A Decision-Aid For Nodes in Command and Control Systems Based on Cognitive Probability Logic

Dr. I.R. Goodman *
Code D44215, SPAWARSYSCEN,
San Diego, CA 92152-7446
Tel: 619-553-4014
Email: goodman@spawar.navy.mil

Abstract

The principles underlying this paper can be applied not only to C2I systems, but also to many other complex structures, such as those involving medical, fault or configuration diagnoses and analyses.

In short, a new relatively simple look-up table of formulas is presented which can be used as a practical aid in C2 decision-making nodes for deducing conditional or unconditional conclusions from (conditional or unconditional) premises in probability form. This results from a recent breakthrough yielding a new *Cognitive Probability Logic* -- or *Logic of Averages*. This logic is actually a natural weighting modification of Adams' well-known *High Probability Logic*. Consequently, a number of long-standing conflicts between ordinary probability logic and "commonsense" reasoning are resolved for the first time, including the well-known *transitivity-syllogism problem*. These results are based upon completely rigorous universal *second order* probability principles, together with use of the newly emerging field of *product space conditional event algebra*. Surprisingly, both disciplines are actually technically entirely within the purview of classical logic and basic probability theory. Applications to linguistic-based information can also be obtained by use of these techniques, together with one-point random set coverage representations of fuzzy logic and an extension of product space conditional event algebra, dubbed *relational event algebra.*

1. Introduction

1.1 *Overall Objectives*

The chief thrust of this paper is to provide the basis for a new decision-aid in the form of a relatively simple look-up table of averaged probability values of conditional or unconditional potential deductions (or conclusions) that a C2 node decision maker may consider, given a finite set of conditional or unconditional premises. However, as in many endeavors, it is equally important to present both the skeletal derivation of certain of these results, and provide a background on the controversy and progress involved in this area. Detailed proofs, as well as additional important results, will be given in [Goodman & Nguyen, b].

---

*This work was supported by the Office of Naval Research (ONR) In-House Laboratory Independent Research Program at SSC-SD, under FY99 IR Project No. ZU58.

Among the key components in any C2 (or actually C4I) system are its decision nodes operating on incoming information, which make interactive decisions with other nodes. These take into account the distributive and hierarchical nature of the overall system, as well as all relevant geographical -- and even political -- constraints present. Such incoming information may typically arise from sensor-based sources, with uncertainties measured through probability considerations or from human-based intelligence sources, possibly in linguistic form and modeled through some combination of fuzzy logic or probability, for example. It is obvious that all such node-based decision making is crucially dependent upon the initial data screening and data fusion aspects. In particular, it is essential to be able to have a reasonably efficient, yet mathematically / logically sound, method for dealing with deductions resulting from probability-based premises.

1.2 *The Basic Issue*

In spite of the above remarks, surprisingly at first glance, *there exists to this day, a large class of simple-appearing probability-based premises for which commonsense or intuitive analysis, as well as classical logic, yields deductions quite different from those obtained via the application of standard probability techniques.* One fundamental example of this will be given in Section 5.

This paper provides the beginning of a remedy to the above difficulty, utilizing, in tandem, two newly developed mathematical tools – *product space conditional event algebra* (PSCEA) and the use of *second order probabilities* (SOP). In addition, these results can be further extended to encompass linguistic-based information via two additional tools: *one-point random set coverage representation of first order fuzzy logic* (RSC)-- in conjunction with fuzzy logic (FL), classical logic (CL), and traditional probability theory (PL) -- and *relational event algebra* (REA). All of these tools enhance the scope of applicability of FL, CL, and PL to many classes of real-world problems, previously either ignored or only superficially treated. Because of space limitations, for completeness only the essential aspects of RSC and REA are provided. A basic reference, detailing use of RSC and REA in treating linguistic-based deductions, is detailed in [Goodman & Nguyen, a].

2. *Notation and Some Basic Relations*

This paper depends in a key way on the fundamental identification, beginning with the Stone Representation Theorem, of propositional calculus – and in fact, aspects of its further development into first and higher order predicate calculus, all part of CL – with (concrete) boolean algebras of events or sets and their metalevel extensions. For background on the basic theory of CL (as well as even a separate Chapter on basic PL) see, e.g., [Copi, 1986]; for applications, see, in addition to the Copi reference, [Bergmann *et al*., 1980], [Oesterle, 1963], and [Neblett, 1985]. (The last reference, combines use of CL with induction and empirical approaches, but avoids probability.) For boolean algebra(s) and the Stone Theorem, the reader should consult, e.g., [Mendelson, 1970]. Although the identification between CL and boolean algebra concepts has been known for a long time, the two areas for the most part have developed separately – with separate notation and concepts. Thus, in this work, references at times will be relative to the literature of CL, while at other times to boolean algebra and related disciplines.

Throughout this work we employ certain symbols to indicate generic and special events (or sets or propositions): lower case letters near the beginning of the roman alphabet, with or without subscripts, such as a, a_1, b, b_2, c, d, for generic events; $\varnothing$ for the null event; Ω for the universal event (a relative usage). In addition, we use particular symbols to indicate specially designated boolean operators and relations, including: &, or absence of any symbol when no ambiguity arises, for conjunction (or intersection or "and-ing") of events; $\vee$ for disjunction (or union or "or-ing") of events; prime as in a' for negation (or complement) of event a relative to Ω. More generally, the event difference between say a and b is indicated by a-b, which is the same as ab', i.e., $a\text{-}b = ab'$, whence we also have $a' = \Omega\text{-}a$. In addition, event inclusion or partial ordering is indicated by the same standard symbol for numerical partial ordering $\leq$. So, $a \leq b$ is the relation that a is a subevent (or subset) of b. Thus, we always have $\varnothing \leq a \leq \Omega$. Also, $a < b$ indicates a is a proper subevent of b, i.e., $a \leq b$, but not $(a = b)$, i.e., $a \neq b$. On the other hand, when dealing with classes or collections of events – usually indicated by capital letters (which may or may not be italicized to indicate some additional property -- see below) near the beginning of the alphabet, with or without subscripts – such as $A_2 = \{a, b, c_4\}$ or $B = \{b_1, b_2,\dots\}$, we employ the notations $A_2 \cap B$, $A_2 \cup B$, to indicate, the classes resulting from the intersection and union of the classes A_2 and B, respectively. If, say, class $A = \{c, d, \dots\}$ is a subclass of class $C = (a, b, c, d,\dots\}$, we write $A \subseteq C$. If the relation is proper, then analogous to the event level, we write $A \subset C$, etc. The italicized capital letter *B* – with or without subscripts -- is reserved for either a boolean algebra or sigma algebra of events. (Recall that the former is a class of events closed under all finite applications of conjunctions, disjunctions and negation, while the latter is similarly closed under both finite and countable infinite applications of such operators.)

We indicate a probability space by the triple (Ω,B,P), where Ω (as well as $\varnothing$) are in *B* and, using functional notation, $P:B\rightarrow[0,1]$ is a probability measure over *B*. P (or P(.)) may or may not have subscripts, depending on the situation and $P(a|b) = P(ab)/P(b)$ indicates, as usual, the conditional probability of a given b, provided $P(b) > 0$. In addition, we reserve the double arrow symbol $\Rightarrow$ to indicate the binary boolean operator known as the *material conditional*, whereby for any events a, b (usually in some common boolean algebra *B*),

$$b\Rightarrow a = b' \vee a = b' \vee ab, \tag{2.1}$$

(the second equality holding by the law of absorption). The material conditional is one basic way classical logic interprets the natural language conditional expression "if b, then a". For background on the use of the material conditional operator – as well as a number of "paradoxes" connected with it, see, e.g., [Copi, 1986] and [Bergmann *et al.*, 1980]. It follows immediately from eq.(2.1) that the probability evaluation of the material conditional is

$$P(b\Rightarrow a) = P(b') + P(ab) = 1\text{-}P(b) + P(ab) = 1\text{-}P(a'b). \tag{2.2}$$

Recall that one fundamental connection among conditional probability, the probability of the material conditional and subevent ordering is that for any a, b in Ω with $P(b) > 0$,

$$P(a|b) = 1 \quad \text{iff} \quad P(a'b) = 0 \;\; \text{iff} \;\; P(b \leq a) = 1 \text{ (slightly abusing notation) iff } \; P(b\Rightarrow a) = 1. \tag{2.3}$$

One other important relation among the material conditional, conditional probability, and conjunction is provided in the easily proven (but not that well-known) relation [Goodman *et al.*, 1997]

$$P(b\Rightarrow a) = P(a|b) + P(a'|b)P(b') \geq P(a|b) = P(ab) + P(a|b)P(b') \geq P(ab), \qquad (2.4)$$

where it is assumed that $P(b) > 0$. Strict inequality holds in general, in place of the right and left $\geq$ relations. Equality holds on the left side of eq.(2.4) iff $[P(b) = 1$ or $P(a|b) = 1]$. Equality holds on the right side iff $[P(b) = 1$ or $P(ab) = 0]$.

Two other special binary boolean operators should be mentioned: the boolean *symmetric sum* (or difference) Δ, defined as

$$a\Delta b = a'b \vee ab', \qquad (2.5)$$

with probability evaluation

$$P(a\Delta b) = P(a'b) + P(ab') = P(a) + P(b) - 2P(ab), \qquad (2.6i)$$

whence

$$P(a\Delta b \mid a\vee b) = (P(a) + P(b) - 2P(ab)) / (P(a) + P(b) - P(ab)), \qquad (2.6ii)$$

etc. The *double conditional* or *biconditional* $\Leftrightarrow$ defined as

$$a\Leftrightarrow b = (b\Rightarrow a) \,\&\, (a\Rightarrow b) = ab \vee a'b' = (a\Delta b)', \qquad (2.7)$$

with probability evaluation

$$P(a\Leftrightarrow b) = P(ab) + P(a'b') = 1-P(a\Delta b) = 1-P(a)-P(b) +2P(ab). \qquad (2.8)$$

We usually make the identifications

$$[P(a = b) = 1] \text{ iff } [P(a\leq b) = P(b\leq a) = 1] \text{ iff } P(a\Leftrightarrow b) = 1 . \qquad (2.9)$$

The evaluations $P(a\Delta b)$ and $P(a\Delta b \mid a \vee b)$ as functions of events a and b in B are legitimate (pseudo)metrics over B. See [Kappos, 1969, Chapter 1] for use of the first; see [Goodman *et al.*, 1997, Part III] for use of both metrics and additional background and applications to determining measures of similarity, as well as testing of hypotheses of similarity of events. Cartesian product of events is indicated by $\times$, with the superscript j as in $b^j\times c$ meaning the iterated cartesian product $b\times\ldots\times b\times c$ (j b-factors) with $b^0\times c$ interpreted as c. In addition, the following multivariate notation will be employed throughout: Finite (nonvacuous) index sets will be usually denoted as J, K, or L, with or without subscripts. Given a collection of events a_j in B, for given boolean algebra B, j in J: a_J denotes the family of events $(a_j)_{j \text{ in } J}$; $\&(a_J)$ denotes $\underset{j \text{ in } J}{\&}(a_j)$; $\vee(a_J)$ denotes $\underset{j \text{ in } J}{\vee}(a_j)$; $\&(b')_J$ denotes $\underset{j \text{ in } J}{\&}(b_j')$; $(ab)_J$ denotes $(a_jb_j)_{j \text{ in } J}$; $(ab, b)_J$ denotes $(a_jb_j, b_j)_{j \text{ in } J}$, etc. More generally, any finite well-defined combination of boolean (or CL) operators &, $\vee$, (.)′, applicable to *any* collection of events a_J indexed by J, is denoted generically as comb(&,v,(.)′;J) (not depending upon any particular a_J) with corresponding event formation comb(&,v,(.)′;J)(a_J) , or more simply, comb(&,v,(.)′)(a_J) in B. In a similar vein, given a_j, b_j in B, j in J, $\&(b\Rightarrow a)_J$ denotes $\underset{j \text{ in } J}{\&}(b_j\Rightarrow a_j)$, $\times(a_J)$ denotes $\underset{j \text{ in } J}{\times}(a_j)$, $a_J \leq b_J$ denotes $(a_j \leq b_j$, j in J), etc.

Other special symbols will be introduced as needed.

3. *Executive Summary of the Key Mathematical Tools*

First, note that FL, beginning with [Zadeh, 1965, 1975], has grown to become an important alternative to the use of PL in modeling many real-world situations, being especially valuable in capturing the essence of linguistic-based information (essentially initiated in [Zadeh, 1978a,b]). While Zadeh originally considered only the fuzzy logic conjunction, disjunction operator pair (min, max), later he and others realized the worth of the alternative operator pair (prod, probsum), and other classes of operator pairs, as well, where "prod" is ordinary product and probsum is the DeMorgan transform of prod. (See, e.g., [Goodman & Nguyen, 1985, Chapter 2] for more details.)

3.1 *PSCEA*

3.1.1 *Introduction to CEA*

Given any probability space $(\Omega,\mathcal{B},P)$, with P arbitrarily variable, conditional event algebra (CEA) is an attempt at providing (1) a rigorous and systematic interpretation of the natural language conditional statements "if b, then a", denoted for simplicity as (a|b), with a denoted as the *consequent* and b as the *antecedent* of (a|b), for any a, b in $\mathcal{B}$ and (2) obtaining the probability of any finite boolean operator combination of such conditional statements, when the statements individually are naturally evaluated as corresponding conditional probabilities. Certainly, if the probabilities of the conditional statements are not interpreted as conditional probabilities, but as the probabilities of the material conditional, then one can readily obtain – at least theoretically – any desired finite combination of boolean operators, such as conjunctions, disjunctions or negations. This is because all such operations applied to material conditional forms in $\mathcal{B}$ remain in $\mathcal{B}$ and are determined and/or simplified according to the usual laws of boolean algebra. In turn, one can then evaluate the probability of any such combination – again, at least theoretically, if one knows all required contributing probabilities of its components. However, inspection of eqs. (2.2) and (2.4) easily shows that a great discrepancy exists between the probability of the material conditional and the corresponding conditional probability evaluation, whenever the antecedent has a small probability value. In fact, for this range of antecedent probability values, the value of the overall material conditional form $P(b\Rightarrow a)$ is close to unity, and hence also close to its "opposite" conclusion probability $P(b\Rightarrow a')$. On the other hand, no matter how small the value of the probability of the antecedent, one always has the simple – but critical – relation

$$P(a'|b) = 1\text{-}P(a|b) \tag{3.1}$$

which nontrivially depends on the ratio of P(ab) to P(b). Other, more logical / algebraic objections (or "paradoxes") to the use of the material conditional operator can be found, e.g., in [Copi, 1986], though if one goes no further than employing the material conditional in a logical context only – with no probability evaluations – it appears as a reasonable interpretation of conditioning.

In a natural sense, no such issue arises when the antecedents of all events for a given problem are restricted to be identical, such as symbolized as (a|b), (c|b), (d|b),…, since in that case one can replace the original probability space by the conditional probability – or trace -- space (Ω, B, P_b), or actually equivalently, (b, Bb,P_b), assuming P(b) > 0, where

$$Bb = \{cb: c \text{ in } B\},\ P_b = P(.\,|b) \text{ (conditional probability for b fixed).} \qquad (3.2)$$

Then, all issues involving (a|b), (c|b), (d|b),…. with respect to P reduce essentially to corresponding ones involving ab, cb, db,… with respect to P_b . On the other hand, often in real-world problems, the antecedents of conditional expressions *change* within a given problem. Of course, if, e.g., they change in a particular patterned way, we may not need to introduce any new concepts, such as in the computation of P(ac | b) from knowledge of P(a | bc) and P(b|c) because of the chaining relation

$$P(ac \mid b) = P(a \mid bc)\cdot P(c|b). \qquad (3.3)$$

However, often the antecedents of conditional expressions in a given problem differ in a way that cannot be treated by purely numerical techniques, such as the use of chaining in eq.(3.2) or by related properties of conditional probabilities. For example, one basic motivation for attempting to develop CEA (in general, many such can exist) is to be able to compare, in a universally systematic and quantitative manner, similarity and differences of inference rules whose validity is interpreted via naturally associated conditional probabilities. This is because of the following analogy extended to conditional probabilities: Given a and b in B, a basic way to compare them for similarity in terms of P is via any of the probability metrics $P(.\Delta..)$ or $P(.\Delta..| \ . \vee..)$ described in eq.(2.6). Note, however, that the full evaluation of these metrics depends on not only knowing the individual probabilities P(a), P(b), but also in a critical way, knowing the conjunctive probability P(ab). Even for this simple case, where a, b, ab all lie in B, if ab is not available, or even if ab is known, but P(ab) is not obtainable, then the metrics cannot be completely determined. All of the above is applicable to the situation where a and b are themselves actually some finite combination of other events e_1, e_2, e_3, f_1, f_2, f_3 in B, such as $a = e_1 \vee e_2$, $b = f_1 \vee f_2$, $c = e_1 \vee e_3$, $b = f_1 \vee f_3$, etc.

Now, consider two inference rules, say,

R_1 = “if it snows or temperature is below 30 degr.F, then enemy will attack in Sectors A or B”,
R_2 = “if it snows or temperature is between 20 and 45 degr.F, then enemy will attack in Sectors B or C”. (3.4)

Denoting the relevant events in B, part of probability space (Ω,B,P),

f_1 = “it snows”, f_2= “temp. ≤ 30 degr.F”, f_3= temp. is between 20 and 45 degr. F,
e_1 = “enemy attacks in Sector A”, e_2 = “enemy attacks in Sector B”,
e_3 = “enemy attacks in Sector C”, (3.5)

then

$$a = e_1 \vee e_2\,, \quad b = f_1 \vee f_2\,, \quad c = e_2 \vee e_3\,, \quad d = f_1 \vee f_3\,, \qquad (3.6)$$

and it is reasonable to interpret symbolically

$$R_1 = (a \mid b)\,, \quad R_2 = (c|d). \qquad (3.7)$$

Furthermore, assume we can obtain the evaluation of P for all of the individual events e_i, f_j, as well as for some some collection of boolean operators among them (conjunctions, disjunctions, negations, and certain combinations of such), with the conditional probability compatibility conditions

$$P(R_1) = P(a|b)\ ,\ \ P(R_2) = P(c|d). \tag{3.8}$$

This may possibly abuse notation, in that R_1, R_2 may very well lie in some space B_o properly containing B and require extending the original P to some P_o. Suppose we can extend eq.(2.6) somehow in a natural sense so that the ordinary events a and b there are replaced by inference rules (a|b), (c|d), we must to be able to compute not only P(a|b), P(c|d), but also P((a|b)&(c|d)) -- or, if need be, $P_o((a|b)\&_o(c|d))$, where $\&_o$ is some extension of ordinary conjunction to B_o. In order to compute such conjunctive probability computations, it is hoped that B_o will be such that $(a|b)\&_o(c|d)$ in B_o is sufficiently simple that $P_o((a|b)\&_o(c|d))$ requires only a finite (and hopefully, relatively small) number of computations involving P applied to certain combinations of boolean operators applied to the antecedents and consequents a,b,c,d – and hence the e_i and f_j -- in *B*.

