtue of the basic combination theorem, it is

$$\mathrm{e}(\mathcal{E}) = \bigcup_{\Gamma(\mathcal{E})} \left[\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)\right],$$

then probabilistic knowledge combination can be regarded as the derivation of the conditional probabilities

$$\mathbf{P}\big(\hat{\mathrm{e}}_1(\mathcal{E}_1) \cap \hat{\mathrm{e}}_2(\mathcal{E}_2)/\xi_1, \xi_2\big)\,,$$

that allow computation of

$$\mathbf{P}(\mathrm{e}(\mathcal{E})/\xi_1, \xi_2)\,,$$

from the conditional probabilities

$$\mathbf{P}(\cdot/\xi_1) = m_1(\cdot)\,, \quad \mathbf{P}(\cdot/\xi_2) = m_2(\cdot)\,.$$

IV-6. Probability Masses in the Product Universe

IV-6-1. The Additive Combination Theorem

Theorem (Additive Combination Theorem): *Let $\widehat{\mathbf{P}}$ be a probability defined on the product epistemic algebra $\widehat{\Omega}_{\mathbf{E}}(\mathcal{U}_\times)$ of $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$, and let m be its associated probability mass. Then, if $\mathcal{E}$ is in $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$ it is*

$$m(\mathcal{E}) = \sum_{\Gamma(\mathcal{E})} \widehat{\mathbf{P}}\big(\hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2)\big)\,, \tag{IV.2}$$

where the sum is over all pairs $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Phi(\mathcal{S}_1) \times \Phi(\mathcal{S}_2)$ such that $(\mathcal{E}_1, \mathcal{E}_2)$ is in $\Gamma(\mathcal{E})$.

Furthermore, if $\widehat{\mathbf{P}}$ is the epistemic probability induced by $\mathbf{P}$ in $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$ and if $\mathbf{P}(\mathbf{\Pi}_\otimes) > 0$, then

$$m(\mathcal{E}) = \kappa \sum_{\Gamma(\mathcal{E})} \mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2)\big)\,,$$

where

$$\kappa = \mathbf{P}(\mathbf{\Pi}_\otimes)^{-1}$$

is the only real constant that makes

$$\sum_{\mathcal{E}} m(\mathcal{E}) = 1.$$

Proof: The first part of the theorem follows at once from the basic combination theorem, the additive properties of a probability, and the obvious fact that the sets

$$\hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2),\ \hat{\mathbf{e}}_1(\mathcal{E}_1') \cap \hat{\mathbf{e}}_2(\mathcal{E}_2')\,, \quad \mathcal{E}_1, \mathcal{E}_1' \in \Phi(\mathcal{S}_1),\ \text{and}\ \mathcal{E}_2, \mathcal{E}_2' \in \Phi(\mathcal{S}_2)\,,$$

are disjoint if $(\mathcal{E}_1, \mathcal{E}_2) \neq (\mathcal{E}_1', \mathcal{E}_2')$.

The second part of the theorem is also an immediate consequence of the proposition proved above, if it is noted that

$$\kappa = P(\mathbf{\Pi}_\otimes)^{-1}\,,$$

and that

$$\widehat{\mathbf{P}}\big(\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)\big) = \sum_{\mathcal{E}} m(\mathcal{E}) = \mathbf{P}(\mathbf{\Pi}_\otimes / \mathbf{\Pi}_\otimes) = 1\,. \blacksquare$$

V

THE COMBINATION OF INDEPENDENT EVIDENCE

This section is devoted to the derivation of Dempster's rule of combination from the bases provided by the additive combination theorem and certain independence assumptions. The nature of these assumptions and their meaning are discussed in terms of other concepts that were introduced in previous sections.

V-1. Probabilistic Independence

V-1-1. Independent Subalgebras

Let Ω_1 and Ω_2 be subalgebras of the subset algebra Ω of a set $\mathbf{X}$. Let $\mathbf{P}$ be a probability on $\mathbf{X}$ with subset algebra Ω. Then the subset algebras Ω_1 and Ω_2 are said to be *independent with respect to the probability* $\mathbf{P}$ if and only if

$$\mathbf{P}(\omega_1 \cap \omega_2) = \mathbf{P}(\omega_1)\mathbf{P}(\omega_2)$$

whenever ω_1 is in Ω_1 and ω_2 is in Ω_2.