Happily, such CEA indeed exist and allow us to address the above – and many other types – of problems. (See [Goodman *et. al*., 1999] for this, as well as a number of other motivating issues behind the use of CEA. See [Goodman et al., 1997, Part III] for more details on not only the practical implementation of the above problem on comparison of inference rules, but when a form of CEA (PSCEA) is employed in conjunction with the use of SOP (see Sections 3.1.2 and 3.4 for some background on SOP as utilized here) an extension to actual formal testing of hypotheses for sameness vs. distinctness can be achieved.

More formally, a CEA is a pair of mappings ((.|..),τ), where for any (*B*,P) as above, we denote for simplicity, $\tau(\mathit{B},P) = (\mathit{B}_o,P_o)$ (with B_o not dependent on any P or P_o) and any a,b,…, a_j,b_j,... in *B*, *conditional event* (a|b) in B_o , where, in general the collection of all *first orde*r conditional events {(a|b): a, b in *B*} may or may not be a proper subclass of B_o, such that:

(1) B_o is a set with an algebraic structure (not necessarily itself boolean) which has formal counterparts of all of the boolean (or CL) operators &, ∨, (.)′,… over *B*, which for convenience are also denoted by the same symbols -- unless there is a problem of ambiguity, in which case appropriate subscripts are attached. All *conditional events* (a|b) lie in B_o, with (a|b) identifiable with (ab |b).

(2) The relation a ↔ (a|Ω), for all a in *B*, is an isomorphic imbedding of *B* into B_o.

(3) For each choice of P, there is an extension $P_o{:}\mathit{B}_o \to [0,1]$ such that the compatibility relation holds:

$$P_o((a|b)) = P(a|b),\ \text{for all P, provided } P(b) > 0. \tag{3.9}$$

Thus, assumptions (1)-(3) imply that, for any finite well-defined combination of boolean operators &, ∨, (.)′, over *B*, and hence comb(&, ∨, (.)′;J) over *B*, the operator counterparts over B_o when applied to a finite collection of conditional events $(a|b)_J = (a_j|b_j)_{j \text{ in } J}$, for any a_j, b_j in *B* , and hence $(a_j|b_j)$ in B_o, j in J, have well-defined counterparts, comb(&, ∨, (.)′)($(a|b)_J$), in B_o. In turn, P_o(comb(&, ∨, (.)′)($(a|b)_J$)) can then be fully evaluated. In conjunction with this, a reasonable additional property to require any CEA to satisfy is that

(4) For any J and any comb(&, ∨, (.)′; J)), there exist finite index sets L, K_i and operator combinations comb(&, ∨, (.)′;K_i) , i in L, and finite well-defined combination of any or all of the four basic arithmetic operators, G(·,+,-,·/..;L), with all of these quantities,

in general, dependent upon comb(&, $\vee$, (.)′; J)), such that for any finite collection of first order conditional events $(a|b)_J$ as above,

$$P_o(\text{comb}(\&, \vee, (.)')((a|b)_J)) = G(.,+,-,\cdot/..;L)\Big(\big(P\big(\text{comb}(\&, \vee, (.)')((ab,b)_{K_i})\big)\big)_{i \text{ in } L}\Big). \qquad (3.10)$$

That is, we can obtain the evaluation of the probability extension of any finite boolean operator combination of conditional events as a finite combination of arithmetic operators acting upon finite boolean combinations of ordinary events consisting of the antecedents and consequents of the conditional events in question. Obviously, it is desirable not only for (4) to hold, for any given J, but for index sets L and K_i, i in L, to be as small as possible and for G to be as simple as possible.

3.1.2 *PSCEA and Related CEA's*

PSCEA is a further development of earlier non-boolean-structured algebras or logics of conditionals that are compatible with corresponding conditional probability evaluations– also called conditional event algebras (CEA's). These non-boolean CEA's are all related isomophically to some corresponding three-valued logic. Conversely, a fundamental theorem shows that any three-valued logic is isomorphic to some corresponding CEA [Goodman et al., 1991b]).

First, and foremost, Adams' partially developed a CEA [Adams, 1975] in the sense that he did not specify what the underlying space was that contained the conditional events, but did specify the negation, conjunction-like and disjunction-like operators (as well as other ones, not considered here). The purpose of introducing, in effect, this CEA by Adams was to characterize in a complete algebraic sense his *high probability* deduction scheme of which we will elaborate on later (see Section 4 *et passim*). Independently, some years later, Calabrese [Calabrese, 1987, 1994] considered the same operators, as Adams, but (unlike Adams) developed a full CEA with explicit interpretations for the conditional events involved, as well as postulating that all second order conditional events – i.e., conditional events with antecedents and consequents themselves being conditional events -- could be identified as first order conditional events. The non-boolean conjunction-like or (as Adams has dubbed it) "quasi-conjunction" operator $\&_{AC}$ and its DeMorgan dual $\vee_{AC}$ (or "quasi-disjunction") operator play a crucial role in deduction, both possessing certain similarities, yet critical differences, with respect to the corresponding natural boolean conjunction and disjunction operators of PSCEA. $\&_{AC}$ and $\vee_{AC}$ -- together with the universal negation operator valid for essentially any (non-intuitionistic) three-valued logic of interest --

$$(a|b)' = (a'|b) = (a'b \mid b)\ , \text{ all a, b in } B, \qquad (3.11)$$

correspond isomorphically to Sobocinski's three-valued logic [Rescher, 1969].

Another non-boolean approach to CEA based initially upon the interpretation of conditional events as cosets in a union of quotient boolean algebras – or equivalently, as appropriately determined intervals of events with corresponding operators being functional image extensions of the ordinary boolean operators acting upon unconditional events – is provided in [Goodman et al., 1991b, 1991c], the special issue *of IEEE Transactions on Systems, Man & Cybernetics* [Dubois et al., 1994]. Apropos to this CEA, we use the subscript DGNW (apropos to the originators of this CEA, including initially B. DeFinetti in the 1930's who first realized its isomorphism with a corresponding fragment of Lukasiewicz' three-valued min, max logic (again, see [Goodman et al., 1991b] for a history of this CEA). We will see that a basic imbedding of both the AC and DGNW CEA's into PSCEA plays a natural role in probabilistic deduction.

PSCEA – unlike its predecessors -- now permits us to represent any simple or compound conditional expression in a legitimate boolean (or sigma) algebra of conditional events. In general, this boolean algebra and associated probability space do not coincide with the boolean algebra and probability space in which the initial (relative) unconditional events making up the antecedents and consequents of the conditional expressions lie. In fact, the resulting probability space containing the desired conditional events is the countable infinite cartesian product probability space formed from the original probability space connected with the unconditional events, all of whose factor or individual marginal spaces are identical to the initial probability space, a standard type of construction in probability theory [Neveu, 1965, Section III.3 *et passim*]. Moreover, a well-known result (Lewis' Triviality Theorem – see [Lewis, 1976], as well as later results in [Eells & Skyrms, 1994]) precludes, in general, the spaces from being the same. The conditional events themselves in the product probability space are simply the algebraic counterparts of the power series expansions of arithmetic divisions, with complements playing the role of subtraction, disjoint disjunctions the role of addition, and cartesian products playing the role of products and (when reiterated) powers. Also, this space contains natural isomorphic probability-preserving imbeds of all of the original unconditional events. Independent of the investigations of Goodman & Nguyen, earlier Van Fraasen [Van Fraasen, 1976] derived preliminary results for PSCEA. Later, independent of Van Fraasen, Goodman, and Nguyen, McGee, via a game-theoretic approach (without obtaining the space and events explicitly) obtained the PSCEA conjunction operator (see [McGee, 1989] and an extension of his technique in [Goodman & Nguyen, 1995]).

In turn, *all* of the standard laws of boolean algebra and probability can then be applied to such conditional events – much as standard probability has been applied to unconditional events and logical operators and relations among them – but with an additional number of calculations involved that can grow exponentially as the number of arguments increase. Nevertheless, all finite logical combinations of conditional events of PSCEA can be put in finite "closed-computable" form. (For expositions on PSCEA, see, e.g., [Goodman & Nguyen, 1995], [Goodman *et al.*, 1997, Part III], and [Goodman & Kramer, 1997].)

One basic application of PSCEA– as utilized in this paper -- is its use as a rigorous basis for uniting, for the first time, classical deductive logic, commonsense reasoning, and probability logic (see Section 6). Another use of PSCEA (see also the example at the beginning of Section 3.1.1) is to compare in a universal quantitative manner similarity and differences of inference rules whose validity is interpreted via naturally associated conditional probabilities. This also makes use of the tool SOP to be explained below and the fact that all probability spaces can be made into (pseudo-)metric spaces using relatively simple unconditional and conditional probabilities involving the boolean symmetric sum operator – as considered, e.g., in [Kappos, 1969, Chapter 1]. For additional background, history, and applications of PSCEA, see, e.g., [Goodman & Kramer, 1997], [Goodman *et al.*, 1997, Part III], [Goodman, 1998a], and [Goodman & Nguyen, 1998].

Summarizing here only the barest properties, for any a, b, c, d,… in B, for given probability space (Ω,B,P) and countable infinite product probability space derived from it, (Ω_o,B_o,P_o), where

$$\Omega_o = \Omega\times\Omega\times... , \tag{3.12}$$

conditional events (a|b), (c|d),… in B_o, where, e.g., (a|b) is defined directly and recursively and, provided P(b) > 0, evaluated as

$$(a|b) = (ab \mid b) = \bigvee_{j=0}^{+\infty} (b')^j \times ab \times \Omega_o = (ab \times \Omega_o) \vee (b'\times(a|b)) ; \; P_o((a|b)) = \sum_{j=0}^{+\infty} (P(b'))^j P(ab) = P(a|b). \tag{3.13}$$

We call any conditional event (a|b) *nontrivial* (or proper) iff $\varnothing$ < a < b, where b may or may not be Ω. If b = Ω, there is the natural identification (isomorphic probability preserving imbedding) of unconditional event a in B with (a|Ω) in B_o. But, again, note that they are not identical and Lewis' Theorem is respected. (In fact, this issue is discussed in detail in [Goodman *et al.*, 1997], p. 395.) Then, for any P and a, b, c, d in B, with P(b) > 0, where required, indicating all boolean operators and relations extending the usual ones for (Ω,B,P) to (Ω_o,B_o,P_o) by the same symbols, when unambiguous,

$$a \leftrightarrow (a|\Omega) = a\times\Omega_o , \quad P(a) = P_o((a|\Omega)). \tag{3.14}$$

$$(a|b) = \varnothing_o = (\varnothing \mid b) \text{ iff } ab = \varnothing ; \quad (a|b) = \Omega_o = (b|b) = (\Omega|\Omega) \text{ iff } a \geq b > \varnothing. \tag{3.15}$$

$$(a|b)' = (a'|b) = (a'b \mid b) , \; P_o((a|b)') = 1- P(a|b) = P(a'|b) , \tag{3.16}$$

$$(a|b)\&(c|d), \; (a|b) \vee (c|d) \text{ in } B_o , \tag{3.17}$$

$$P_o((a|b)\&(c|d)) = [P(abcd) + (P(abd')P(c|d)) + (P(cdb')P(a|b)] / P(b \vee d) , \tag{3.18}$$

$$P_o((a|b) \vee (c|d)) = P(a|b) + P(c|d) - P_o((a|b)\&(c|d)), \tag{3.19}$$

$$P_o((a|b) \mid (c|d)) = P_o((a|b)\&(c|d)) / P(c|d) , \text{ etc.} \tag{3.20}$$

If a, b, c, d in B arbitrary such that (a|b), (c|d) are nontrivial, then:

$$\begin{aligned}(a|b) \leq (c|d) \;\; &\text{iff} \;\; (a|b)\&(c|d)' = \varnothing \;\; \text{iff} \; [(a|b) = (a|b)\&(c|d)] \;\; \text{iff} \; (c|d) = (a|b) \vee (c|d) \\ &\text{iff} \;\; [ab \leq cd \;\; \text{and} \;\; b\Rightarrow a \leq d\Rightarrow c] \; \text{iff} \; [ab \leq cd \;\; \text{and} \;\; a'b \leq c'd] \\ &\text{iff} \;\; [P(a|b) \leq P(c|d), \text{ for all (well-defined) P}] ,\end{aligned} \tag{3.21}$$

$$\begin{aligned}(a|b) = (c|d) \;\; &\text{iff} \;\; [(a|b) \leq (c|d) \;\; \text{and} \;\; (c|d) \leq (a|b)] \;\; \text{iff} \;\; (a|b)\&(c|d) = (a|b)\vee(c|d) \\ &\text{iff} \;\; [ab = cd \;\; \text{and} \;\; b = d] \\ &\text{iff} \;\; [P(a|b) = P(c|d), \text{ for all (well-defined) P}].\end{aligned} \tag{3.22}$$

Many other important properties hold, as well as full algebraic characterizations (again, see above-cited references). Other critical properties, used as the basis for deduction investigations within the framework of PSCEA are provided in later sections of this paper.

Consider now the imbedding of the non-boolean AC and DGNW into PSCEA. Though these operators in their original setting were all commutative, associative DeMorgan, and well-defined over all of the conditional events they acted upon, due to the boolean structure of B_o – especially using eq.(3.15) -- a discrepancy exists between these operators acting in their original spaces and acting over B_o, even in imbedded form. (See [Goodman & Nguyen b] for a more complete analysis.) However, this can be completely remedied for our purpose here: These operators will defined in a non-associative way for each number of arguments separately and restricted to only nontrivial first order conditional event arguments in B_o. Eq.(3.22) guarantees that this produces well-defined operators for all four conjunction and disjunction AC and DGNW operators relative to B_o. Let us denote the class of first order nontrivial conditional events in B_o as B_o* . Hence, for any finite index set J, $\&_{AC}$, $\vee_{AC}$, $\&_{DGNW}$, $\vee_{DGNW}$: $(B_o*)^J \rightarrow B_o$ are now all well-defined and formally the same operators as in their original context with the exception of the identifications of eq.(3.15), needed for B_o. In addition, we can also obtain in recursive form the full evaluations algebraically and probabilistically of PSCEA &, $\vee$: $(B_o*)^J \rightarrow B_o$,as well as their extensions to all of $B_o{}^J$. Using multivariate notation introduced earlier, for arbitary $\varnothing < a_j < b_j$ in B, j in J, where now $(a|b)_J$ indicates $(a_j|b_j)_{j \text{ in } J}$, etc., first define

$$\alpha((a,b)_{J,K}) = \&(b')_{J-K} \,\&\, \&a_K , \tag{3.23}$$

$$\alpha((a,b)_J) = \bigvee_{\varnothing \neq K \subseteq J}(\alpha((a,b)_{J,K})), \tag{3.24}$$

$$\alpha_o((a,b)_J) = \bigvee_{\varnothing \neq K \subseteq J}\big(\alpha((a,b)_{J,K}) \times \&(a|b)_{J-K}\big). \tag{3.25}$$

Next, using eqs.(3.23), (3.24), define and derive for $\&_{AC}:(B_o*)^J \rightarrow B_o$,

$$\&_{AC}(a|b)_J = (\&(b \Rightarrow a)_J \mid \vee (b_J)) = (\alpha((a,b)_J) \mid \vee (b_J)) , \tag{3.26}$$

with evaluations

$$P_o(\&_{AC}(a|b)_J) = P(\alpha((a,b)_J) \mid \vee (b_J)) = P(\alpha((a,b)_J)) / P(\vee (b_J)), \tag{3.27}$$

$$P(\alpha((a,b)_J)) = \sum_{\varnothing \neq K \subseteq J} P(\alpha((a,b)_{J,K})). \tag{3.28}$$

Note that we can also determine recursively the full general form in B_o of the PSCEA conjunction of any finite number of first order conditional events by extending the recursive form in eq.(3.13) (see eq.(3.30) and straightforward combinatorics involving the cartesian products (see also, e.g., [Goodman & Nguyen, 1995]):

$$\&(a\,|b)_J = [\alpha_o((a,b)_J) \mid \vee (b_J)]. \tag{3.29}$$

The notation $[\alpha\,|b]$ for any b in B and α in B_o stands for the analogue of the direct definition in eq.(3.13), but formally replaces a in B by $\alpha\,\&(b \times \Omega_o)$ in B_o:

$$[\alpha\,|b] = \bigvee_{j=0}^{+\infty} \big((b')^j \times (\alpha\,\&(b \times \Omega_o))\big) = \alpha\&(b \times \Omega_o) \;\vee\; (b' \times [\alpha|b]) \quad \text{in } B_o, \tag{3.30}$$

yielding

$$P_o([\alpha\,|b]) = P_o(\alpha\,\&(b \times \Omega_o)) / P(b) = P_o(\alpha|b \times \Omega_o). \tag{3.31}$$

Then, the recursive evaluation of eq.(3.29), using eqs.(3.30), (3.31) yields

$$P_o(\&(a\,|b)_J) = P_o(\alpha_o((a,b)_J)) / P(b_J) , \tag{3.32}$$

$$P_o(\alpha_o((a,b)_J)) = \sum_{\varnothing \neq K \subseteq J}\big(P(\alpha((a,b)_{J,K})) \cdot P_o(\&(a|b)_{J-K} \big). \tag{3.33}$$

Note the close analogy of forms between &AC and PSCEA &: formally speaking, &AC is the same as (PSCEA)& with all cartesian product factors omitted. But this difference is enough to make the former non-boolean, while the latter is boolean. This also immediately shows $\&_{AC}$

always dominates & in size. Since $\vee_{AC}$ is defined as the DeMorgan transform of $\&_{AC}$, it readily follows, using eq.(3.26) that

$$\vee_{AC}(a|b)_J = (\vee(a)_J \mid \vee (b_J)), \tag{3.34}$$

with evaluation

$$P_o(\vee_{AC}(a|b)_J) = P(\vee(a)_J \mid \vee (b_J)). \tag{3.35}$$

Analogously, $\vee(a|b)_J$ and $P_o(\vee(a|b)_J)$ can be obtained; the latter, e.g., via the usual Poincaré alternating sign expansion, and dually, $\vee_{AC}$ becomes dominated by PSCEA $\vee$.

Next, we consider $\&_{DGNW}:(B_o*)^J \to B_o$, where, for any $(a|b)_J$ in $(B_o*)^J$, the above-cited references show

$$\&_{DGNW}((a|b)_J) = (\&(a_J) \mid \vee(a'b)_J) \vee (\&(a_J)), \tag{3.36}$$

with DeMorgan transform

$$\vee_{DGNW}((a|b)_J) = (\vee(a_J) \mid \&(a'b)_J) \vee (\vee(a_J)), \tag{3.37}$$

and with straightforward probability evaluations. Additional ordering properties involving PSCEA &, $\vee$, and the imbedding of $\&_{AC}$, $\vee_{AC}$, $\&_{DGNW}$, $\vee_{DGNW}$ with respect to PSCEA ordering are shown in Section 4.

3.2 *RSC*

The field of RSC began in the mid 1970's to 1980's with the realization that a fundamental link exists between FL and PL via random sets, with emphasis on nested or level random sets (see, e.g., [Orlov, 1978], [Nguyen, 1979], [Goodman, 1981], [Höhle, 1981]). In a related direction, around the same time, non-random (isomorphic) counterparts between fuzzy sets and collections of level sets were pointed out in [Negoita & Ralescu, 1975]; but little progress has taken place beyond these original ideas. In a somewhat different, but non-probabilistic, direction, a possible alternative to the representation of certain FL concepts via RSC may well lie in generalized techniques based upon traditional interval analysis (such as originally exposited in [Alefeld & Herzberger, 1983]), typified by [Aubin & Frankowska, 1990].

Prior to the development of RSC, emphasis on the use of random sets -- which are no more than actually set-valued random quantities – was focused overwhelmingly in the obvious area: geometry and pattern recognition. To this day, this still remains a chief area of activity. (To see this, peruse, e.g., through the earlier monographs of [Kendall & Moran, 1963], [Harding & Kendall, 1974], [Matheron, 1975], and [Ripley, 1981], and the more recent works of [Stoyan & Stoyan, 1995], [Goutsias *et al.*, 1998].)

However, by extending the scope of random sets beyond random rectangles, circles, and even numerical ranges, to truly general set-valued forms, it is seen that there is one basic function associated with the local behavior of any random set – namely, its *one-point coverage function*, i.e., the probability of the *coverage event* that any arbitrary, but fixed, point is covered by (or is in) the random set. This plays a role in partially specifying a random set, analogous to the role expectation plays in partially specifying a random variable or random vector. If one's knowledge of a random set extends beyond its one-point coverage function to its (arbitrary, but fixed) two-point coverage function,…, and in fact to its (arbitrary, but fixed) set coverage function, this is indeed the well-known *Dempster-Shafer doubt function,* which, under mild assumptions, uniquely characterizes the random set in question. See, e.g. [Shafer, 1976], [Nguyen, 1978], and [Goodman & Nguyen, 1985, Chapters 3-5] for a further general exposition on this area, as well as relations to Choquet's earlier work [Choquet, 1954] on *infinite capacities*, and the similar, but distinct, Dempster-Shafer belief and plausibility functions, the knowledge of either also fully characterizing the corresponding random set. These quantities, as well, serve as natural lower and upper bounding functions on the induced probability measure under a particular function of a random variable when the actual range values of that function are only known up to being in prescribed events (rather than being actual points) [Dempster, 1967]. (See also a generalization of this in [Walley, 1991].)

On the other hand, the relatively weak knowledge of the one-point coverage function of any random set is just sufficient in many situations to construct a natural bridge (i.e., a homomorphic-like relation) between fuzzy logic concepts and corresponding probability ones. The following summary of results is fully documented in [Goodman & Nguyen, 1985, 1999], [Goodman, 1994, 1999], [Goodman *et al.*, 1997, Part III], and [Goodman & Kramer, 1997]:
In brief, compatible with the above comments, it can be demonstrated that any random subset of a finite set (with obvious modifications for the extensions of the result to the infinite set situation) is uniquely determined by two quantities: its one-point coverage function and some appropriately chosen *copula*, i.e., joint cumulative distribution function (cdf) over the unit n-cube all of whose marginal cdf's are uniform ones over the unit interval. (For background on copulas, see, e.g., [Goodman & Nguyen, 1985, Chapter 2], [Sklar, 1973], [Schweizer & Sklar, 1983], and [Dall'Aglio *et al*., 1991].) Noting that the one-point coverage function of any random set is also a fuzzy set (membership function), the above result demonstrates that *any* random subset of a finite set is partially specified by an appropriate (and uniquely corresponding) fuzzy set over its domain, and also by choosing an appropriate copula, it is fully determined by both. Conversely, it can be shown that *any* choice of fuzzy set membership function over a finite domain and any choice of copula, produces a uniquely corresponding random subset of that domain whose one-point coverage function matches the given fuzzy set. Because, for a given fuzzy set, any copula can be chosen generating a random set satisfying the last statement, and that in general, distinct copulas produce distinct random sets, a full many-to-one natural relation exists between random sets and fuzzy sets. Moreover, all of the above results can be extended to establishing a many-to-one (one-point coverage) relation between any finite collection of fuzzy sets over finite domains and appropriate joint collections of random subsets of those domains.

In turn, for any choice of copula and corresponding *cocopula* (the DeMorgan dual of the copula – e.g., see again [Goodman & Nguyen, 1985, Chapter 2]), leading to a corresponding FL conjunction and disjunction operator pair, the above one-point coverage relations between fuzzy sets and random sets can then be augmented to *homomorphic-like relations* between such fuzzy logic operators separately and corresponding CL operators acting upon one-point coverage events resulting from the random sets involved. (For an explanation of this term and documentation of these results, again see the above Goodman & Nguyen and Goodman & Kramer references.) It should be noted at this point that the FL operator pairs (min, max) and (prod, probsum) are also copula, cocopula pairs. By appropriately restricting the class of copulas and corresponding cocopulas – which still includes (min, max) and (prod, probsum), these RSC relations may be further extended to represent in a homomorphic-like sense through random sets arbitrary finite combinations of well-defined (and requiring non-repeatable arguments for the choice of (prod, probsum)) fuzzy logic conjunction and disjunction operators.

The above RSC relations can be specialized to a wide variety of well-known FL concepts, including: Zadeh's now universally used "extension principle" (the fuzzy logic counterpart of probability theory's propagation or transformation of errors / randomness - see, e.g. the texts of [Dubois & Prade, 1980] and [Nguyen & Walker, 1997] for background); a new approach to defining conditional fuzzy sets which extend the PSCEA approach to conditional events; and, in conjunction with REA (see below), fuzzy logic modifiers can be interpreted probabilistically (Again, see [Goodman, 1999] and [Goodman & Nguyen, 1999] for details.) In short, RSC serves as both a means to interpret many fuzzy logic concepts in a probabilistic context and as a guide for defining new fuzzy ideas, based upon homomorphic-like random set counterparts. *Most importantly here*, RSC and REA (see just below) *applied to cognitive probability logic*, as discussed in Sections 4-6, allow that logic to be extended to treat linguistic-based information. It is also hoped that the sound homomorphic-like relations established between FL and PL via use of RSC will dispel misconceptions concerning the nature and relations between both disciplines, such as in [Elkan, 1993]. (See [Goodman, 1998b] for a survey of the controversy.)

3.3 *REA*

REA includes PSCEA as a special case. (For basic background on REA, see [Goodman *et al.*, 1997, Part III], [Goodman & Kramer, 1997], and [Goodman & Nguyen, 1998, 1999].)

First, recall that conditional probabilities are simply divisions of probabilities of consequent events by antecedent events (with the proviso, in effect, that the consequent is a subevent or subset of the antecedent). That is, conditional probabilities are simply special cases of functions of probabilities with values lying in the unit interval, subject to certain (algebraic or logical) constraints among the events involved. More generally, models of real-world situations may involve functions of probabilities, also with values in the unit interval, but other than just divisions, *both explicitly and implicitly*. These two terms will be clarified as follows: One class of examples illustrating models based upon explicit probabilities consists of weighted linear combinations of probabilities of possibly overlapping events representing fused expert judgments of complex situations; one class of examples illustrating based upon implicit probabilities is in the use of operators over the unit interval or n-cube with range in the unit interval as FL *modifiers*, representing such linguistic concepts as the one-argument (n = 1) modifiers "very", "much", "more or less", "not extremely" – these often being in the form of exponential or power functions over the unit interval, or more generally, as monotonic (increasing functions from 0 up to 1 in the first three cases and decreasing down from 1 to 0 in the last case). Even more generally, the functions could be unimodal, analytic, etc. But, from the above discussion concerning RSC, the fuzzy set membership functions that such FL modifiers act upon, themselves are the same as probabilities of one-point coverage events. Hence, using RSC, one can always interpret such FL modifier statements as models which are (explicit) functions of probabilities. (For more specifics, see [Goodman & Kramer, 1997] and [Goodman & Nguyen, 1999].)

Thus, a number of real-world situations can arise which are modeled as various functions of input probabilities. Then, analogous to the role that PSCEA plays in representing conditional expressions compatible with conditional probability evaluations, one seeks REA to obtain a probability space which contains, in a natural imbedded sense, all of the initial unconditional events and, as well, for each such model a single "relational" (replacing "conditional") event which is compatible with the model, i.e., the probability evaluation of the event (for this extended space) coincides with the given function of probabilities the model represents.

Numerical-valued function $f:[0,1]^n \to [0,1]$	**Relational event counterpart $f_o: \boldsymbol{B}^n \to \boldsymbol{B}_o$ satisfying for all n event arguments a,b,c,.. in *B*, possibly constrained, and all P, $P_o(f_o(a,b,c,\ldots)) = f(P(a),P(b),P(c),\ldots)$**	**Restrictions, Explanations**
Addition: $P(a) + P(b)$	$a \vee b$	n=2; a, b disjoint
Subtraction: $P(b) - P(a)$	$b\text{-}a = a'b$	n=2; a is subevent of b
Multiplication: $P(a) \cdot P(b)$	$a \times b$	n=2
Constant functions: λ in [0,1]	$\theta(\lambda)$	see [Goodman *et al.*,1997, Part III] for explanation
Positive integer powers: $(P(a))^m$	$a^m = a \times \ldots \times a$ (m factors)	n=1
General real positive powers: $(P(a))^r$	$a^r = a^{[r]} \times a^{\{r\}}$; see above for $a^{[r]}$; $a^{\{r\}} = \bigvee_{j=0}^{+\infty} (a')^j \times a \times \theta(\lambda_{j,r})$	n=1; r real > 0; $[\cdot]$ greatest integer function, $\{\cdot\} = [\cdot] - \cdot$, fractional part function, $\lambda_{j,r} = (1-\{r\}) \cdot (1-(\{r\}/2)) \cdots (1-(\{r\}/j))$
Finite polynomials: $\sum_{j=0}^{m} \ddot{e}_j \cdot (P(a))^j$	$\bigvee_{j=0}^{m-1} (a^j \times a' \times \theta(\lambda^{(j)})) \vee (a^n \times \theta(\lambda^{(m)}))$	n=1; all $0 \le \lambda_j \le 1$, $\lambda^{(j)} = \lambda_0 + \lambda_1 + \ldots + \lambda_j \le 1$
Infinite series / analytic functions around 0 : $\sum_{j=0}^{+\infty} \ddot{e}_j \cdot (P(a))^j$	$\bigvee_{j=0}^{+\infty} (a^j \times a' \times \theta(\lambda^{(j)}))$	n=1; all $0 \le \lambda_j \le 1$, $\lambda^{(j)} = \lambda_0 + \lambda_1 + \ldots + \lambda_j \le 1$
Weighted affine functions in multiple arguments: $\sum_{j=1}^{n} \ddot{e}_j \cdot P(a_j) + \ddot{e}_0$ Ex: $\lambda_0 + \lambda_1 P(a_1) + \lambda_2 P(a_2)$	$\bigvee_{(\kappa \text{ in } \{\varnothing,\Omega\}^{\{1,\ldots,n\}})} (a_\kappa \times \theta(\lambda_\kappa))$ Ex: $a_1 a_2 \times \theta(\lambda_0+\lambda_1+\lambda_2) \vee a_1 a_2' \times \theta(\lambda_0+\lambda_1) \vee a_1' a_2 \times \theta(\lambda_0+\lambda_2) \vee a_1' a_2' \times \theta\lambda_0)$	$n \ge 1$; all $0 \le \lambda_j \le 1$, $\lambda_\kappa = \sum_{j \text{ in } \kappa^{-1}(\Omega)} \lambda_j + \lambda_0 \le 1$; $a_\kappa = \mathop{\&}_{j=1}^{n} (a_j \Delta(\kappa(j))')$
Exponentials: $e^{-P(a)}$	$\bigvee_{j=0}^{+\infty} ((a')^j \times a \times \theta(\mu^{(j)}))$	n=1; $\mu^{(j)} = (1/e) \cdot \sum_{k=0}^{j} 1/k!$, j=0,1,2,...
Division: Conditional events: $P(a)/P(b) = P(a\|b)$	$(a\|b) = \bigvee_{j=0}^{+\infty} ((b')^j \times a \times \Omega_o)$ in B_o;	n=2; a is subevent of b (see Section 3.1)

Table 1. Examples of Relational Events Corresponding to Given Numerical Functions.

Indeed, it has been shown that REA can be applied to a wide class of unary and n-ary argument functions of probabilities, including linear, bilinear, exponential, restricted analytic and polynomial forms (again, see the above listed references). In summary, the basic problem REA – including CEA – addresses and solves, can be formulated mathematically as follows for the case of two models to be compared or combined:

Given probability space $(\Omega, \mathcal{B}, P)$, with P variable, construct probability space $(\Omega_o, \mathcal{B}_o, P_o)$ extending the initial space in an isomorphic probability preserving way (such as in the case of PSCEA) with Ω, $\mathcal{B}$, Ω_o, $\mathcal{B}_o$ not depending on any choice of P and corresponding P_o, so that for given functions f,g: $[0,1]^n \to [0,1]$: we can find event-valued functions f_o, g_o: $\mathcal{B}^n \to \mathcal{B}_o$ such that for all P (well-defined) and all $a_1,\ldots, a_n$ in $\mathcal{B}$ – subject to some possible fixed constraints –

$$P_o(f_o(a_1,\ldots,a_n)) = f(P(a_1),\ldots,P(a_n)), \quad P_o(g_o(a_1,\ldots,a_n)) = g(P(a_1),\ldots,P(a_n)). \tag{3.38}$$

Obviously, the relational events here are $f_o(a_1,\ldots,a_n)$, $g_o(a_1,\ldots,a_n)$ in $\mathcal{B}_o$, representing in eq.(3.38) the numerical models $f(P(a_1),\ldots,P(a_n))$, $g(P(a_1),\ldots,P(a_n))$, respectively. Equivalently, this can be thought of as a commutative relationship between f, g acting upon the $P(a_j)$'s and P_o evaluating the relational events. One can then logically combine $f_o(a_1,\ldots,a_n)$ with $g_o(a_1,\ldots,a_n)$, as needed, relative to $\mathcal{B}_o$ and then evaluate this combination via P_o. We provide a table below of basic functions of probabilities and corresponding relational event pairs (f, f_o) for a number of situations. Finally, the tie-in between REA and Lindley's result on characterizing a large class of functions of (finitely additive conditional) probabilities via a two person zero sum game [Lindley, 1982] should be mentioned. (See the discussion in Section 4.4.)

3.4 *SOP*

SOP techniques have been around for some time, as documented in the text of [Aitchison, 1986]. However, in-depth applications to: problems of logical deduction [Goodman & Nguyen, 1998] and updating information [Goodman & Nguyen, b] have only recently been attempted, the latter in connection with [Grove & Halpern, 1997] for the "Judy Benjamin Problem". In a nutshell, the SOP approach is theoretically applicable to any situation where one seeks initially, relative to the same possible choice of probability measure, the determination of bounds on the (unconditional or conditional) probability of certain events of interest, given a collection of bounds or constraints on the probability of the same and/or other events. SOP is then carried out in three or, optionally four, basic steps:

(1) One forms either the collection of all (nonvacuous) *relative atomic events* generated by all of the initial relevant events involved in the problem or any finite refinement of such.

(2) One identifies any possible probability measure for this problem in a natural way as a corresponding constant vector, where each entry corresponds to a particular atom being evaluated by that probability measure. Thus, the size of the vector matches the cardinality of the set of distinct atomic events involved and the components of the vector lie in the unit interval and add up to unity. (Kosko in several papers and texts – see, e.g., [Kosko, 1990, 1997] – extends this idea to the class of fuzzy sets over a finite domain.) Thus, the collection of all possible probability measures here is identifiable either as the face of a multidimensional simplex (due to the unity sum constraint) or, if the last component is omitted, as a full multidimensional simplex of one less dimension.

(3) The bayesian aspect of the approach commences when one chooses a prior (and hence, second order) distribution over the simplex representing the possible probability measures relative to their evaluations of the atoms. Thus, in effect, the deterministic possible probability measures of interest are replaced by a joint random vector and all given constraints on the probability of events of interest become corresponding constraints on (parts of the) joint random vector. Finally, the initial problem is modified by replacing the goal of determining the bounds on the possible probability values (subject to the same constraints) by the evaluation of the posterior expectation of the joint random vector. Often, it is reasonable to choose the prior to be a uniform, or more generally (and more flexibly), a dirichlet one.

(4) For either choice in step (3), one can also show, using an extension [Goodman & Nguyen a] of Lukacs' characterization of the gamma and dirichlet distributions [Lukacs, 1955], that a transformation exists which converts the problem to an analogous one of computing the posterior expectation subject to constraints, but no longer confined to the simplex. Instead the constraints involve the all-positive orthant of multidimensional space and the transformed random vector consists of independent marginal gamma random variables. In the case of a joint uniform assumption, this reduces to independent exponential random variables with common parameter being unity.

Because the above-mentioned theorem will be used in Section 6, a full statement of it is provided here. First, recall [Johnson & Kotz, 1972], [Wilks, 1963] the following standard probability distributions. For any fixed positive real numbers q, r, the gamma distribution Gam(q,r) with corresponding probability density function (pdf) $gam_{q,r}$ is given for its non-zero values as

$$gam_{q,r}(x) = \gamma_{q,r} \cdot x^{q-1} \cdot e^{-x/r}\text{ , if } x > 0\text{ ;}\qquad \gamma_{q,r} = 1/(r^{q} \cdot \Gamma(q)), \tag{3.39}$$

where $\Gamma(.)$ is the standard gamma function. For any $\underset{\sim}{q} = (q_1,\ldots,q_{n+1})$, with each q_j a fixed positive real number, the dirichlet distribution $Dir(\underset{\sim}{q})$, with corresponding joint pdf $dir_{\underset{\sim}{q}}$ is given for its non-zero values as

$$dir_{\underset{\sim}{q}}(y_1,\ldots,y_n) = \lambda_{\underset{\sim}{q}} \cdot y_1^{q_1-1} \cdots y_n^{q_n-1} \cdot (1-y_1-\ldots-y_n)^{q_{n+1}-1}\text{, if } (y_1,\ldots,y_n) \text{ in } Q_n\text{ ;} \tag{3.40i}$$

$$\lambda_{\underset{\sim}{q}} = \Gamma(q_1+\ldots+q_{n+1}) / (\Gamma(q_1)\cdots\Gamma(q_{n+1}))\text{ ;} \tag{3.40ii}$$

and *full n-simplex*

$$Q_n = \{(y_1,\ldots,y_n)\text{: } 0 \leq y_j \leq 1,\ j=1,\ldots,n;\ y_1+\ldots+y_n \leq 1\}. \tag{3.41}$$

Theorem 3.1. [Goodman & Nguyen, a] Extension of Lukacs' characterization to dirichlet distributions.

Let $Y_1,\ldots,Y_n$ be any n positive non-degenerate random variables. Then:

(I) $Y_1,\ldots,Y_n$ are jointly distributed as $Dir(\underset{\sim}{q})$ iff there exist n+1 positive non-degenerate random variables $X_1,\ldots,X_{n+1}$ such that they are mutually independent, with each $Y_j =$ $X_j / (X_1+\ldots+X_{n+1})$ and being independent of $X_1+\ldots+X_{n+1}$, j =1,…,n.