V-1-2. Independence Semantics

Independence between two subset algebras with respect to a probability is the formalization of the notion that values of that probability function on one of the subset algebras do not *depend* (in a functional sense) on the probability values on the other. In this case, the conditional probability $\mathbf{P}(\cdot/\omega_1)$, defined on Ω_2 for any member ω_1 of Ω_1 such that $\mathbf{P}(\omega_1) > 0$, is the same as the unconditioned probability $\mathbf{P}$ over Ω_2.

If two subset algebras, Ω_1 and Ω_2, are independent with respect to the probability $\mathbf{P}$, then the values of $\mathbf{P}$ over the smallest subset algebra Ω that contains them will be known once $\mathbf{P}$ is specified separately over Ω_1 and Ω_2. In addition, probability values over one of those algebras may be varied without affecting the values over the other.

Sometimes the values of $\mathbf{P}$ over Ω_1 and Ω_2 are thought of as being the values of two different probability functions $\mathbf{P_1}$ and $\mathbf{P_2}$, which are then said to be independent. This is one of the most frequent characterizations of statistical independence, which we do not use here, as it tends to obscure (by concealing the role of a more encompassing probability function $\mathbf{P}$), the context under which probabilistic independence reduces to functional independence.

V-1-3. The Dempster Combination Formula

Theorem (Dempster [6]): *Let $\widehat{\mathbf{P}}$ be the epistemic probability induced in $\mathcal{U}_\otimes(\mathcal{S}_1 \otimes \mathcal{S}_2)$ by a probability $\mathbf{P}$ defined in the product epistemic algebra $\widehat{\Omega}_{\mathbf{E}}(\mathcal{U}_\times)$ of $\mathcal{U}(\mathcal{S}_1)\times\mathcal{U}(\mathcal{S}_2)$. Assume that $\mathbf{P}(\mathbf{\Pi}_\otimes) > 0$. Assume also that the marginal epistemic algebras $\Omega^1_{\mathbf{E}}(\mathcal{U}_\times)$ and $\Omega^2_{\mathbf{E}}(\mathcal{U}_\times)$ of $\mathcal{U}(\mathcal{S}_1)\times\mathcal{U}(\mathcal{S}_2)$ are independent with respect to* $\mathbf{P}$.

Let m be the probability mass associated with the probability $\widehat{\mathbf{P}}$ in $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$, and m_1, and m_2 be the probability masses associated with $\mathbf{P}$ in the frames of discernment $\Phi(\mathcal{S}_1)$ and $\Phi(\mathcal{S}_2)$, respectively. If $\mathcal{E}$ is in $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$, then

$$m(\mathcal{E}) = \kappa \sum_{\Gamma(\mathcal{E})} m_1(\mathcal{E}_1)m_2(\mathcal{E}_2)\,, \tag{V.1}$$

where the sum is over all pairs $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Phi(S_1) \times \Phi(S_2)$ such that $(\mathcal{E}_1, \mathcal{E}_2)$ is in $\Gamma(\mathcal{E})$, and where κ is the only real constant such that

$$\sum_{\mathcal{E}} m(\mathcal{E}) = 1 .$$

Proof: Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be such that $\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2) \subseteq \mathbf{\Pi}_\otimes$, and let $\kappa = \mathbf{P}(\mathbf{\Pi}_\otimes)^{-1}$. Then, by the hypothesis of independence, it follows that

$$\mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1) \cap \tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big)\mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = m_1(\mathcal{E}_1)\, m_2(\mathcal{E}_2) .$$

The theorem follows at once from this equation and the additive combination theorem.∎

V-1-4. Remark

The above formula is usually given for the case where $S_1 \equiv S_2$. The form given here allows combination of knowledge over different frames of discernment.

V-1-5. Independence as Sensor Independence

In most applications, the underlying probability $\mathbf{P}$ describes the statistical behavior of a group of sensors, or observers of a real world system. Under these conditions, independence of the marginal epistemic algebras can be interpreted as the lack of influence of errors made by one sensing device on the statistical behavior of the other. It is important to note, however, that the sensors (which, after all, are designed to perform accordingly) will be affected by the state of the world.