(II) In (I), each X_j may always be chosen so that each is independently distributed Gam(q_j,r), where each q_j is the jth component of $\underset{\sim}{q}$ and r>0 is arbitrary fixed. ■

4. *The Four Tools and Their Relations to PL and Associated Logics*

4.1 *Preliminary Remarks and Notation*

The chief application of SOP here, in conjunction with the use of CEA, REA, and RSC, is to the development of *Cognitive Probability Logic* (CPL). This logic successfully addresses, for the first time in a fully quantitative manner, the following quandary: We know that use of the classical logic operator representing "if, then" (see Section 2) produces a whole host of valid – as well as invalid -- deductive relations. It would be desirable to have, in some sense, a sort of PL validity, so that a number of deductions that are valid for the material conditional interpretation of "if, then" would also be PL-valid. In the following development, for purpose of efficiency, we extend our use of multivariable notation and introduce a basic assumption which will be used a number of times later:

First, for any probability space (Ω,B,P), finite (nonvacuous) index set J, and any events a_j, b_j in B, recall – as well as define -- the following symbols: a_J for $(a_j)_{j \text{ in } J}$; b_J for $(b_j)_{j \text{ in } J}$; $(b\Rightarrow a)_J$ for $(b_j\Rightarrow a_j)_{j \text{ in } J}$; $(a|b)_J$ for $(a_j|b_j)_{j \text{ in } J}$, where each conditional event $(a_j|b_j)$ in B_o; $P(b\Rightarrow a)_J$ for $(P(b_j\Rightarrow a_j))_{j \text{ in } J}$, t_J in $[0,1]^J$ for $(t_j)_{j \text{ in } J}$, where each real number t_j in [0,1]; in particular, letting $t_j = 1$, all j in J, $1_J = (1)_{j \text{ in } J}$, $\&(a_J)$ for $\underset{j \text{ in } J}{\&}(a_j)$, $\&(a|b)_J$ for $\underset{j \text{ in } J}{\&}(a_j|b_j)$, etc.

Next, introduce our two basic assumptions, where, depending on the situation, the first or the stronger second holds:

Assumption I: (Ω,B,P) is any given probability space, with PSCEA extension, denoted, as usual, by (Ω_o,B_o,P_o); J is a finite index set, $(a|b)_J$ in B_o^J is a family of nontrivial conditional events, each in B_o, i.e., $\varnothing < a_j < b_j$, j in J; (c|d) in B_o is a nontrivial conditional event, i.e., $\varnothing < c < d$; and P – and hence P_o – is arbitrarily variable but, subject to the restriction that $0 < P(\vee(b_J))$. When all $b_j = \Omega = d$, apropos to the isomorphic probability-preserving imbedding of B in B_o discussed in Section 3.1.2 (but with the caution of Lewis' theorem being present), we identify a_j in B with $(a_j|\Omega)$ in B_o, j in J, and c in B with $(c|\Omega)$ in B_o.

Assumption II: Assumption I, together with the assumption that $P(b)_J > 0_J$, i.e., $P(b_j) > 0$, for all j in J and all b_J and P satisfying Assumption I. ■

Note also that under Assumption II, apropos to the discussion following eq.(3.22), the PSCEA-imbedded AC and DGNW operators $\&_{AC}$, $\vee_{AC}$, $\&_{DGNW}$, $\vee_{DGNW}$: $(B_o^*)^J \rightarrow B_o$ are all well-defined slightly modified versions of the original non-boolean operator counterparts (relative to the inclusion of trivial conditional event arguments and trivial conditional event range values)conditions), where we recall B_o^* is the set of all nontrivial (first order) conditional events in B_o. In brief review, it is readily seen by inspection, that the usual definition of *tautological deduction* in CL is actually the same as *conjunctive deduction* relative to the boolean algebra interpretation. (For the former situation, see, e.g., [Copi, 1986] and [Bergmann *et al.*, 1980].) Thus, we say that (unconditional) *premise* a_J (tautologically) *deduces* (unconditional) conclusion c in CL iff a_J *conjunctively deduces* c in CL, i.e., iff

$$\&(a_J) \leq c . \tag{4.1}$$

Extending this, we will say that *linguistic premise* ("if b_j, then a_j") $_{j \text{ in } J}$ deduces conclusion "if d, then c" *in the material conditional sense* iff

$$\&(b\Rightarrow a)_J \leq d\Rightarrow c\ , \tag{4.2}$$

i.e., when $(b\Rightarrow a)_J$ deduces $d\Rightarrow c$ in CL, noting the obvious reduction when all $b_j = d = \Omega$. Let us denote the logic based upon this reasoning as *Material Conditional Logic* (MCL), an actual fragment of CL.

4.2 *Extending Deduction to PL*

Next, it is desirable to be able to extend CL in some way to PL: a number of attempts have been made, including the use of conditional probabilities, one of which we will discuss below. (See, e.g., [Hailperin, 1996] and [Rescher, 1969, Section 2.27, *et passim*] for general background.) In any case, relative to PL, under Assumption I, in light of the conjunctive form in eq.(4.1), *one basic function that measures the degree that given premise* a_J *deduces conclusion* c *is simply the conditional probability*

$$P(c \mid \&(a_J)), \tag{4.3}$$

assuming $P(\&(a_J)) > 0$. In fact, in view of eq.(2.3) and standard results,

$$[P(c \mid \&(a_J)) = 1, \text{ for all P so that } P(\&(a_J)) > 0] \quad \text{iff} \quad \&(a_J) \leq c. \tag{4.4}$$

But, since the quantity in eq.(4.3) is dependent upon a particular probability measure P, a reasonable alternative measure not dependent directly upon any one probability measure -- yet apparently dependent upon probability in general -- is the *minimum conclusion function* (following Adams' ideas – see any of his references) $\text{minconc}(a_J;c):[0,1]^J \to [0,1]$, where, for any $0 \leq t_j \leq 1$, j in J, letting $t_J = (t_j)_{j \text{ in } J}$,

$$\text{minconc}(a_J;c)(t_J) = \inf\{P(c): \text{all prob. meas. } P{:}\mathcal{B} \to [0,1] \text{ with } P(a_j) \geq t_j, \text{ j in J}\}. \tag{4.5}$$

Again, analogous to eqs.(4.3) and (4.4), for a given P, one could use as a *basic function that measures the degree that a given (conditional) premise* $(a|b)_J$ *deduces (conditional) conclusion* $(c|d)$ *is simply the conditional probability with respect to PSCEA*

$$P_o((c|d) \mid \&(a|b)_J) = P_o((c|d)\&\&(a|b)_J) / P(c|d), \tag{4.6}$$

etc., since PSCEA obeys all the laws of PL. In fact, relative to $\mathcal{B}_o$,

$$\text{if } \&(a|b)_J \leq (c|d), \text{ then for all P with } P_o(\&(a|b)_J) > 0,\ P_o((c|d) \mid \&(a|b)_J) = 1. \tag{4.7}$$

Analogous to the above comments for the unconditional situation (following eq.(4.4)), we may seek to find only a general dependency on probability for deduction by extending the minconc function to possible deduction schemes involving conditional statements "if b_j, then a_j" in the premise and "if d, then c" in the (potential) conclusion whether interpreted, e.g., via probabilities of material conditionals or conditional probabilities in the obvious senses via the definition in eq.(4.5),

$$\text{minconc}((b\Rightarrow a)_J;d\Rightarrow c)(t_J) = \inf\{P(d\Rightarrow c): \text{all prob. meas. } P{:}\mathcal{B} \to [0,1] \text{ with } P(b_j\Rightarrow a_j) \geq t_j, \text{ j in J}\}, \tag{4.8}$$

$$\text{minconc}((a|b)_J;(c|d))(t_J) = \inf\{P(c|d): \text{all prob. meas. } P{:}\mathcal{B} \to [0,1] \text{ with } P(a_j|b_j) \geq t_j, \text{ j in J}\}, \quad (4.9)$$

noting for either eq.(4.8) or (4.9), minconc $[0,1]^J \to [0,1]$ has same domain and range.

Adams has produced several reasonable forms of probability logic, including *High Probability Logic* (HPL) and *Certainty Probability Logic* (CPL) which intimately involve the use of the minconc function: Under Assumption II, HPL yields as valid deductions those premise, conclusion pairs $(a|b)_J$, $(c|d)$ so that for any P, whenever $P(a|b)_J$ is "sufficiently high", $P(c|d)$ is "high"; while CPL yields as valid deductions those pairs such that for any P, whenever $P(a|b)_J$ is unity, so is $P(c|d)$. (See, e.g., [Adams, 1975, 1996, 1998], as well as work of [Schurz, 1983] and [Goodman & Nguyen, b] in clarifying certain of Adams' proofs and results.) These definitions certainly apply to unconditional expressions as well. In terms of the minconc function, Adams' HPL and CPL can be conveniently defined as follows in a PSCEA context, under Assumption II, for any pair $(a|b)_J$, $(c|d)$:

$$(a|b)_J \text{ deduces } (c|d) \text{ in HPL sense, written as } (a|b)_J \leq_{HPL} (c|d),$$
$$\text{iff} \qquad \lim_{(t_J \uparrow 1_J,\ \text{uniformly})} (\text{minconc}((a|b)_J;(c|d))\,(t_J)) = 1; \qquad (4.10)$$

under only Assumption II,

$$(a|b)_J \text{ deduces } (c|d) \text{ in CPL sense, written as } (a|b)_J \leq_{CPL} (c|d),$$
$$\text{iff} \qquad \text{minconc}((a|b)_J;(c|d))\,(1_J)\,) = 1, \quad \text{if } P(b_J) > 0, \qquad (4.11)$$

or, weakening this to Assumption I, replace the latter in eq.(4.11) by

$$\text{minconc}((a|b)_{J\text{-}K};(c|d))(1_{J\text{-}K})) = 1, \text{ if there is some } \varnothing \neq K \subset J \text{ with } P(b_K) = 0_K.$$

4.3 *Fundamental Relations among PL and CL Deductions*

The following theorems characterize and relate HPL, CPL, MCL, and CL deductions for unconditional and conditional events in PSCEA, as well as with the imbedded AC and DGNW operators.

Theorem 4.1. [Goodman & Nguyen 1998, b]

Noting that there are four possible pure ordering relations between $(a|b)_J$ and $(c|d)$ with respect to &, $\vee$ over $\mathcal{B}_o$, the following characterizations hold involving the imbedding $\&_{AC}$, $\vee_{AC}$, $\&_{DGNW}$, $\vee_{DGNW}$, under Assumption II:

(i) $(a|b)_J \leq (c|d)$, i.e., $\underset{\text{j in J}}{\text{And}}((a_j|b_j) \leq (c|d))$ iff $\vee(a|b)_J \leq (c|d)$ iff $\vee_{DGNW}(a|b)_J \leq (c|d)$.

(ii) $(c|d) \leq (a|b)_J$, i.e., $\underset{\text{j in J}}{\text{And}}((c|d) \leq (a_j|b_j))$ iff $(c|d) \leq \&(a|b)_J$ iff $(c|d) \leq \&_{DGNW}(a|b)_J$.

(iii) $\&(a|b)_J \leq (c|d)$ iff $\underset{\varnothing \neq K \subseteq J}{\text{Or}}(\&_{AC}(a|b)_K \leq (c|d))$.

(iv) $(c|d) \leq \vee(a|b)_J$ iff $\underset{\varnothing \neq K \subseteq J}{\text{Or}}((c|d) \leq \vee_{AC}(a|b)_K)$.

Proof: (i) follows directly from use of the criterion in middle right of eq.(3.21). (iii) is shown in {Goodman & Nguyen, b]. By DeMorgan duality: (ii) follows from (i), (iv) from (iii). ■

Corollary 4.1. Under Assumption II,

$$\&_{DGNW}(a|b)_J \leq \&(a|b)_J \leq \&_{AC}(a|b)_J \leq \vee_{AC}(a|b)_J \leq \vee(a|b)_J \leq \vee_{DGNW}(a|b)_J .$$

Proof: Theorem 4.1(ii) shows $\&_{DGNW}(a|b)_J \leq \&(a|b)_J$ and inspection of the very definitions readily show $\&(a|b)_J \leq \&_{AC}(a|b)_J$. The remainder of results follows from DeMorgan duality. ■

In addition to Adams characterizing HPL purely algebraically, via a certain non-boolean-structured conditional event algebra ("quasi"-conjunction) operator (again, see the basic Adams references), more recently it has been demonstrated that HPL deduction when placed in a PSCEA setting is fully characterized as the natural extension of CL conjunctive deduction:

Theorem 4.2. [Adams, 1996], [Goodman & Nguyen, b] Combination of Adams' algebraic characterization of HPL via $\&_{AC}$ and use of & in PSCEA. Under Assumption II,

$$(a|b)_J \leq_{HPL} (c|d) \quad \text{iff} \quad \underset{\emptyset \neq K \subseteq J}{\mathrm{Or}}\left(\&_{AC}(\underline{a},\underline{b},K) \leq (c|d) \right) \quad \text{iff} \quad \&(a|b)_J \leq (c|d). \tag{4.12}$$

Proof: Follows immediately from Adams' algebraic characterization (left-hand side equivalence shown in [Adams, 1996] and right-hand side equivalence shown in Theorem 4.1(iii). ■

Thus, Theorem 4.2 shows that standard CL deduction (again see, e.g. [Copi, 1986], [Bergmann *et al.*, 1980]), which is actually conjunctive deduction relative to a boolean algebra interpretation, extends directly to Adams' HPL in a PSCEA setting. Obviously, this enhances the possible use of $P_o((c|d) \mid \&(a|b)_J)$ as a measure of the degree of deduction of $(a|b)_J$ with respect to $(c|d)$, for a given P (again, see eqs.(4.6), (4.7)). Adams has also provided other characterizing criteria, as well as full axiomatizability, for HPL deduction [Adams, 1996, 1998].

In a related direction, the following *negative* conclusion proven by Adams throughout his works, dispels the belief that the $\underset{t_J \to 1}{\mathrm{limit}}(\,\mathrm{minconc}((a|b)_J;(c|d))(t_J))$ somehow can be used as a natural measure of the degree of deduction of premise $(a|b)_J$ with respect to (potential) conclusion $(c|d)$.

Theorem 4.3. [Adams, 1996]
Under Assumption II, either $\underset{t_J \to 1}{\mathrm{limit}}(\,\mathrm{minconc}(a|b)_J;(c|d))(t_J)) = 1$ or 0, with no nontrivial values in between, accordingly as $(a|b)_J \leq_{HPL} (c|d)$ or not$[(a|b)_J \leq_{HPL} (c|d)]$.
Proof: See above Adams' reference. ■
It also appears that, in many cases, Theorem 4.3 can be strengthened to include non-limiting t_J, but this remains to be investigated.

Theorem 4.4. ([Adams, 1996] modified) Characterization of CPL being the same as MCL.
Under Assumption I:

$$\begin{aligned}(a|b)_J \leq_{CPL} (c|d) \quad &\text{iff} \quad \text{for all P (with } P(b_1 \vee b_2) > 0\text{), if } P(\&(b \Rightarrow a)_J) = 1\text{, then } P(d \Rightarrow c) = 1 \\ &\text{iff} \quad \&(b \Rightarrow a)_J \leq d \Rightarrow c\,. \end{aligned} \tag{4.13}$$

Proof: It is instructive to show this here. Note that the top equivalence readily follows from the definition in eq.(4.11). Apropos to eq.(2.3), for any P such that $P(b_j) > 0$, j in J, $P(d) > 0$,

$$1_J = P(a|b)_J \text{ iff } 1_J = P(b\Rightarrow a)_J \text{ iff } 1 = P(\&(b\Rightarrow a)_J); \qquad 1 = P(c|d) \text{ iff } 1 = P(d\Rightarrow c). \tag{4.14}$$

Hence, by monotonicity of probability, if RHS(4.13) holds, then eq.(4.14) shows holds. Conversely, we show first LHS(4.13) implies RHS(4.13), first for J = {1}, where, for convenience, we replace a_1 by a, b_1 by b. Suppose LHS(4.13) so holds, but RHS(4.13) doesn't. Thus, $\emptyset < (b\Rightarrow a)c'd$, implying either Case 1: $\emptyset < ac'd$ or Case 2: $ac'd = \emptyset$ and $\emptyset < b'c'd$:

Case 1: Let ω in $ac'd$ and define P(ω) = 1, yielding P(b) = P(d) = 1 > 0, satisfying Assumption II, with $P(a|b) = P(c'|d) = 1$, whence $P(c|d) = 0$, contradicting CPL property on LHS(4.9), where $P(a|b) = 1$ must imply $P(c|d) = 1$.

Case 2: Since by Assumption I, $\emptyset < a = ac'd \vee ac \vee ad'$, one of these must be nonvacuous, implying two Subcases: Subcase 1: $\emptyset < ac$ and $\emptyset < b'c'd$ or Subcase 2: $\emptyset = ac$ and $\emptyset < ad'$ and $\emptyset < b'c'd$.

Subcase 1: Let ω_1 in $b'c'd$, ω_2 in ac and define $P(\omega_1) = P(\omega_2) = ½$. This yields P(b) = ½ > 0, P(d) = 1 > 0, satisfying Assumption I, with P(a|b) = (1/2)/(1/2) = 1, but P(c|d) = (1/2)/1 = ½, contradicting CPL property that P(a|b) = 1 must imply P(c|d) = 1.

Subcase 2: Let ω_1 in $b'cd$, ω_2 in ad' and define $P(\omega_1) = P(\omega_2) = ½$. This yields P(b) = P(d) = ½ > 0, thus satisfying Assumption I, but $P(a|b) = P(c'|d)$ = (1/2)/(1/2) =1, implying P(c|d) = 0, contradicting CPL property that P(a|b) = 1 must imply P(c|d) = 1.

Hence, for J = {1}, LHS(4.13) implies RHS(4.13). For J = {1,2}, under Assumption I, note that

$$(b_1\Rightarrow a_1)\&(b_2\Rightarrow a_2) = (b_1\vee b_2) \Rightarrow \alpha, \tag{4.15}$$

where

$$\alpha = (b_1\Rightarrow a_1)\&(b_2\Rightarrow a_2)\,\&(b_1\vee b_2) = a_1b_2' \vee a_2b_1' . \tag{4.16}$$

For any P satisfying Assumption I, note that $P(b_1\vee b_2) > 0$. Furthermore, similar to eq.(4.14),

$$1= P(b_1\Rightarrow a_1) = P(b_2\Rightarrow a_2) \text{ iff } P(b_1\Rightarrow a_1)\&(b_2\Rightarrow a_2) = 1 \text{ iff } P((b_1\vee b_2) \Rightarrow \alpha) = 1. \tag{4.17}$$

When any equivalent part of eq.(4.17) holds, note that we cannot have P(α) = 0, since that implies $P(b_1\vee b_2) = 0$, contradicting Assumption I.

In any case, for J = {1,2}, LHS(4.13) implies RHS(4.13) by simply considering the result for J= {1} with b replaced by $b_1\vee b_2$ and a by α, etc., taking into account all boundary cases, such as $\alpha = b_1\vee b_2$, whence it can be shown that this implies, e.g., $a_j'b_j = \emptyset$, contradicting Assumption I. The procedure can then be continued inductively in this way for arbitrary finite J. Hence, LHS(4.13) implies RHS(4.13) in general. ■

Next, the basic relation between HPL and CPL deduction is shown:

Theorem 4.5. ([Adams, 1996], modified for PSCEA setting)

(i) HPL deduction implies CPL deduction in the following slightly restricted sense: Under Assumption II, for any $(a|b)_J \leq_{HPL} (c|d)$ and P,

if $P(a|b)_J = 1_J$, then $P(c|d) = 1$.

(ii) In (i), if P does not satisfy Assumption I, the conclusion may not hold. In particular, suppose $\alpha_1,\ldots,\alpha_5$ are any events in B which are nonvacuous, mutually disjoint and with $\alpha_1\vee\ldots\vee\alpha_5 \le \Omega$. Let J = {1}, $a_1 = \alpha_1$, $b_1 = \alpha_1 \vee \alpha_3 \vee \alpha_4$, $c = \alpha_1 \vee \alpha_2$, $d = \Omega$. Clearly, $(a_1|b_1)$, $(c|d)$ (identified with c) are nontrivial conditional events in B_o, and by the criterion in eq.(3.21), $(a_1|b_1) \le (c|d)$ (in fact, < holds here), whence by Theorem 4.2, $(a_1|b_1) \le_{HPL} (c|d)$ easily holds. On the other hand, choose any P such, $P(\alpha_1) = P(\alpha_3) = P(\alpha_4) = 0 < P(\alpha_5)$, implying $0 \le P(\alpha_2) \le 1- P(\alpha_5) < 1$. Thus, $P(b_1) = 0$, violating Assumption I, yielding $P(b_1\Rightarrow a_1) = 1 > P(\alpha_2) = P(d\Rightarrow c)$, meaning not $[(a|b) \le_{CPL} (c|d)]$.

(iii) Under Assumption I, there are pairs $(a|b)_J$, $(c|d)$ such that $(a|b)_J\le_{CPL} (c|d)$ holds but where not $[(a|b)_J\le_{HPL} (c|d)]$ holds, such as those corresponding to transitivity-syllogism, positive conjunction, and strengthening antecedents in Table 4.1 (as explained in Section 4.4).

Proof: (i) is shown immediately by using Theorem 4.2 (RHS(4.12)) and basic properties of probabilities. ■

The final theorem in this series shows for unconditional premises and conclusions the reduction of HPL and CPL to CL and related deductions.

Theorem 4.6. (See also [Goodman & Nguyen, 1998]). Then, under Assumption I,

(i) $(a|\Omega)_J \le_{HPL} (c|\Omega)$ iff $(a|\Omega)_J \le_{CPL} (c|\Omega)$ iff $\&(a_J) \le c$
iff $\lim\limits_{t_J \to 1} \big(\mathrm{minconc}((a|\Omega)_J;(c|\Omega))(t_J)\big) = 1$, etc.

(ii) $(a|b)_J \le_{CPL} (c|d)$ iff $(b\Rightarrow a \mid \Omega)_J \le_{HPL} (d\Rightarrow c|\Omega)$ iff $(b\Rightarrow a \mid \Omega)_J \le_{CPL} (d\Rightarrow c|\Omega)$
iff $\lim\limits_{t_J \to 1} \big(\mathrm{minconc}((b\Rightarrow a)_J;d\Rightarrow c)(t_J)\big) = 1$, etc.

Proof: (i): In Theorems 4.2 and 4.4 simply set all antecedents $b_j = \Omega = d$, j in J, etc.
(ii): Apply (i) above to Theorem 4.4. ■

4.4 *A Listing of Some Patterned Deduction Schemes for Application to CPL and/or HPL*

Next, consider the following list of some well-known patterned potential deductions given in linguistic form (see, e.g., [Pearl, 1988], [Dubois & Prade, 1993], and [Kraus *et al.*, 1990]. For any propositions a, b, b_j, c, d,…in B :

1. *Disjunction (Adams' R4)*: "if b, then a" and "if c, then a" <u>deduces</u> "if (b ∨ c), then a"
   This implies *Sure Thing Principle*: "if bc, then a" and "if bc′, then a" <u>deduces</u> "if b, then a" .
   In turn, this specializes to: *Disjoint Antecedents*: "if b, then a" and "if b′, then a" <u>deduces</u> a
2. *Bayes (Adams' R5)*: "if b, then a" and "if ab, then c" <u>deduces</u> "if b, then c"
3. *Cautious Monotonicity (or Triangle or Adams' R6)*: "if b, then a" and "if b, then c" <u>deduces</u> "if bc, then a", as well as "if ab, then c"
4. *PSCEA Order (Adams' R7)*: Under the assumption that ab ≤ cd and c′d ≤ a′b: "if b, then a" <u>deduces</u> "if d, then c". This implies with no assumptions:

"if b, then ac" <u>deduces</u> "if b, then c" , as well as "if ab, then c";
"if ab, then c" <u>deduces</u> "if b, then c"

5. *Reflexivity (Identity)*: "if b, then a" <u>deduces</u> "if b, then a"
6. *Cut*: "if b, then a" and "if ab, then c" <u>deduces</u> "if b, then ca", as well as "if b, then c"
7. *Exceptions*: "if bc, then a" and "if b, then a′" <u>deduces</u> "if b, then c′"
8. *Equivalence*: "if b, then a" and "if a, then b" <u>deduces</u> $a \Leftrightarrow b$
9. *Strict Modus Ponens*: "if b, then a" and b <u>deduces</u> ab , as well as a
10. *General Modus Ponens*: "if $(b \vee c)$, then a" and b <u>deduces</u> ab, as well as a
11. *Conditional Bound 1*: "if b, then a" <u>deduces</u> $b \Rightarrow a$
12. *Conditional Bound 2*: ab <u>deduces</u> "if b, then a"
13. *Transitivity-Syllogism*:"if c, then b" and "if b, then a" <u>deduces</u> "if c, then a"
14. *Contraposition*: "if b, then a" <u>deduces</u> "if not a, then not b"
15. *Positive conjunction*:"if b, then a" and "if c, then a" <u>deduces</u> "if bc, then a"
16. *Strengthening of Antecedents*: "if b, then a" <u>deduces</u> "if bc, then a"
17. *Penguin Triangle*: "if b, then a" and "if c, then b" and "if c, then d" and "if d, then a′b" <u>deduces</u> "if c, then a′"
18. *Modified Penguin Triangle*: Under the assumption $d \le a'b$: "if b, then a" and "if c, then b" and "if c, then d" <u>deduces</u> "if c, then a′"
19. *Consequent 1*: if b, then a" <u>deduces</u> a
20. *Consequent 2*: "if b, then a" <u>deduces</u> b
21. *Consequent 3*: a <u>deduces</u> "if b, then a"
22. *Consequent 4*: b <u>deduces</u> "if b, then a"
23. *Nixon Diamond*: "if c, then ab" and "if a, then d" and if b, then d′" <u>deduces</u> "if c, then d"
24. *Reverse Conditional Bound 1*: $b \Rightarrow a$ <u>deduces</u> "if b, then a"
25. *Reverse Conditional Bound 2*: "if b, then a" <u>deduces</u> ab
26. *Abduction*: "if b, then a" and a <u>deduces</u> b
27. *Induction*: Under the assumption that J is a finite index set, all b_j, j in J are mutually disjoint, and $\vee (b_J c) < c$: $\underset{\text{j in J}}{\text{And}}$ "if $b_j c$, then a" <u>deduces</u> "if c, then a"
28. *Augmented Induction*: Under the assumption that J is a finite index set, all b_j, j in J are mutually disjoint, and $\vee(b_J)c < c$: $\underset{\text{j in J}}{\text{And}}$ "if $b_j c$, then a" and " if c, then $\vee(b_J)$" <u>deduces</u> "if c, then a"

As before, we consider two basic interpretations of "if then": Approach 1 via the material conditional and MCL = CPL and Approach 2, via conditional probability and HPL. Table 4.1 in Section 4.5 shows which of the deduction schemes are valid for either approach, utilizing in a straightforward way the basic criteria in Theorem 4.4 (for MCL = CPL) and Theorem 4.2 (for HPL).

4.5 *Discrepancies between CPL and HPL and Effect on Choice of PL as a Basic Measure of Uncertainty*

Despite many reasonable properties holding for HPL, such as indicated in Theorem 4.2 and by "YES" in the last column of Table 4.1 below, unfortunately a number of intuitively appealing deduction schemes are *invalid* for HPL, in general, including the very natural schemes 13-16, which are all valid for MCL. (Details on the failure of the one of the most important deduction

No. and Name of Possible Deduction Scheme	**Valid for CPL ?**	**Valid for HPL ?**
1. Disjunction	YES	YES
2. Bayes	YES	YES
3. Cautious Monotonicity	YES	YES
4. PSCEA Order	YES	YES
5. Reflexivity	YES	YES
6. Cut	YES	YES
7. Exceptions	YES	YES
8. Equivalence	YES	YES
9. Strict Modus Ponens	YES	YES
10. General Modus Ponens	YES	YES
11. Conditional Bounds 1	YES	YES
12. Conditional Bounds 2	YES	YES
13. Transitivity-Syllogism	YES	NO
14. Contraposition	YES	NO
15. Positive Conjunction	YES	NO
16. Strengthening Anteced.	YES	NO
17. Penguin Triangle	NO	NO
18. Modified Penguin Triangle	YES	YES
19. Consequent 1	NO	NO
20. Consequent 2	NO	NO
21. Consequent 3	YES	NO
22. Consequent 4	NO	NO
23. Nixon Diamond	YES	NO
24. Reverse Cond. Bnd. 1	YES	NO
25. Reverse Cond. Bnd. 2	NO	NO
26. Abduction	NO	NO
27. Induction	NO	NO
28. Augmented Induction	YES	YES

Table 4.1. Validity or Non-Validity of Various Deduction Schemes for CPL and HPL.

schemes, transitivity-syllogism (no. 13), is provided in Section 5.) This shows a basic disconnect between HPL and CL; yet it is generally purported that probability theory is a natural extension of CL to account for uncertainties and errors. In fact, PL has been characterized as the unique approach to the modeling of propositions with uncertainty, among all possible approaches, satisfying certain natural relations. (See, e.g., [Aczél, 1966], [Cox, 1946, 1979] -- but see Halpern's negative claims for the characterization for finite domains [Halpern, 1999], and more recently, P. Snow's counter claims to that [Snow, 1998]. In [Lindley, 1982] the classical two person zero sum game characterization of PL was extended as originally given, e.g., in [DeFinetti, 1974], where the latter considered sums of squared loss between the 1-0 evaluation of occurrence vs. non-occurrence of membership functions of collections of ordinary events vs. corresponding uncertainty values pre-assigned to their occurrences-non-occurrences. Lindley considered a much more general class of loss functions and showed the set of all admissible uncertainty functions, for each choice of such a loss function, coincides with a corresponding nontrivial nondecreasing fixed *functional composition*of arbitrary (finitely additive conditional) probability measures. By suitably varying such loss functions, the functional compositions can vary arbitrarily over essentially all nondecreasing functions (composed with arbitrary finitely additive conditional probability measures). But, each such functional composition, in general, is quite distinct from the identity composition on probabilities, and only reduces to actual probabilities when the loss function becomes summed squared loss. However, Lindley's conclusion from this sound mathematical result is another matter: that only probability need be considered as the reasonable measure of uncertainty. This is due to his confusion of probability with functions of probability, which include certain Dempster-Shafer functions, as well as, in an asymptotic sense, FL-related functions. (See, the second part of [Lindley, 1982] as well as, e.g., [Goodman *et al.*, 1991a].) On the other hand, whenever the REA problem is solvable for such functional compositions, indeed those compositions themselves become probabilities in a higher order space. Even if a superior universal characterization of the usual measure-theoretic definition of PL can be achieved, relative to all competing approaches to modeling uncertainty, the very basis for connecting PL to real-world phenomena, i.e., *randomness*, still appears to be a problem of great difficulty. See, e.g., [Ayton *et al.*, 1989, 1991] for the psychological aspect of the issue, [Kyburg, 1983] for an attempt at an axiomatic approach, and [Kalman, 1994] for an overall negative conclusion concerning use of probability for the real-world.

The above-mentioned perplexing discrepancies between CL and PL or HPL, as well as the basic controversies surrounding probability have led to intensive investigations and the development *of non-monotonic and default* logics as outlined, e.g., in [Kraus *et al.*, 1990], [Pearl, 1988, 1990] and [Pearl & Goldszmidt, 1996], or other nonstandard and "hybrid" logics as in [Dubois & Prade, 1993, 1996]. Essentially, a monotonic logic is one in which once a premise set of events deduces a particular conclusion, any addition to the premise set will continue to deduce the same conclusion. A non-monotonic logic allows for the possibility of switching back and forth between deducing or not deducing a conclusion, depending on what collection of new events (or conditional events) are added to the original premise collection. However, all is not lost, in that application of SOP to a number of deduction schemes, including all three mentioned above (schemes 13-16), leads to valid deductions in the asymptotic sense that when the premise probability constraints are required to be "high", so will the conclusion probabilities be "high", but on the average. More details of the resulting non-monotonic logic can also be found here in Section 6 and [Goodman & Nguyen b]. Moreover, by again straightforward (but rather tedious use in some deduction schemes), posterior expectations for deductions can be obtained for all of the above-mentioned deduction schemes (and more), not just in the asymptotic sense, but for a wide variety of constrained probability levels, thereby allowing practical quantitative implementations. (See again Section 6.)

5. *A Motivating Example: More Details on the Problem of Transitivity-Syllogism Deduction*

In a very natural sense, it can be stated that probability theory extends classical logic to deal with situations where one seeks to quantify the percentage of time an event occurs or does not occur, or equivalently, where some uncertainty may be present in the occurrence of events. Yet, as stated earlier, there exists a gap between many of the concepts in classical logic and seemingly corresponding ones in probability theory. One basic example that illustrates this somewhat surprising lack of continuity between classical logic and probability theory is the transitivity-syllogism problem (deduction scheme 13 of Table 4.1) which can well occur at any given C2 decision node: Let a, b, c be three events of interest, all lying in some boolean algebra *B*. If $a \le b$ and $b \le c$, then clearly $a \le c$. This is essentially the same form as the classical "barbara" syllogism whose origins go back to Ancient Greece with the well known example "all men are mortal", "I am a man"; therefore, "I am mortal". (See, e.g., [Prior *et al*., 1968] for a history of this problem.) However, let us add a little uncertainty between a and b and between b and c, such as in the real-world interpretation

a = "enemy ships will be stationed in Sector A tomorrow night",
b = "fog condition B is expected to hold tomorrow afternoon and night",
c = "it is expected tonight for war condition C to be declared and for the temperature to be around 40 degrees F".

Denoting P(.) for probability, suppose, for simplicity, at this node, the *only* information that is received "now" is the premise P(if b, then a) = s, P(if c, then b) = t, where s, t are somewhat high values, such as 0.92, 0.85, respectively, based on previous intelligence-gathering sources and past performances. Otherwise, we have no knowledge of the subevent relations among the three events. The individuals at this decision node must, as soon as possible, transmit their assessment of the key value P(if c, then a) in a secure one-way manner to another specified node. What should they conclude P(if c, then a) is ? We first present two well-tried approaches to this problem which at first glance should produce satisfactory solutions to the quandary, but instead yield great difficulties. In all of the following we can reasonably suppose that a, b, c all lie in common boolean algebra B relative to probability space (Ω, B, P).

Approach 1. One models P(if b, then a) – and analogously, P(if c, then b) and P(if c, then a) – by the material conditional operator Þ , as discussed in the previous section. (Actually, this is the approach of [Nilsson, 1986] and others.)

From the very definition of the material conditional operator, by identifying:

$$\text{“if b, then a” with } b \Rightarrow a\text{ ; “if c, then b” with } c \Rightarrow b\text{ ; “if c, then a” with } c \Rightarrow a\text{ ,} \tag{5.1}$$

we have from eq.(5.1) and the basic assumption,

$$P(\text{if b, then a}) = P(b \Rightarrow a) = 0.92\ ,\quad P(\text{if c, then b}) = P(c \Rightarrow b) = 0.85\ , \tag{5.2i}$$

$$P(\text{if c, then a}) = P(c \Rightarrow a) = \ ? \tag{5.2ii}$$

Note that now a, b, c, b⇒a, c⇒b, c⇒a all are in B and subject to the usual laws of boolean algebra (or classical logic); it follows that the conjunction

$$(b \Rightarrow a)\&(c \Rightarrow b) = b'c' \vee abc' \vee abc = b'c' \vee ab\Omega = b'c' \vee ab \le c' \vee a = c \Rightarrow a\ . \tag{5.3}$$

Although according to our comments in Section 4 this is sufficient to show MCL deduction (see eq.(4.2) and ensuing remarks), consider some further details: Applying the usual monotonicity law of probability to eq.(5.3), we have, for any choice of P,

$$P((b \Rightarrow a)\&(c \Rightarrow b)) \le P(c \Rightarrow a). \tag{5.4}$$

Next, we apply another basic property of all probability spaces, the Fréchet-Hailperin bounds which provide the tightest upper and lower bounds for the probability of the conjunction or disjunction of events in terms of the probabilities of the individual (or marginal) events (see, e.g., [Hailperin, 1965, 1984]. For the case of two events, these bounds reduce to simply

$$\max(P(a)+P(b)-1,0) \le P(a\&b) \le \min(P(a), P(b)) \tag{5.5i}$$

$$(\le)\quad \max(P(a), P(b)) \le P(a \vee b) \le \min(P(a)+P(b),1), \tag{5.5ii}$$

where, in general, strict inequality holds and, e.g., equality holds on right-hand side of eq.(5.5i) iff equality holds on left-hand side of eq.(5.5ii) iff (slightly abusing notation) either $P(a \leq b$ or $b \leq a) = 1$, i.e., iff $P(ab') = 0$ or $P(a'b) = 0$.