Finally, it is important to note that using the interpretation of epistemic probability presented in Sections III-3-11 and IV-5-6, the assumption of independence of the evidential bodies is equivalent to the condition

$$\mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1) \cap \tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big)\mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) , \quad \text{if } \mathcal{E}_1 \wedge \mathcal{E}_2 \not\Leftrightarrow \varphi .$$

V-1-6. Combination with Certain Support Functions

Consider now the case where the probability $\mathbf{P}$ is such that

$$m_2(\mathcal{F}) = \mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{F})\big) = 1 ,$$

for some sentence $\mathcal{F}$ in the frame of discernment $\Phi(S_2)$. Then, it is clear that

$$\mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1) \cap \tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = 0 ,$$

if $\mathcal{E}_2 \not\Leftrightarrow \mathcal{F}$.

Furthermore,

$$\mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{F})\big) = \mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{F})\big) + \sum_{\mathcal{E}_2 \not\Leftrightarrow \mathcal{F}} \mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\mathbf{e}_1(\mathcal{E}_1) \times \mathcal{U}(S_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big) .$$

These equations imply that

$$\mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1) \cap \tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big)\mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) .$$

The marginal epistemic algebras are, therefore, always independent whenever one of the corresponding marginal probabilities is focused on a single epistemic set.

When $\mathcal{U}(S_1) \equiv \mathcal{U}(S_2) \equiv \mathcal{U}(S)$, then application of the Dempster combination formula yields the well known *Dempster's Rule of Conditioning*

$$S(\mathcal{E}/\mathcal{F}) = \frac{S(\mathcal{E} \vee \neg\mathcal{F}) - S(\neg\mathcal{F})}{1 - S(\neg\mathcal{F})} .$$

Note also that

$$S(\mathcal{E}/\mathcal{F}) = \mathbf{P}\big(\mathbf{k}(\mathcal{E} \vee \neg\mathcal{F})/\overline{\mathbf{k}(\neg\mathcal{F})}\big) = \frac{\mathbf{P}^*\big(\mathbf{k}(\mathcal{E} \vee \neg\mathcal{F}) \cap \mathbf{t}(\mathcal{F})\big)}{\mathbf{P}^*\big(\mathbf{t}(\mathcal{F})\big)} .$$

VI

SIMPLE FORMS OF DEPENDENT EVIDENCE

This section presents the results of applying the additive combination theorem to the combination of evidential bodies that depend on each other in relatively simple forms described through mappings between frames of discernment that are called *compatibility relations.* Two classes of compatibility relations, corresponding to deterministic and probabilistic relationships, are investigated.

The results presented in this section can be described more accurately as formulas for the *translation* or *projection* of knowledge, rather than as expressions for the integration of distinct evidential bodies. The developments described in this section, however, generalize previous results of the Dempster-Shafer theory, emphasizing the conceptual power of the formal structures presented in previous sections. More importantly, as discussed in detail below, these developments provide insight into the nature of problems entailed in combining evidence under various assumptions of probabilistic dependence. These combination formulas are derived and studied in a related work [15].

VI-1. Compatibility Between Frames of Discernment

VI-1-1. Probabilistic Compatibility Relations

If the probabilistic distribution of the state of knowledge of a rational agent is a function of the state of knowledge of another rational agent, then a probability distribution characterizing the uncertainty of the latter can be used to derive the probability distribution for the states of knowledge of the former.

Theorem: *Let $\widehat{\mathbf{P}}$ be the epistemic probability induced in $\mathcal{U}_{\otimes}(\mathcal{S}_1 \otimes \mathcal{S}_2)$ by a probability $\mathbf{P}$ defined in the product epistemic algebra $\widehat{\Omega}_{\mathbf{E}}(\mathcal{U}_{\times})$ of $\mathcal{U}(\mathcal{S}_1) \times \mathcal{U}(\mathcal{S}_2)$. Assume that $\mathbf{P}(\Pi_{\otimes}) > 0$.*

Further, assume that the conditional mass assignments with respect to an objective sentence $\mathcal{E}_1$ are given by the function

$$\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = \mathbf{P}\big(\tilde{\mathrm{e}}_2(\mathcal{E}_2)/\tilde{\mathrm{e}}_1(\mathcal{E}_1)\big),$$

defined in $\Phi(\mathcal{S}_2)$ for every $\mathcal{E}_1$ in $\Phi(\mathcal{S}_1)$ such that $\mathbf{P}(\tilde{\mathrm{e}}_1(\mathcal{E}_1)) > 0$.

Assume also that $\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = 0$ if $\mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2) \not\subseteq \Pi_{\otimes}$.

Let m and m_1 be the probability masses associated with $\mathbf{P}$ in the subset algebras $\Omega_{\mathbf{E}}(\mathcal{U}_{\otimes})$ and $\Omega^1_{\mathbf{E}}(\mathcal{U}_{\otimes})$, respectively. If $\mathcal{E}$ is in $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$, then

$$m(\mathcal{E}) = \sum_{\Gamma(\mathcal{E})} \tilde{m}(\mathcal{E}_2/\mathcal{E}_1) m_1(\mathcal{E}_1)\,, \tag{VI.1}$$

where the sum is over all pairs $(\mathcal{E}_1, \mathcal{E}_2)$ in $\Phi(\mathcal{S}_1) \times \Phi(\mathcal{S}_2)$ such that $(\mathcal{E}_1, \mathcal{E}_2)$ is in $\Gamma(\mathcal{E})$.

Proof: Since $\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = 0$ whenever $\mathbf{e}_1(\mathcal{E}_1) \times \mathbf{e}_2(\mathcal{E}_2) \not\subseteq \mathbf{\Pi}_\otimes$, and since

$$\sum_{\mathcal{E}_2} \tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)/\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big) = 1,$$

then, if $\sum_{\mathcal{E}_2}^*$ denotes a sum over all $\mathcal{E}_2$ in $\Phi(\mathcal{S}_2)$ such that the pair $(\mathcal{E}_1, \mathcal{E}_2)$ is in $\Gamma(\mathcal{E})$, it is

$$\sum_{\mathcal{E}} m(\mathcal{E}) = \sum_{\mathcal{E}_1} \sum_{\mathcal{E}_2}{}^* \tilde{m}(\mathcal{E}_2/\mathcal{E}_1) m_1(\mathcal{E}_1) = \sum_{\mathcal{E}_1} m_1(\mathcal{E}_1) \sum_{\mathcal{E}_2} \tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = 1.$$

Therefore, $\kappa = \mathbf{P}(\mathbf{\Pi}_\otimes) = 1$, and the theorem follows at once from the additive combination theorem.∎

VI-1-2. Deterministic Compatibility Relations

Assume now that, to each sentence $\mathcal{E}_1$ in the frame of discernment $\Phi(\mathcal{S}_1)$ there corresponds a sentence in the frame of discernment $\Phi(\mathcal{S}_2)$, defined by a mapping

$$\Psi : \Phi(\mathcal{S}_1) \mapsto \Phi(\mathcal{S}_2),$$

which in turn defines a conditional probability $\mathbf{P}\big(\cdot/\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big)$ which is focused on the sentence $\Psi(\mathcal{E}_1)$ in the frame of discernment $\Phi(\mathcal{S}_2)$ with an associated probability mass assignment given by

$$\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = \begin{cases} 1, & \text{if } \mathcal{E}_2 = \Psi(\mathcal{E}_1); \\ 0, & \text{otherwise.} \end{cases}$$

In this case, the formula of the above theorem can be simplified to

$$m(\mathcal{E}) = \sum_{\Gamma(\mathcal{E})} m_1(\mathcal{E}_1), \quad \mathcal{E} \text{ in } \Phi(\mathcal{S}_1 \otimes \mathcal{S}_2),$$

where the sum is over all $\mathcal{E}_1$ in $\mathcal{S}_1$ such that $\big(\mathcal{E}_1, \Psi(\mathcal{E}_1)\big)$ is in $\Gamma(\mathcal{E})$.

If, for a given $\mathcal{E}_2$ in $\Phi(\mathcal{S}_2)$, the left hand sides are now added over all sentences $\mathcal{E}$ in the frame of discernment $\Phi(\mathcal{S}_1 \otimes \mathcal{S}_2)$ such that $(\mathcal{E}_1, \mathcal{E}_2) = \big(\mathcal{E}_1, \Psi(\mathcal{E}_1)\big)$, then the familiar result for the marginal probability mass m_2 is obtained:

$$m_2(\mathcal{E}_2) = \sum_{\Psi^{-1}(\mathcal{E}_2)} m_1(\mathcal{E}_1).$$

This result is usually described as a *translation* formula that allows mapping of probabilistic knowledge defined on one frame of discernment into another frame of discernment related to the former by deterministic compatibility relations.