Considering the left-hand side of eq.(5.5i) and replacing there a by $b \Rightarrow a$ and b by $c \Rightarrow b$, shows that eq.(5.4) expands to include a lower bound

$$\max(P(b \Rightarrow a) + P(c \Rightarrow b) - 1, 0) \leq P((b \Rightarrow a)\&(c \Rightarrow b)) \leq P(c \Rightarrow a). \quad (5.6)$$

Next, using the premise levels s, t, together with eq.(5.1) in eq.(5.6), finally shows

$$\max(s + t - 1, 0) \leq P(c \Rightarrow a) = P(\text{if } c, \text{ then } a) \quad (5.7)$$

Thus, for sufficiently large values of s and t, such as above, where $s = 0.92$, $t = 0.85$, eq.(5.7) shows that

$$0.77 \leq P(\text{if } c, \text{ then } a). \quad (5.8)$$

This is certainly a reasonable conclusion, agreeing with "commonsense" reasoning that if the pattern of probability of b to c and a to b are high, so should be the probability of a to c – though possibly somewhat degraded in value as seen here – provided we have no other information on the relations of a, b, c to each other. Thus, it would seem that the material conditional operator satisfactorily solves our problem. However, though the mathematical concepts leading from eqs.(5.1) to (5.8) are quite sound, one important factor is missing: Note, that while it is possible to have $P(b \Rightarrow a) = 0.92$ and $P(c \Rightarrow b) = 0.85$ – and consistent with eq.(5.8), one must have $0.77 \leq P(c \Rightarrow a)$ – the actual values of P(b), P(ab), P(c), P(bc) can be, e.g.,

$$P(b) = 0.10, \; P(ab) = 0.02, \; P(c) = 0.20, \; P(bc) = 0.05, \quad (5.9)$$

implying

$$P(a'b) = 0.08, \quad P(b'c) = 0.15). \quad (5.10)$$

Eq.(5.9) holds compatibly with all of the above equations by simple use of eq.(2.2) applied to eq.(5.2i) But, by being consistent with respect to the interpretation of P(if b, then a) via $P(b \Rightarrow a)$, we should now interpret

$$P(\text{if } b, \text{ then } a') = P(b \Rightarrow a'), \quad P(\text{if } c, \text{ then } b') = P(c \Rightarrow b'). \quad (5.11)$$

Thus, computing the values in eq.(5.11), via eqs.(5.9), (5.10), readily shows

$$P(\text{if } b, \text{ then } a') = 0.98, \quad P(\text{if } c, \text{ then } b') = 0.97. \quad (5.12)$$

Comparing eq.(5.12) with the basic assumption given in eq.(5.2) shows

$$P(\text{if } b, \text{ then } a) = 0.92; \; P(\text{if } b, \text{ then } a') = 0.98; P(\text{if } c, \text{ then } b) = 0.85; \; P(\text{if } c, \text{ then } b') = 0.97. \quad (5.13)$$

Eq.(5.13) illustrates a most disquieting situation: *there is relatively little difference between the probability of the negation of the consequents and the affirmation of the consequents for the same antecedents*. Similar difficulties arise, if P(b) and P(c) in eq.(5.9) were any relatively small values, with P(ab), P(bc) determined accordingly. On the other hand, if in place of eq.(5.9), the antecedent probabilities were medium or even large in value, it is easily seen that the differences between the probabilities of the negation and affirmation of consequents, for common antecedent grows larger, producing a more satisfactory result.

Thus, we must conclude that while the use of the material conditional in interpreting the probability of a conditional statement provides a reasonable formal solution to the transitivity-syllogism problem overall, it can be a poor model of the individual conditional expressions due to its insensitivity to consequents when antecedents have low probabilities.

Approach 2. One models P(if b, then a) – and analogously, P(if c, then b) and P(if c, then a) – by the well-known conditional probability counterparts (as considered also in the last section)

$$P(\text{if b, then a}) = P(a|b)\ ,\quad P(\text{if c, then b}) = P(b|c),\ \ P(\text{if c, then a}) = P(a|c). \tag{5.14}$$

Before attempting to address the transitivity-syllogism problem, we note the obvious difference between this interpretation and the material conditional one: no discrepancy can arise as in Approach 1 given in eqs. (5.9)-(5.13), because of the basic properties

$$P(a'|b) = 1-P(a|b),\ \ P(b'|c) = 1-P(b|c)\ ,\ P(a'|c) = 1-P(a|c), \tag{5.15}$$

as well as the fact that conditional probability is sensitive only to the ratio of (conjoined) consequent probability to antecedent probability -- it is not essentially forced to approach unity when the antecedent probabilities grow small.

Returning now to the main problem at hand, consider Figures 5.1 and 5.2, where probability space (Ω,*B*,P) and events a, b, c in *B* are interpreted so that in both figures below: Ω represents a solid rectangle in two dimensions; a, b, c are solid triangles of different sizes and shapes; P is uniform probability measure, and thus all conditional probabilities are proportional to the relative areas of the triangles involved.

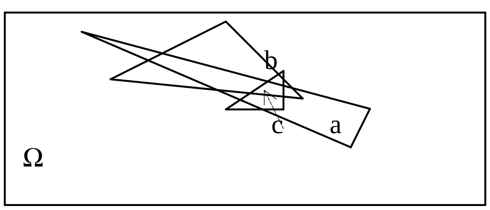

Figure 5.1. Situation with All Moderate Values for P(a|b), P(b|c), P(a|c)

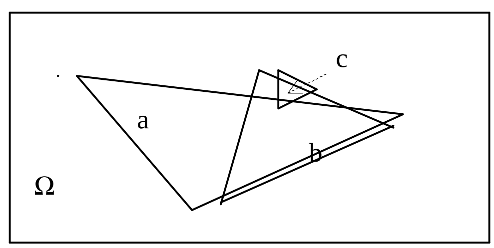

Figure 5.2. Situation with High Values for P(a|b), P(b|c), but P(a|c) Low

Thus, the above illustrations show how events a, b, c and probabilities P can be chosen so that – by a simple continuity argument – one can pick situations where P(a|b), P(b|c) can be moderate or high and P(a|c) can take any value between 0 and 1, depending on the construction. Note also that in these two figures, while the underlying probability distribution over Ω is uniform, it does *not* induce equal probabilities over the (at most 8 relative) atoms formed by intersecting various combinations of affirming or negating a, b, c. Indeed, the particular situations the figures describe are essentially determined by the non-uniform pattern of such values.

Note that certain situations can arise involving additional general relations among a, b, c which will yield a satisfactory solution to the problem. For example, it is easily shown that if a, b, c in *B* and P are such that there are real numbers $0 < \delta, \varepsilon_1, \varepsilon_2 < 1$ so that

$$P(abc') \leq \ \delta\cdot P(abc)\ \ ,\quad P(a|b) \geq 1-\varepsilon_1,\quad P(b|c) \geq 1-\ \varepsilon_2\ , \tag{5.16}$$

then

$$(1+\delta)P(a \mid bc) \geq (P(abc) + P(abc'))/ P(b) = P(a|b) \geq 1-\varepsilon_1 \ .$$

Hence,

$$P(a \mid bc) \geq (1-\varepsilon_1)/(1+\delta). \tag{5.17}$$

Also,

$$P(a|c) \geq P(ab \mid c) = P(a |bc)P(b|c). \tag{5.18}$$

Then, combining eqs.(5.16)-(5.18), yields finally

$$P(a|c) \geq (1-\varepsilon_1)(1-\varepsilon_2) /(1+\delta), \tag{5.19}$$

which obviously includes the special case of $b \leq c$, whereupon δ may be chosen as 0 in eq.(5.19). Other than special cases such as the above, the situation in Approach 2 is also not satisfactory: No conditional probability counterpart of eq.(5.7) can exist in general. On the other hand, Approach 2 does not suffer from the difficulties of Approach 1 relative to negations discussed above.

Thus, both seeming "natural" approaches to dealing with the transitivity-syllogism issue are unsatisfactory, for two different reasons. In a similar vein, one can provide other examples where the material conditional "works" compatibly with "commonsense" reasoning, but its probability counterpart doesn't. Some examples of this include for any a, b, c in *B*, apropos to rows 14-16 of Table 4.1:

Contraposition: When premise P(if b, then a) is high, we want deduction P(if a′, then b′) high
Positive Conjunction: When premise collection P(if b, then a), P(if c, then a) are both high, we want deduction P(a | bc) high
Strengthening Antecedent: When premise P(a|b) high, we want deduction P(a | bc) high

The next section provides a means for addressing the quandary developed here, not only for the transitive syllogism problem, but also for a wide variety of deduction problems, including the above examples and, in fact, the listing of the 28 potential deduction schemes in Table 4.1.

6. *Summary of Cognitive Probability Logic*

6.1 *Introduction and Basis for Choice*

The results of Section 6 stem from a recent mathematical breakthrough, documented more fully in [Goodman & Nguyen, b], resulting in a new logic, called *Cognitive Probability Logic* (or the *Logic of Expectations / Averages* -- EPL). In fact, the underlying structure of this logic can be shown to be actually a *natural averaging modification of Adams' HPL* discussed in Section 4. Consequently, a number of long-standing conflicts between ordinary probability logic and "commonsense" reasoning are resolved for the first time, including the well-known transitivity-syllogism problem discussed previously. CPL draws heavily upon SOP, as outlined in Section 3.4. In addition, practical applications to the systematic understanding of incoming linguistic-based information can also be obtained by use of the above-mentioned techniques, together with one-point random set coverage representations of fuzzy logic. (See also [Goodman & Nguyen, 1999] for more details.) Because of the already long length of this paper, most of the long proofs involved are either omitted, or very much abridged (as in the case of the transitivity-syllogism deduction scheme in Section 6.2).

The novelty of the result discussed in this presentation is that one can reconcile directly both commonsense reasoning and classical probability in a fully rigorous and efficiently implementable way. The "trick" in accomplishing this involves the basic assumption that, in actuality, commonsense reasoning in a large number of cases – and certainly including transitivity as well as the other above-named types of deductions – is actually based on a fixed abstracted pattern, not on the specific received information. Thus, for the transitivity syllogism problem, when one concludes that it is natural to have P(a|c) to be relatively high when both P(a|b) and P(b|c) are high, *one is generally not considering what the specific events a, b, c stand for, nor the particular type of probability measure P is, but simply what the pattern is, provided no other possible relevant information is present.* (If it were, then one would add other constraints, but again stop at some fixed abstracted pattern for further analysis.) The transitivity pattern remains valid for completely different situations, such as for a = enemy ship A, b= ship class B, c = ship class C, with uncertainty measured via the naturally corresponding conditional probabilities. In turn, based on the above assumption, it is natural to seek a mathematical model of the situation where the abstracted pattern of the knowledge structure is taken into account, yet does not contradict standard probability logic. The apparent answer to this problem was found by computing in place of a specific P(a|c) – which, as stated above, one could not obtain in general anyway, unless one knew all of P (relative to the eight atoms) – an *averaged value* of P(a|c), taking into account the variability of possible choices of P, as well as a, b, c, i.e., considering a second order probability approach – as opposed to the usual first order one, where a specific P is used and analyzed.

At this point, it is of some interest to see, via brief survey, how the transitivity-syllogism and related problems have been treated from a cognitive psychology viewpoint. [Freeman, 1993] discusses various methods of implementing reasoning and commonsense reasoning, syllogisms, but interestingly does not suggest any randomization of patterns in syllogistic reasoning. [Johnson-Laird, 1983], especially Chapters 2-5, discusses commonsense reasoning and claims many mistakes are made; he suggests non-truth functional connectives as a possibility for modeling such, but does not present a specific psychological theory of reasoning that treats syllogistic reasoning as well as probability concepts. However, Johnson-Laird discusses separately syllogistic reasoning in depth and describes a number of numerical experiments run by him and other colleagues apropos to percentage of (CL) "correct" conclusions and how to reason syllogistically; but again, he presents no tie-ins with probability logic, even though he utilizes probability concepts in other sections of his work. [Mayer, 1992] briefly considers probability, but only applied to the famous Tversky & Kahneman experiments on probability judgment and decisions, not relating to his presentation on categorical and conditional reasoning (Chapter 5, pp. 114-151). However, he does discuss a number of experiments concerning percentage of times deductions (involving conditional or unconditional statements) are compatible with CL conclusions. [Solso, 1979], especially Chapter 14, pp. 371-382, treats syllogistic reasoning and problem solving, but as so with the other above-mentioned writers, only considers the CL interpretation.

Returning to the SOP approach to the modeling of deductions, one follows the general three or four step procedure outlined in Section 3.4, formally identifying P as an appropriately chosen random vector -- most naturally assumed to be uniformly distributed over the relative atoms comprising the events of interest, but more generally, as in the treatment of the "Judy Benjamin" updating problem mentioned earlier, assumed to be a dirichlet distribution [Goodman & Nguyen, a]. At least under the joint uniform distribution assumption, though the effort involved in achieving closed-form expressions involves a considerable number of computations and specialized probability techniques, it is not intractable for any of the above-mentioned deduction forms in Table 6.1, as well as for many others.

6.2 *Definition and Basic Properties of EPL*

For any variable probability space (Ω,B,P), using again multivariable notation, where, e.g., $P(a|b)_J = (P(a_j|b_j))_{j \text{ in } J}$, consider any nontrivial finite premise collection $(a|b)_J$ in $(B_o{}^*)^J$ and nontrivial (potential) conclusion event $(c|d)$ in B_o. Then, referring to eq.(4.6), in place of the function $\text{minconc}((a|b)_J;(c|d)): [0,1]^J \to [0,1]$, we seek to evaluate

$$\text{meanconc}((a|b)_J;(c|d)):[0,1]^J \to [0,1], \tag{6.1}$$

where now, for any t_J in [0,1]J, comparing with the definition in eq.(4.5),

$$\text{minconc}((a|b)_J;(c|d))(t_J) \le \text{meanconc}((a|b)_J;(c|d))(t_J) = E_{\mathsf{P}}(P(c|d) \mid P(a|b)_J = t_J) . \tag{6.2}$$

Here, P is some random vector representing the distribution of probability spaces (Ω,B,P). More specifically, one carries out the SOP procedure outlined in Section 3.4, where the atoms generated by the collection of all consequents and antecedents involved in the potential deduction are first considered. P is identified as the random probability function or random vector resulting from evaluating random probability measure P over all of the atoms, assumed to be a joint uniform one. If under Assumption II, $(a|b)_J \le_{HPL} (c|d)$, then comparing eq.(4.10) with eq.(6.2) shows

$$\underset{(t_J \uparrow 1_J,\ \text{uniformly})}{\text{limit}} (\ \text{minconc}_o((a|b)_J;(c|d))\ (t_J)\ ) = 1 \quad \text{implies} \quad \underset{(t_J \uparrow 1_J,\ \text{uniformly})}{\text{limit}} (\ \text{meanconc}((a|b)_J;(c|d))(t_J)\ ) = 1. \tag{6.3}$$

Analogous to the definition of HPL deduction, we denote EPL deduction as

$$(a|b)_J \le_{EPL} (c|d)) \ \text{ iff } \ \underset{(t_J \uparrow 1_J,\ \text{uniformly})}{\text{limit}} (\ \text{meanconc}((a|b)_J;(c|d))(t_J)\ ) = 1. \tag{6.4}$$

Note also that for any random variable Z in [0,1] – as is P(c|d) as a function of random vector P -- we have the fundamental result that since $Z^2 \le Z$ over [0,1],

$$\text{Var}(Z) = E(Z^2) - (E(Z))^2 \le E(Z) - (E(Z))^2 = (E(Z))(1-E(Z)). \tag{6.5}$$

Hence,

$$(a|b)_J \le_{EPL} (c|d)) \ \text{implies} \ \underset{(t_J \uparrow 1_J,\ \text{uniformly})}{\text{limit}} (\ \text{Var}_{\mathsf{P}}(P(c|d) \mid P(a|b)_J = t_J)) = 0 , \tag{6.6}$$

which, in turn, from standard results (such as via the Chebychev Inequality) implies

$$\underset{(t_J \uparrow 1_J,\ \text{uniformly})}{\text{limit in probability}} \left(\mathsf{P}(P(c|d) \mid P(a|b)_J = t_J) \right) = 1. \tag{6.7}$$

The next result shows that EPL lies properly in between HPL and CPL:

Theorem 6.1 Under Assumption I,
(i) $(a|b)_J \leq_{HPL} (c|d)$ implies properly $(a|b) \leq_{EPL} (c|d)$,
(ii) $(a|b)_J \leq_{EPL} (c|d)$ implies properly $(a|b) \leq_{CPL} (c|d)$.

Proof: (i) follows immediately from eq.(6.3). (ii) is shown in [Goodman & Nguyen, b], independent of the general proof in [Bamber]. ■

Table 6.1 (Section 6.4) illustrates the relations in Theorem 6.1 by determining not only EPL deduction (or non-deduction) via the limiting froms in the definition in eq.(6.4), but, in fact, obtaining closed-form expressions for meanconc((a|b)J;(c|d)(t_J), for nonlimiting t_J in $[0,1]^J$, for many of the 28 types of deduction schemes considered there (and originally in Table 4.1).

6.3 *Outline of Proof of Form of Meanconc for Transitivity-Syllogism*

For the syllogism problem, e.g., P(a|b), P(b|c) correspond to scalar functions of P, and therefore correspond to random variables. Then, in turn, the constraints P(a|b) = s, P(b|c) = t , for any fixed real s, t in the unit interval, become corresponding events in the second order probability space over which P occurs. In this case, there are eight relative atoms, formed in the usual way out of events a, b, c in *B*, with respect to variable probability space (Ω,*B*,P) yielding the corresponding atoms

$$\alpha_1 = abc,\ \alpha_2 = abc',\ \alpha_3 = ab'c,\ \alpha_4 = ab'c',\ \alpha_5 = a'bc,\ \alpha_6 = a'bc',\ \alpha_7 = a'b'c,\ \alpha_8 = a'b'c' \qquad (6.8)$$

forming the vector of atoms

$$\underline{\alpha} = (\alpha_1,\ldots,\alpha_7,\alpha_8). \qquad (6.9)$$

Their variable evaluation by any one choice of P as a vector of probabilities of the atoms is

$$\underline{y}^{(1)} = \underline{P}(\underline{a}) = (y_1,\ldots,y_7,y_8) = (P(\alpha_1),\ldots,P(\alpha_7),P(\alpha_8)), \qquad (6.10)$$

with the obvious constraints

$$0 \leq y_j \leq 1,\ j=1,\ldots,7,8; \quad y_1+\ldots+y_7+y_8 = 1. \qquad (6.11)$$

As (Ω,*B*,P) varies, so does a, b, c, P, and, in turn, $\underline{\alpha}$ and $\underline{Y}$, clearly covering the region

$$Q_{(8)} = \{\underline{y}^{(1)};\ \underline{y}^{(1)} \text{ given in eq.(6.10) satisfies eq.(6.13) arbitrarily}). \qquad (6.12)$$

From now on, it is more convenient to eliminate the variable y_8, using eq.(6.11), and only consider the corresponding region Q_7, the full 7-dimensional simplex (see eq.(3.41))

$$Q_7 = \{\underline{y}: \underline{y} \text{ given in eq.(6.12) satisfies eq.(6.14) arbitrarily}\}, \qquad (6.13)$$

where

$$\underline{y} = (y_1,\ldots y_6,y_7), \qquad 0 \leq y_j \leq 1\ ,\ j=1,\ldots,6,7; \quad y_1+\ldots+y_6+y_7 \leq 1. \qquad (6.14)$$

In turn, the variables (random or deterministic) representing the two premise conditional probabilities P(a|b), P(b|c) as functions of $\underline{y}$ are, using eqs.(6.8) and (6.10),

$$P(a|b) = (y_1 + y_2) / (y_1 + y_2 + y_5 + y_6)\ , \quad P(b|c) = (y_1 + y_5) / (y_1 + y_3 + y_5 + y_7), \qquad (6.15)$$

while the variable representing the conclusion probability is

$$P(a|c) = (y_1 + y_3) / ((y_1 + y_3 + y_5 + y_7). \qquad (6.16)$$

Then, using eqs.(6.15), (6.16), the premise constraints become, for any choice of real fixed s, t, with $0.5 \le s, t < 1$,

$$P(a|b) = s \text{ iff } (y_1 + y_2) / (y_1 + y_2 + y_5 + y_6) = s \text{ iff } (1-s)y_1 - sy_5 = -(1-s)y_2 + sy_6 ,$$
$$P(b|c) = t \text{ iff } (y_1 + y_5) / (y_1 + y_3 + y_5 + y_7) = t \text{ iff } (1-t)y_1 + (1-t)y_5 = ty_3 + ty_7 . \qquad (6.17)$$