VI-1-3. Dependence on Support Sets

We will now discuss the combination of dependent knowledge when conditional mass functions defined on one frame of discernment as a function of the *known* truth of a sentence in another frame of discernment have been provided. Unlike previous examples, where it was assumed that dependence information was available in the form of the conditional probability values $\mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)/\tilde{\mathbf{e}}_1(\mathcal{E}_1)\big)$, we shall now assume that information is provided in the form of the conditional probability functions

$$\mathbf{P}\big(\cdot/\tilde{\mathbf{k}}_1(\mathcal{E}_1)\big) : \Omega^2_{\mathrm{E}}(\mathcal{U}_\times) \mapsto [0,1],$$

defined for every marginal epistemic support set $\tilde{\mathbf{k}}_1(\mathcal{E}_1)$, with $\mathcal{E}_1$ in $\Phi(\mathcal{S}_1)$, such that $\mathbf{P}\big(\tilde{\mathbf{k}}_1(\mathcal{E}_1)\big) > 0$.

Note that, if such information is provided, the marginal probabilities for the marginal algebra $\Omega^2_{\mathrm{E}}(\mathcal{U}_\times)$ are known because

$$m_2(\mathcal{E}_2) = \mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2) \,/\, \tilde{\mathbf{k}}_1(\Theta_1)\big).$$

Moreover, any specification of the conditional probability $\mathbf{P}\big(\cdot/\tilde{\mathbf{k}}_1(\mathcal{E}_1)\big)$ *must* be expected to be further constrained, as shown by the following proposition:

Proposition: *Assume that* $\mathbf{P}$ *is a probability defined in the product epistemic algebra of* $\mathcal{U}(S_1)\times\mathcal{U}(S_2)$. *Let*

$$\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = \mathbf{P}\big(\tilde{\mathbf{e}}_2(\mathcal{E}_2)/\tilde{\mathbf{k}}_1(\mathcal{E}_1)\big)\,,$$

with $\mathcal{E}_1$ *in* $\Phi(S_1)$, $\mathcal{E}_2$ *in* $\Phi(S_2)$; *and assume also that* $\mathbf{P}\big(\tilde{\mathbf{k}}_1(\mathcal{E}_1)\big) > 0$. *Further, let* S_1 *be the support function associated with* $\mathbf{P}$ *in the marginal epistemic algebra* $\Omega^1_{\mathbf{E}}$ *of* $\mathcal{U}(S_1)\times\mathcal{U}(S_2)$.

Then, for every $\mathcal{E}_2$ *in* $\Phi(S_2)$, *the function*

$$\tilde{S}_1(\cdot\odot\mathcal{E}_2):\Phi(S_1)\mapsto[0,1]\,,\quad \mathcal{E}_2 \text{ in } \Phi(S_2)\,,$$

defined by

$$\tilde{S}_1(\mathcal{E}_1\odot\mathcal{E}_2) = \mathbf{P}\big(\tilde{\mathbf{k}}_1(\mathcal{E}_1)\cap\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \tilde{m}(\mathcal{E}_2/\mathcal{E}_1)\,S_1(\mathcal{E}_1)\,,$$

satisfies Shafer's axiom III.5.

Proof: The proposition follows immediately from the relation

$$\tilde{\mathbf{k}}_1(\mathcal{E}_1)\cap\tilde{\mathbf{e}}_2(\mathcal{E}_2) = \bigcup_{\mathcal{G}\Rightarrow\mathcal{E}_1}\left[\tilde{\mathbf{e}}_1(\mathcal{G})\cap\tilde{\mathbf{e}}_2(\mathcal{E}_2)\right],$$

where the union is over all $\mathcal{G}$ in $\Phi(S_1)$ such that $\mathcal{G}\Rightarrow\mathcal{E}_1$, by application of the combinatorial theory results leading to the inequality III.5.∎

Corollary:

$$\tilde{m}_1(\mathcal{E}_1\odot\mathcal{E}_2) = \mathbf{P}\big(\tilde{\mathbf{e}}_1(\mathcal{E}_1)\cap\tilde{\mathbf{e}}_2(\mathcal{E}_2)\big) = \sum_{\mathcal{G}\Rightarrow\mathcal{E}_1}(-1)^{|\mathcal{E}_1-\mathcal{G}|}\tilde{S}_1(\mathcal{G}\odot\mathcal{E}_2)\,,$$

where the sum is over all $\mathcal{G}$ *in* $\Phi(S_1)$ *such that* $\mathcal{G}\Rightarrow\mathcal{E}_1$.