Hence, if we now introduce $\underline{Y} = (Y_1,\ldots,Y_6,Y_7)$ as a random vector, whose typical outcome is denoted by $y = (y_1,\ldots,y_6,y_7)$ as in eq.(6.14) subject to the constraint of being in Q_7 given in eq.(6.13), the random vector and random variable counterparts of eqs.(6.11)-(6.13) all hold with the replacement of each y_j by Y_j. Denote the basic constraint event for the premise as

$$A_{s,t} = \{\underline{y}: \underline{y} \text{ in } Q_7 \text{ and } P(a|b) = s, P(b|c) = t\} = \{\underline{y}: \underline{y} \text{ in } Q_7 \text{ and } \underline{y} \text{ satisfies eq.(6.19)}\} \subset Q_7, \qquad (6.18)$$

where, using eq.(6.17),

$$(1-s)y_1 - sy_5 = -(1-s)y_2 + sy_6 ; \qquad (1-t)y_1 + (1-t)y_5 = ty_3 + ty_7. \qquad (6.19)$$

In turn, the simultaneous linear equations in eq.(6.19) are easily seen to be always solvable because of the choice of s, t, yielding the formal solution

$$y_1 = h_1(v(y_3,y_7),y_2,y_6) , \quad y_5 = h_5(v(y_3,y_7),y_2,y_6), \ \underline{y} \text{ arbitrary}, \qquad (6.20)$$

where the functions h_1, h_2, and v are defined as

$$v(y_3,y_7) = y_3 + y_7 , \quad h_1(v,y_2,y_6) = -(1-s)y_2 + sy_6 + (st/(1-t))v ,$$

$$h_5(v,y_2,y_6) = (1-s)y_2 - sy_6 + ((1-s)t/(1-t))v, \text{ for all } \underline{y} . \qquad (6.21)$$

Substituting eq.(6.21) into eq.(6.16) yields, up to requiring $\underline{y}$ to be also in Q_7: If $\underline{y}$ in $A_{s,t}$, then P(a|c) as a function of $\underline{y}$ in eq.(6.16) becomes, after simplifying,

$$P(a|c) = -(1-s)(1-t)(y_2/v) + s(1-t)(y_6/v) + st + (1-t)(y_3/v). \qquad (6.22)$$

Hence, indicating as before, the corresponding random variables Y_j to each possible outcome y_j, V to v, etc., eq.(6.22) shows

$$\text{meanconc}((a|b), (b|c);(a|c))(s,t)) = E_P[P(a|c) \mid A_{s,t}]$$

$$= -(1-s)(1-t)\cdot E_P[(Y_2/V) \mid A_{s,t}] + s(1-t) \cdot E_P[(Y_6/V) \mid A_{s,t}] + st + (1-t) \cdot E_P[(Y_3/V) \mid A_{s,t}]. \qquad (6.23)$$

Next, consider use of Theorem 3.1, where the original random vector $P = (Y_1,\ldots,Y_7)$ that is jointly uniformly distributed over Q_7 (actually $Dir(1_8)$, see eq.(3.40))is transformed to a new random vector $Q = (X_1,\ldots,X_7,X_8)$ over the positive 7-orthant. Theorem 3.1 then shows, $Y_j = X_j / X_o$, j=1,…,7, 8, where $X_o = X_1+\ldots+X_7+X_8$, $V = U/X_o$, $U = X_3 + X_7$, with all X_j being independently identically distributed as Gam(1,1), i.e., Expo(1), with common pdf g_1, given for non-zero values as

$$g_1(x) = e^{-x}, \text{ for all } x > 0. \qquad (6.24)$$

Then, these relations substituted into eq.(6.23) and taking into account the transformation's effect on $A_{s,t}$, and cancelling out X_o, where possible, yields

$$\text{meanconc}((a|b), (b|c);(a|c))(s,t)) = E_Q[P(a|c) \mid A_{s,t}]$$

$$= -(1-s)(1-t)\cdot E_Q[(X_2/U) \mid A_{s,t}] + s(1-t)\cdot E_Q[(X_6/U) \mid A_{s,t}] + st + (1-t)\cdot E_Q[(X_3/U) \mid A_{s,t}]. \quad (6.25)$$

Now, in terms of the X_j (and U)

$$A_{s,t} \text{ holds iff } X_1 = h_1(U,X_2,X_6) \text{ and } X_5 = h_5(U,X_2,X_6), \qquad (6.26)$$

with (deterministic) functions h_1 and h_5 the same as before. Note that eqs.(6.25), (6.26) show that the only random variables involved in the final computations for meanconc here are X_1, X_2, X_6 and $U = X_3+X_7$, with X_1, X_2, X_6, U, $W = X_3/U$ all being mutually independent (the latter two from each other by standard properties of gamma distributions). U is distributed as Gam(2,1), with pdf g_2, and $W = X_3/U$ is distributed beta(1,1) = Unif[0,1], with pdf g_o given for non-zero values as

$$g_2(u) = u\cdot e^{-u}, \text{ for all } u > 0; \quad g_o(w) = 1, \text{ for all } 0 < w < 1. \qquad (6.27)$$

In particular, the above remarks show the independence of W from $A_{s,t}$, whence

$$E_Q[W \mid A_{s,t}] = E_Q[W] = ½$$

and now eq.(6.25) becomes

$$\text{meanconc}((a|b), (b|c);(a|c))(s,t)) = E_Q[P(a|c) \mid A_{s,t}]$$

$$= -(1-s)(1-t)\cdot E_Q[(X_2/U) \mid A_{s,t}] + s(1-t)\cdot E_Q[(X_6/U) \mid A_{s,t}] + st + (1/2)(1-t). \qquad (6.28)$$

Thus, it remains to compute $E_Q[(X_j/U) \mid A_{s,t}]$, j = 2, 6. Note also by the independence of X_1 and X_5 from $h_j(U,X_2,X_6)$ and the fact that the non-zero pdf values for X_1 and X_6 is the positive real line, letting region

$$R_o = \{(u,x_2,x_6): 0 < u,x_2,x_6, \text{ such that } A_{s,t} \text{ holds with } x_j = h_j(u,x_2,x_6) > 0, j = 1, 5\}, \qquad (6.29)$$

$$E_Q((X_j/U) \mid A_{s,t}) = N_j(s,t) / D(s,t)\ , \qquad (6.30)$$

where, for convenience, indicating the pdfs involved generically by p(.,..,…) with appropriate subscripts, etc.,

$$N_j(s,t) = \int_{R_o} (x_j/u)\cdot p(A_{s,t} \mid u, x_2, x_6)\cdot p(u, x_2, x_6)\,du\,dx_2\,dx_6\ , \ j=2, 6, \qquad (6.31)$$

$$D(s,t) = Q(A_{s,t}) = \int_{R_o} p(A_{s,t} \mid u, x_2, x_6) \cdot p(u, x_2, x_6) du dx_2 dx_6 . \qquad (6.32)$$

Also, by independence of the relevant random variables, for outcomes (u, x_2, x_6) in R_o,

$$p(A_{s,t} \mid u,x_2,x_6) = g_1(h_1(u,x_2,x_6)).g_1(h_5(u,x_2,x_6)) = e^{-h_1(u,x_2,x_6)-h_5(u,x_2,x_6)} = e^{-(t/(1-t))u}. \qquad (6.33)$$

Utilizing further constraints required to satisfy R_o and produce positive $h_j(u,x_2,x_6)$, followed by a good deal of integral computations and intensive crosschecking,

$$N_2(s,t) = [s^2(1-s)t(1-t)(1-s(1-t) + t + t^2)] / [(s + t - st)(1 - s(1-t))^2] , \qquad (6.34)$$

$$N_6(s,t) = [st(1-s)^2(1-t)(s + (2-s)t + t^2)] / [(s+ t - st)^2(1- s(1-t))], \qquad (6.35)$$

$$D(s,t) = \left(st(1-s)(1-t)^2\left(t(1+2t) + [s(1-s)(1-t)(2 + 3t - t^2)]\right)\right) / \left[(s +t - st)^2(1- s(1-t))^2\right]. \qquad (6.36)$$

Finally, substituting eqs.(6.34)-(6.36) into eq.(6.30), and then into eq.(6.28), yields after further manipulations, for any real s, t, with $½ \le s, t < 1$,

$$\text{meanconc}((a|b), (b|c);(a|c))(s,t)) = E_P [P(a|c) \mid P(a|b) = s, P(b|c) = t] = st + (1-t)/2 - f(s,t), \qquad (6.37)$$

where f(s,t) is the additional correction term

$$f(s,t) = [s(1-s)(2s-1)t(1-t^2)] / [t +2t^2 + (s(1-s)(1-t)(2+3t-t^2))]. \qquad (6.38)$$

Inspection of eqs.(6.37) and (6.38) shows general agreement with the commonsense conclusion: the averaged value of P(a|c) *has value generally lower than both* P(a|b) *and* P(b|c)*, but is not that much lower, and as s, t approach unity, so does the averaged value of* P(a|c). *Furthermore, the expectation in eq.(6.37) may be interpreted as the averaged degree (over all worlds or probability assignments) to which premise set at level* (s, t), (P(a|b) = s and P(b|c) = t)*, yields the conclusion "if c then a" relative to a conditional probability interpretation*

6.4 *Tabulation of Both Limiting and Non-Limiting Evaluations of Meanconc for Various Deduction Schemes and Concluding Remarks*

The interested reader will find all relevant details omitted here in the sketch of the proof for meanconc for the transitivity-syllogism pattern in [Goodman & Nguyen, b]. In addition, detailed proofs for a number of other meanconc computations in Table 6.1 are also presented there. Apropos to Theorem 6.1, Table 6.1 below illustrates the proper implication aspect of HPL deduction relative to EPL, by noting, e.g., rows 13-16 (including transitivity syllogism, contraposition, etc.) among others; the table also emonstrates the proper implication aspect of EPL deduction relative to CPL, by noting row 23 (Nixon Diamond).

In summary, this paper has presented an introduction to a new type of logic, Cognitive Probability Logic (EPL) and its key associated function – meanconc – for calculating the degree of averaged validity of a potential conclusion with respect to given premises in the form of conditional or unconditional probability constraints for various nonlimiting levels.

Name and Number of Deduction Scheme $(a\|b)_J$ Potent. Deducing $(c\|d)$	**Given Levels of Premises: $P(a\|b)_J \geq t_J$, for minconc, $P(a\|b)_J = t_J$, for meanconc**	**Con-clus. (c|d)**	**Minconc($(a\|b)_J$; $(c\|d)$)(t_J)**	**Meanconc($(a\|b)_J$; $(c\|d)$)(t_J)**	**Valid For CPL ?**	**Valid for EPL ?**	**Valid for HPL ?**
1. Disjunction	P(a|b) = s, P(a|c) = t	$(a\|b\vee c)$	≥ max(s+t-1,0)	≥ max(s+t-1,0)	YES	YES	YES
2. Bayes	P(a|b) = s, P(c|ab) = t	(c|b)	≥ st	≥ st	YES	YES	YES
3. Cautious Monotonicity	P(a|b) = s, P(c|b) = t	(a|bc)	≥ max(s+t-1,0)	≥ max(s+t-1,0)	YES	YES	YES
4.PSCEA Order	P(a|b) = t, for $\varnothing < a < b$, $\varnothing < c < d$	(c|d)	≥ t	≥ t	YES	YES	YES
5. Reflexivity	P(a|b) = t	(a|b)	t	t	YES	YES	YES
6. Cut	P(a|b) = s, P(c| ab) = t	(ac|b)	≥ st	≥ st	YES	YES	YES
7. Exceptions	P(a | bc) =s, P(a′|b) = t	(c|b)	≥ max(s+t-1,0)	≥ max(s+t-1,0)	YES	YES	YES
8. Equival.	P(a|b) = s, P(b|a) = t	$a \Leftrightarrow b$	≥ st	(s+t)/[2(s+t-st)]	YES	YES	YES
9. Strict Modus Ponens	P(a|b) = s, P(b) = t	ab	st	st	YES	YES	YES
10.General Modus Ponens	$P(a\|b\vee c)$ = s, P(b) = t	ab	≥ st	st + (1-t)/2	YES	YES	YES
11.Condition. Bounds 1	P(a|b) = t	$b \Rightarrow a$	≥ t	(2+t)/3	YES	YES	YES
12.Condition. Bounds 2	P(ab) = t	b	≥ t	$2\frac{-t\log(t)-t(1-t)}{(1-t)^2}$	YES	YES	YES
13.Transitiv.-Syllogism	P(a|b) = s, P(b|c) = t	(a|c)	0	$\geq st + (1-t)/2 - \frac{(1-s)s(2s-1)(1-t^2)}{1+2t}$	YES	YES	NO
14.Contra-position	P(a|b) = t	(b|a)	0	$1/t + \frac{(1-t)\log(1-t)}{t^2}$	YES	YES	NO
15. Positive Conjunction	P(a|b) = t, P(a|c) = t	(a|bc)	0	$(1+t)/3 + [((1+t)(2-t)/(3t))\cdot\theta(t)]$, $\theta(t) = (t^2/4)[\log((2-t)/t)]/(1-t) - ((1-t)^2/4)\cdot\log((1+t)/(1-t)$	YES	YES	NO
16.Strengthen. Antecedent	P(a|b) = t	(a|bc)	0	approx. t (complicated, but in closed-form)	YES	YES	NO
17.Penguin Triangle	P(a|b) = r, P(b|c) = s, P(d|c) = t, P(a′b|d) = u	(a′|c)	0	?	NO	NO	NO
18.Modified Penguin Triangle	P(a|b) = r, P(b|c) = s, P(d|c) = t, $d \leq a'b$	(a′|c)	≥ max(s+t-1,0)	≥ max(s+t-1,0)	YES	YES	YES
19.Consequ. 1	P(a|b) = t	a	0	(1+t)/3	NO	NO	NO
20.Consequ. 2	P(a|b) = t	b	0	1/3	NO	NO	NO

21.Consequ. 3	$P(a) = t$	$(a\|b)$	0	$1/2)(1 + g(t))$, $g(t) = [(1-t)\cdot\log(1-t)] / t - (t\cdot\log(t)) / (1-t)$	YES	YES	NO
22. Consequ.4	$P(b) = t$	$(a\|b)$	0	1/2	NO	NO	NO
23 Nixon Diamond	$P(ab\|c) = s$, $P(d\|a) = t$, $P(d'\|b) = t$	$(d\|c)$	0	1/2	YES	NO	NO
24.Reverse Cond. Bnd. 1	$P(b\Rightarrow a) = t$	$(a\|b)$	0	$t + \frac{2(1 - t)\log(1 - t)}{t^2} + \frac{(1-t)(2+t)}{t}$	YES	YES	NO
25.Reverse Cond. Bnd. 2	$P(a\|b) = t$	ab	0	t/3	NO	NO	NO
26. Abduction	$P(a\|b) = s$, $P(a) = t$	b	0	If $s \geq t$: $t/(2s)$, If $s < t$: $\frac{t^3 s(1-t)^2}{2(t^2 - 2st + s)^2}$	NO	NO	NO
27. Induction	For $b_j c$ all disj. $\vee(b_J c) < c$: $P(a \mid b_j\&c) = t_j$, $j=1,\ldots,n$;	$(a\|c)$	0	?	NO	NO	NO
28.Augmented Induction	For $b_j c$ all disj. $\vee(b_J c) < c$: $P(a \mid b_j\&c) = t_j$, $j=1,\ldots,n$; $P(\vee(b_J) \mid c) = s$	$(a\|c)$	$\geq \Pi(t_J) - (1-s)$	$\geq \Pi(t_J) - (1-s)$	YES	YES	YES

Table 6.1. Tabulation of minconc and meanconc Functions and Listing of Validity-Nonvalidity of 28 Selected Potential Deduction Schemes with Respect to CPL, EPL, and HPL.

Evaluation of EPL-validity and meanconc, utilizes SOP under the assumption that the pattern of deduction should be held fixed, not the particular situation. EPL is seen to play a natural role relative to both HPL and CPL. In addition, PSCEA is also demonstrated to be a critical factor in establishing a natural space to formulate the possible deduction schemes for both HPL and EPL.
Tabulations as in Table 6.1 can be used as guidelines for whether certain deduction techniques are really worthwhile, such as the much-used abduction scheme. Of course, the relatively poor showing of abduction – or for that fact, any other CL fallacy which may well have real-world meaning -- may have its performance in Table 6.1 improved upon, by embellishing it with *additional appropriately chosen* premises. But note, one must be careful in adding premises to a particular deduction scheme: improvement – unlike the monotonic logic HPL, as well-characterized numerically by its very definition (eq.(4.10)) in terms of minconc and algebraically via Theorem 4.2– is not guaranteed because of the non-monotonic nature of EPL. Many other CL deduction schemes, including classical fallacies gleaned from standard logic texts (again, see, e.g., the Copi reference), such as the fallacies of "converse accident", "denying the antecedent", "the material conditional", "converting a conditional", etc., can all be tested for their degree of fallaciousness, and it is possible that a number of technically incorrect deduction schemes for CL, may very well, on the average, produce relatively high validity values for the meanconc computations, even without the premise levels t_J approaching unity. Furthermore, the above

comments concerning improving abduction apply here as well. But, of course, all of this must be investigated in future work.

Such expansion of the fundamental results in Table 6.1 may well be expedited by not only application of the situation-specific techniques used to evaluate meanconc for transitivity-syllogism for non-limiting premise probability levels, as illustrated in Section 6.3, but by determining that a general feasible-to-implement technique may exist for evaluating large classes of such problems as particular multiple integrals involving products of powers (due to the initial uniform or dirichlet distribution assumptions) over compact convex regions or polytopes – or via Theorem 3.1 – as related integrals involving negative exponential forms over linearly constrained infinite multi-dimensional regions in the positive orthant. (One basic reference for this possible direction of research is [Bistriczky *et al*., 1994].) In a related vein, the significant results in [Bamber] should be pointed out: the obtaining essentially of a full workable –albeit, somewhat complicated -- algebraic characterization of EPL in its limiting form and a larger class of (limiting form) logics -- based upon the generalization of assumed uniform convergence of the probability levels of the premises to unity here, by "scaled" convergence rates. (In addition, see [Goodman & Nguyen, b] for details of related results.)

The usefulness of EPL can also be extended to linguistic situations, such as the transitivity-syllogism deduction scheme, where typically, one may have the premise "many enemy tanks of class C are of class B" and "almost all tanks of class B are of class A" and the desired conclusion concerns just what rough percentage – or linguistic modifier – should be used to best describe the number of enemy tanks of class C that are of class A ? Here, both REA techniques and RSC can be used together to obtain linguistic results which are fully compatible with the numerical-probabilistic ones of Table 6.1. Because of lack of space here, the interested reader is referred to the recent paper [Goodman & Nguyen, 1999], where Zadeh's linguistic deduction analysis utilizing fuzzy logic (as, e.g., in [Zadeh, 1985]) is so addressed.