Proof: Follows at once from the above proposition by application of the Mōbius inversion.∎

By combining these results with the additive combination theorem, it is possible to derive the main result of this section:

Theorem: *Let the functions* $\tilde{m}(\cdot/\cdot)$ *and* $\tilde{S}_1(\cdot\odot\cdot)$ *be defined as before. Assume also that*

$$\tilde{m}(\mathcal{E}_2/\mathcal{E}_1) = 0,\quad \text{if } \mathbf{e}_1(\mathcal{E}_1)\times\mathbf{e}_2(\mathcal{E}_2)\not\subseteq\mathbf{\Pi}_{\otimes}.$$

If $\mathcal{E}$ *is in* $\Phi(S_1\odot S_2)$, *then it is*

$$\begin{aligned} m(\mathcal{E}) &= \sum_{\Gamma(\mathcal{E})}\tilde{m}_1(\mathcal{E}_1\odot\mathcal{E}_2)\\ &= \sum_{\Gamma(\mathcal{E})}\sum_{\mathcal{G}\Rightarrow\mathcal{E}_1}(-1)^{|\mathcal{E}_1-\mathcal{G}|}\tilde{S}_1(\mathcal{G}\odot\mathcal{E}_2)\\ &= \sum_{\Gamma(\mathcal{E})}\sum_{\mathcal{G}\Rightarrow\mathcal{E}_1}(-1)^{|\mathcal{E}_1-\mathcal{G}|}\tilde{m}(\mathcal{E}_2/\mathcal{G})S_1(\mathcal{G}), \end{aligned}$$

where the sum is over all pairs $(\mathcal{E}_1,\mathcal{E}_2)$ *in* $\Phi(S_1)\times\Phi(S_2)$ *such that* $(\mathcal{E}_1,\mathcal{E}_2)$ *is in* $\mathbf{\Gamma}(\mathcal{E})$ *and over all* $\mathcal{G}$ *in* $\Phi(S_1)$, *such that* $\mathcal{G}\Rightarrow\mathcal{E}_1$.

Proof: Follows immediately from the additive combination theorem and the fact that $\kappa = 1$ (proved as was done before when probabilistic compatibility relations were discussed).∎

VI-2. The Combination of Dependent Evidence

The results presented in this and previous sections illustrate the important role of the results discussed in Section IV (particularly the additive combination theorem) for the derivation of evidence combination rules under a wide variety of assumptions about the nature of the evidential bodies being combined and their relations.

The basic framework in any such problem consists of the specification of probabilities over the marginal epistemic algebras of the Cartesian product $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$. These probability functions will always be consistent, since there will always exist a probability defined over the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$ that extends a probability defined on both $\Omega^1_{\mathrm{E}}(\mathcal{U}_\times)$ and in $\Omega^2_{\mathrm{E}}(\mathcal{U}_\times)$ (for example, that leading to the Dempster combination formula)[6].

In addition to these probabilities, which characterize the uncertainties inherent in each evidential body, other probabilities may be defined over certain subalgebras $\widetilde{\Omega}$ of the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$. For example, in the case of the Dempster formula, a probability is defined directly over that algebra by means of the independence assumption (i.e., defining a value for $\mathbf{P}(\mathrm{e}_1(\mathcal{E}_1) \times \mathrm{e}_2(\mathcal{E}_2))$ as a function of the values of the marginal probabilities). Similarly, in the cases studied in this section, the conditional values $\tilde{m}(\mathcal{E}_2/\mathcal{E}_1)$ specify (together with the mass assignment m_1) probabilities over members of $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$ in a straightforward manner.

The probabilities specified over $\widetilde{\Omega}$ must be consistent, however, with the marginal probabilities defined over the marginal epistemic algebras Ω^1_{E} and Ω^2_{E}. If the constraining probabilities are consistent with the marginal distributions, the problem of combination reduces to that of extending $\mathbf{P}$ (defined over the marginal epistemic algebras and $\widetilde{\Omega}$) to the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}$, thus allowing application of the results of Section IV. As seen in Section III, this is equivalent the computation of the lower probabilities induced by $\mathbf{P}$.

For example, in an important[7] and more complex problem involving three mutually independent bodies of evidence, probabilities are specified in three marginal epistemic algebras of the Cartesian product

$$\mathcal{U}_\times = \mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_3).$$

These probabilities represent the results $\mathbf{P}_{12}$ and $\mathbf{P}_{13}$ of combining one body of evidence with each of the other two (corresponding to probability specification over the marginal algebras $\Omega^{12}_{\mathrm{E}}(\mathcal{U}_\times)$ and $\Omega^{13}_{\mathrm{E}}(\mathcal{U}_\times)$, defined in a way that naturally extends the definitions of Section IV) plus actual knowledge of the probability over the marginal epistemic algebra $\Omega^1_{\mathrm{E}}(\mathcal{U}_\times)$. These interrelated, consistent, probability functions must be combined so as to extend $\mathbf{P}$ to the product epistemic algebra $\widehat{\Omega}_{\mathrm{E}}(\mathcal{U}_\times)$, thus allowing application of the results of Section IV.

A detailed treatment of these problems is given in a related paper [15].

[6] Note, however, that conditioning with respect to the subset $\mathbf{\Pi}_\otimes$ of $\mathcal{U}(S_1) \times \mathcal{U}(S_2)$ may be impossible as $\mathbf{P}(\mathbf{\Pi}_\otimes)$ may be equal to zero indicating inconsistence of the bodies of evidence.

[7] The importance of this problem derives from its practical application to the combination of evidential bodies sharing common knowledge.

REFERENCES

[1] Hintikka,J., *Knowledge and Belief,* Cornell University Press, Ithaca, New York,1962.

[2] Moore, R., *Reasoning about Knowledge and Action,* Technical Note 191, SRI International, Menlo Park, California, 1980.

[3] Carnap, R., *Meaning and Necessity,* University of Chicago Press, Chicago, Illinois, 1956.

[4] Carnap, R., *Logical Foundations of Probability,* University of Chicago Press, Chicago, Illinois, 1962.

[5] Suppes, P., *The Measurement of Belief,* Jrnl. Royal Statist. Soc., Series B, **36**, 160-191, 1974.

[6] Dempster, A., *Upper and Lower Probabilities Induced by a Multivalued Mapping,* Annals of Mathematical Statistics, **38**, 325–339, 1967.

[7] Shafer, G., *A Mathematical Theory of Evidence,* Princeton University Press, Princeton, New Jersey, 1976.

[8] Hall, M. Jr, *Combinatorial Theory,* Second Edition, John Wiley and Sons, 1986.

[9] Neveu, J., *Bases Mathématiques du Calcul des Probabilités,* Masson, Paris, France, 1964.

[10] Halmos, P., *Measure Theory,* Springer-Verlag, New York, New York, 1974.

[11] Quine, W.V.O., *Mathematical Logic,* Harvard University Press, Cambridge, Massachusetts, 1983.

[12] Rosenschein, S.J. and L.P. Kaelbling, *The Synthesis of Digital Machines with Provable Epistemic Properties,* Proc. 1986 Conference on Theoretical Aspects of Reasoning about Knowledge, 83–98, W. Kaufmann, Los Altos, California, 1986.

[13] Savage, L.J., *The Foundations of Statistics,* John Wiley and Sons, New York, 1954.

[14] Nilsson, N., *Probabilistic Logic,* Artificial Intelligence, **28**, 71–88, 1986.

[15] Ruspini, E.H.,*The Computational Treatment of Dependent Evidence,* in preparation.

[16] Humez, P.A. and N.D. Humez, *Alpha to Omega: The Life and Times of the Greek Alphabet,* Godine, Boston, Massachusetts, 1981.

[17] Hoad, T.F. (editor), *The Concise Oxford Dictionary of English Etymology,* Oxford University Press, Oxford, England, 1986.

[18] Rota, G.C., *On the Foundations of Combinatorial Theory, I, Theory of Möbius Functions,* Z. Wahrscheinlichkeitstheorie und. Verw. Gebiete, **2**, 340-368, 1964.

[19] Kyburg, H.E., *Bayesian and Non-Bayesian Evidential Updating,* private communication, 1985.