7. *Acknowledgments.*

In addition to the aforementioned appreciation expressed by the author to the Office of Naval Research relative to the FY99 Independent Research Program under the Science & Technology Office at SSC-SD (Code D14) for general support of this effort, he would like to add a special acknowledgement to Don Bamber, Code D44215 of SSC-SD, for valuable technical discussions and insights concerning Dr. Bamber's recent work on EPL and related issues. Finally, the author wishes to express his appreciation for the overall strong support of this work by the line management of both his recent former SSC-SD Division (Code D42) -- especially to Dr. R. Jaffee (D42) and G.F. Kramer (D422)-- and those of his present Division (Code D44) -- especially J. Grossman (D44), J.L. Martin (D442) and Dr. J.L. Kaiwi (D44215).

8. *References*

[Aczél, 1966] J. Aczél, *Lecture on Functional Equations and Their Applications,* Academic Press, New York, 1966. Especially, Section 7.1.4, pp. 319-324.
[Adams, 1975] E.W. Adams, *The Logic of Conditionals*, D. Reidel, Dordrecht, Holland, 1975.

[Adams, 1996] E.W. Adams, “Four probability preserving properties of inferences” *Journal of Philosophical Logic*, vol. 25, 1996, pp. 1-24.
[Adams, 1998] E.W. Adams, *A Primer of Probability Logic*, CSLI Public., Stanford, CA, 1998.
[Aitchison, 1986] J. Aitchison, *The Statistical Analysis of Compositional Data*, Chapman & Hall, London, U.K., 1986.
[Alefeld & Herzberger, 1983] G. Alefeld & J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.
[Aubin & Frankowska, 1990] J.-P. Aubin & H. Frankowska, *Set-Valued Analysis*, Birkhauser Press, Basel, Switzerland, 1990.
[Ayton *et al*., 1989, 1991] P. Ayton, A.J. Hunt & G. Wright, “Psychological conceptions of randomness”, *Journal of Behavioral Decision Making* 2: 221-238, 1989. Discussion, same journal, vol. 4, pp. 215-226, 1991.
[Bamber] D. Bamber, “Entailment with near surety of scaled assertions of high conditional probability”, to appear in *Journal of Philosophical Logic*.
[Bergmann *et al*., 1980] M. Bergmann, J. Moor & J. Nelson, *The Logic Book*, Random House Publishers, New York, 1980.
[Bisztriczky *et al.*, 1994] T. Bisztriczky, P. McMullen, R. Schneider, & A. I. Weiss (eds.), *Polytopes: Abstract, Convex and Computational*, Kluwer Academic Publishers, Dordrecht, Holland, 1994.
[Calabrese, 1987] P.G. Calabrese, “An algebraic synthesis of the foundations of logic and probability “, *Information Sciences* 42, 1987, pp. 187-237.
[Calabrese, 1994] P.G. Calabrese, “A theory of conditional information with applications”, in journal [Dubois *et al.*, 1994], pp. 1676-1684.
[Choquet, 1954] G. Choquet, “Theory of capacities”, *Annales de l’Institute Fourier V*, Universite de Grenoble, France, !953-1954, pp. 131-295.
[Copi, 1986] I. M. Copi, *Introduction to Logic, 7th Ed.*, MacMillan, New York, 1986**.**
[Cox, 1946] R. Cox, **“**Probability**,** frequency, and reasonable expectation”, *American Journal of Physics* 14(1), 1946, pp. 1-13.
[Cox, 1979] R.T. Cox, “Of inference and inquiry -- an essay in inductive logic”, in *The Maximum Entropy Formalism* (R.D. Levine & M. Tribus, eds.), 119-168, MIT Press, Cambridge, MA, 1979.
[Dall’Aglio *et al.*, 1991] G. Dall’Aglio, S. Kotz, & G. Salinetti (eds.), *Advances in Probability Distributions with Given Marginals*, Kluwer Academic Publishers, Dordrecht, Holland, 1991.
[DeFinetti, 1974] B. DeFinetti, *Theory of Probability*, Vol. 1, John Wiley & Sons, Inc., New York, 1974.
[Dempster, 1967] A.P. Dempster, “Upper and lower probabilities induced by a multivalued mapping”, *Annals of Mathematical Statistics* 38: 325-338, 1967.
[Dubois & Prade, 1980] D. Dubois And H. Prade, *Fuzzy Sets and Systems*, Academic Press, New York, 1980.
[Dubois & Prade, 1993] D. Dubois & H. Prade, *Conditional Objects as Non-Monotonic Consequent Relationships*, Report IRIT/93-08-R, Institut de Recherche en Informatique de Toulouse, April, 1993.
[Dubois & Prade, 1996] D. Dubois & H. Prade, “Non-standard theories of uncertainty in plausible reasoning”, Chapter 1 (pp. 1-32) in *Principles of Knowledge Representat*ion (G. Brewka, ed.), CSLI Publications, Stanford, CA, 1996.

[Dubois *et al.*, 1994] D. Dubois, I.R. Goodman & P.G. Calabrese (eds.), Special Issue on Conditional Event Algebra, *IEEE Transactions on Systems, Man & Cybernetics*, **24**(12), Dec., 1994.
[Eells & Skyrms, 1994] E. Eells and B. Skyrms (eds.), *Probability and Conditionals*, Cambridge University Press, Cambridge, U.K., 1994.
[Elkan, 1993] C. Elkan, "The paradoxical success of fuzzy logic", *Proceedings 11th Conference on AI (AAAI'93)* (held July 11-15, 1993, Wash., D.C.), AAAI/MIT Press, Menlo Park, CA, 1993, pp. 698-703.
[Freeman, 1993] J.B. Freeman, *Thinking Logically: Basic Concepts for Reasoning, 2nd Ed.*, Prentice Hall, Englewood Cliffs, NJ, 1993.
[Goodman, 1981], I.R. Goodman, "Identification of fuzzy sets with a class of canonically-induced random sets", Proceedings 19th IEEE Conference on decision & Control, Albuquerque, NM (held Dec. 10-12, 1980), pp. 352-357. (Based on unpublished work (1976) referenced in [Nguyen, 1979].)
[Goodman, 1991] I.R. Goodman, "Algebraic and probabilistic bases for fuzzy sets and the development of fuzzy conditioning", in book [Goodman *et al.*, 1991c], pp. 1-69.
[Goodman, 1994] I.R. Goodman, "A new characterization of fuzzy logic operators producing homomorphic-like relations with one-point coverages of random sets", in book *Advances in Fuzzy Theory & Technology, Vol. II* (P.P. Wang, ed.), Duke University, Durham, NC, 1994, pp. 133-159.
[Goodman, 1998a] I.R. Goodman, "New perspectives on deduction in data fusion", *Proceedings of the 1998 (Open) Symposium on Multi-Sensor Data Fusion* (held at Georgia Tech. Research Inst., Marietta, GA, March 30-31, 1998), pp. 79-96.
[Goodman, 1998b] I.R. Goodman, "Random sets and fuzzy sets: a special connection", *Proceedings International Conference on Multisource-Multisensor Information Fusion (Fusion'98)*, Vol. 1, Las Vegas NV, July 6-9, 1998, pp. 93-100.
[Goodman, 1999] I.R. Goodman, "Homomorphic-like random set representations for fuzzy logic models using exponentiation with applications to data fusion"" *Information Sciences*, vol. 113, Jan., 1999, pp. 85-112.
[Goodman & Kramer, 1997] I.R. Goodman & G.F. Kramer, "Extension of relational and conditional event algebra to random sets with applications to data fusion" in book [Goutsias *et al.*, 1997], pp. 209-242.
[Goodman & Nguyen, 1985] I.R. Goodman & H.T. Nguyen, *Uncertainty Models for Knowledge-Based Systems*, North-Holland, Amsterdam, 1985.
[Goodman & Nguyen, 1995] I.R. Goodman & H.T. Nguyen, "Mathematical foundations of conditionals and their probabilistic assignments", *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems,* **3**(3), Sept., 1995, pp. 247-339.
[Goodman & Nguyen, 1998] I.R. Goodman & H.T. Nguyen, "Adams' high probability deduction and combination of information in the context of product space conditional event algebra", *Proceedings of the International Conference on Multisource-Multisensor Information Fusion* (*Fusion '98,* held at Las Vegas, NE, July 6-9, 1998), pp. 1-8.
[Goodman & Nguyen, 1999] I.R. Goodman & H.T. Nguyen, "Application of conditional and relational event algebra to the defining of fuzzy logic concepts", to appear in *Proceedings of Conference on Signal Processing, Sensor Fusion, and Target Recognition* (vol. 3720) (part of S.P.I.E. 13th Annual International Symposium on Aerosense, Orlando, FL, April 5-9, 1999).

[Goodman & Nguyen a] I.R. Goodman & H.T. Nguyen, “Probability updating using second order probabilities and conditional event algebra”, to appear in *Information Sciences*.
[Goodman & Nguyen b] I.R Goodman & H.T. Nguyen, “A new application of bayesian analysis to the resolution of anomalies between probability logic and commonsense reasoning”, lecture on outline of results presented at 37th Annual Bayesian Research Conference, Los Angeles, CA, Feb. 18-19, 1999, with full paper to be submitted for publication
[Goodman *et al.*, 1991a] I.R. Goodman, H.T. Nguyen & G.S. Rogers, “On the scoring approach to admissibility of uncertainty measures in expert systems”, *Journal of Mathematical Analysis & Applications* 159(2): 550-594, Aug., 1991.
[Goodman *et al.*, 1991b] I.R. Goodman, H.T. Nguyen & E.A. Walker, *Conditional Inference and Logic for Intelligent Systems: A Theory of Measure-Free Conditioning*, North-Holland, Amsterdam, 1991.
[Goodman *et al.*, 1991c] I.R. Goodman, M.M. Gupta, H.T. Nguyen & G.S. Rogers, eds., *Conditional Logic in Expert Systems*, North-Holland, Amsterdam, 1991.
[Goodman *et al.*, 1997] I.R. Goodman, R.P. Mahler & H.T. Nguyen, *Mathematics of Data Fusion*, Kluwer Academic, Dordrecht, Holland, 1997.
[Goodman *et al.*, 1999] I.R. Goodman. R.P. Mahler & H.T. Nguyen, “What is conditional event algebra and why should you care?”, to appear in *Proceedings of Conference on Signal Processing, Sensor Fusion, and Target Recognition* (vol. 3720) (part of S.P.I.E. 13th Annual International Symposium on Aerosense, held at Orlando, FL, April 5-9, 1999).
[Goutsias *et al.*, 1997] J. Goutsias, R.P. Mahler & H.T. Nguyen, eds., *Random Sets: Theory & Applications*, Springer-Verlag, New York, 1997.
[Grove & Halpern, 1997] J. Grove & J.Y. Halpern, “Probability update: conditioning vs. cross-entropy”, *Proceedings of the 13th Conference on Uncertainty in AI, 1997* (held at Providence, RI, Aug. 1-3, 1997), pp. 208-214.
[Hailperin, 1965] T. Hailperin, “Best possible inequalities for the probability of a logical function of events”, *American Mathematical Monthly* 72, 1965, pp. 343-359.
[Hailperin, 1984] T. Hailperin, “Probability logic”, *Notre Dame Journal of Formal Logic*, 25(3), July, 1984, pp. 198-212.
[Hailperin, 1996] T. Hailperin, *Sentential Probability Logic*, Lehigh University Press, Bethlehem, PA, 1996.
[Halpern, 1999] J.Y. Halpern, “A counterexample to theorems of Cox and Fine”, *Journal of Artificial Intelligence Research* 10, 1999, pp. 67-85.
[Harding & Kendall, 1974] E.F. Harding & D.G. Kendall (eds.), *Stochastic Geometry*, John Wiley & Sons, Inc., New York, 1974.
[Höhle, 1981] U. Höhle, “Representation theorems for L-fuzzy quantities”, *Fuzzy Sets & Systems* **5**, 1981, pp. 83-107; an outgrowth of his Ph.D. dissertation of 1973.
[Johnson & Kotz, 1972] N. Johnson & S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions (vols. I & II) & Continuous Multivariate Distributions* (vol. IV), John Wiley & Sons, Publishers, New York, 1972.
[Johnson-Laird, 1983] P.N. Johnson-Laird, *Mental Models: Toward a Cognitive Science of Language, Inference, and Conciousness*, Harvard University Press, Cambridge, MA, 1983.
[Kalman, 1994] R.E. Kalman, “Randomness re-examined”, *Modeling, Identification & Control* 15(3): 141-151, 1994.
[Kappos, 1969] D. A. Kappos, *Probability Algebra & Stochastic Processes*, Academic Press, New York, 1969, pp. 16-17 *et passim*.

[Kendall & Moran, 1963] M.G. Kendall & P.A. Moran, *Geometrical Probability*, Charles Griffin & Co., Ltd., London, U.K., 1963.

[Kosko, 1990] B. Kosko, "Fuzziness vs. probability", *International Journal of General Systems,* 17: 211-240, 1990.

[Kosko, 1997] B. Kosko, *Fuzzy Engineering*. Prentice-Hall, New York, 1997.

[Kraus *et al.*, 1990] S. Kraus, D. Lehmann & M. Magidor, "Nonmonotonic reasoning, preferential models and cumulative logics", *Artificial Intelligence*, vol. 44, 1990, pp. 167-207.

[Kyburg, 1983] H. Kyburg, *Epistemology and Inference*, University of Minnesota Press, Minneapolis, 1983; especially Part III, Section 10 (pp. 158-179).

[Lewis, 1976] D. Lewis, "Probabilities of conditionals and conditional probabilities", *Philosophical Review* 85, (1976), 297-315.

[Lindley, 1982] D.V. Lindley, "Scoring rules and the inevitability of probability", *International Statistical Review* 50, pp. 1-14, 1982; discussion follows, pp. 14-26.

[Lukacs, 1955], E. Lukacs, "A characterization of the gamma distribution", *Annals of Mathematical Statistics* 26, 1955, pp. 319-324.

[Matheron, 1975] G. Matheron, *Random Sets & Integral Geometry*, John Wiley & Sons, Inc., New York, 1975.

[Mayer, 1992] R.E. Mayer, *Thinking, Problem Solving, Cognition*, 2nd Ed., W.H. Freeman and Co., New York, 1992.

[McGee, 1989] V. McGee, "Conditional probabilities and compounds of conditionals", *Philosophical Review,* 98, No. 4, (1989), pp. 485-541.

[Mendelson, 1970] E. Mendelson, Boolean Algebra and Switching Circuits, Schaum's Outline Series in Mathematics, McGraw-Hill Book Co., New York, 1970.

[Neblett, 1985] W. Neblett, *Sherlock's Logic*, Dorset Press, New York, 1985.

[Negoita & Ralescu, 1975] C.V. Negoita And D.A. Ralescu, "Representation theorems for fuzzy concepts", *Kybernetes* 4, (1975), 169-174.

[Neveu, 1965] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day Publishers, San Francisco, CA, 1965.

[Nguyen, 1978] H.T. Nguyen, "On random sets and belief functions", *Journal of Mathematical Analyses & Applications* 65, pp. 531-542, 1978.

[Nguyen, 1979] H.T. Nguyen, "Some mathematical tools for linguistic probabilities", *Fuzzy Sets & Systems* 2, 1979, pp. 53-65.

[Nguyen & Walker, 1997] H.T. Nguyen & E.A. Walker, *A First Course in Fuzzy Logic*, CRC Press, Boca Raton, FL, 1997.

[Nilsson, 1986] N.J. Nilsson, "Probabilistic logic", *Artificial Intelligence* 28: 71-87, 1986.

[Oesterle, 1963] J.A. Oesterle, Logic: The Art of Defining and Reasoning, 2nd Ed., Prentice-Hall, Englewood Cliffs, NJ, 1963.

[Orlov, 1978] A.I. Orlov, "Fuzzy and random sets", *Applied Multivariate Statistical Analysis of the Journal of the Academy of Sciences*, USSR, pp. 262-280, Moscow (in Russian), 1978.

[Pearl, 1988] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,* Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[Pearl, 1990] J. Pearl, "System Z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning", in *Theoretical Aspects of Reasoning about Knowledge* (Proceedings of the Third Conference TARK, 1990), R. Parikh, ed., Morgan Kaufmann Publishers, San Mateo, CA, pp. 121-135.

[Pearl & Goldszmidt, 1996] J. Pearl & M. Goldszmidt, "Probabilistic foundations of reasoning with conditionals", Chapter 2 (pp. 33-68) in *Principles of Knowledge Representati*on (G. Brewka, ed.), CSLI Publications, Stanford, CA, 1996.
[Prior *et al.*, 1967] N. Prior *et al.*, "History of logic", in *Encyclopedia of Philosophy, vol. 4* (P. Edwards et al., eds., 1967) , Macmillan, New York, 1967, pp. 513-571.
[Rescher, 1969] N. Rescher, *Many-valued Logic*, McGraw-Hill Book Co., New York, 1969.
[Ripley, 1981] B.D. Ripley, *Spatial Statistics*, John Wiley & Sons, Inc., New York, 1981.
[Schurz, 1983] G. Schurz, "Qualitative probabilities for default reasoning", preprint, 1996.
[Schweizer & Sklar, 1983] B. Schweizer & A. Sklar, *Probabilistic Metric Spaces*, Elsevier Science Publishers, New York, 1983.
[Shafer, 1976] Shafer, G. , *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
[Sklar, 1973] A. Sklar, "Random variables, joint distributions and copulas", *Kybernetika* (Czechoslovakia) 9(6), 1973, pp. 449-460.
[Snow, 1998] P. Snow, "On the correctness and reasonableness of Cox's Theorem for finite domains", *Computational Intelligence* 14(3), Aug., 1998, pp. 452-459.
[Solso, 1979] R. L. Solso, *Cognitive Psychology*, Harcourt Brace Jovanovich, San Diego, 1979.
[Stoyan & Stoyan, 1995] D. Stoyan & H. Stoyan, *Fractals, Random Shapes and Point Fields, Methods of Geometrical Statistics*, John Wiley & Sons, Inc., New York, 1995.
[Van Fraasen, 1976] B. Van Fraasen, "Probabilities of conditionals", in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, (W.L. Harper and C.A. Hooker, eds.), D. Reidel, Dordrecht, Holland, 1976, pp. 261-300.
[Walley, 1991] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, London, U.K., 1991.
[Wilks, 1963] S.S. Wilks, *Mathematical Statistics*, John Wiley & Sons, Publishers, New York, 1963; especially, Chapter 7.
[Zadeh, 1965] L.A. Zadeh, "Fuzzy sets", *Information & Control*, 8(3), June, 1965, pp. 338-353.
[Zadeh, 1975] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning – I, II, III", *Information Sciences*; 8, 1975, pp. 199-249; 8, 1975, pp. 301-357; 9, 1975, pp. 43-80.
[Zadeh, 1978a] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets & Systems* 1, 1978, pp. 3-28.
[Zadeh, 1978b] L.A. Zadeh, "PRUF – A meaning representation language for natural language", *International Journal of Man-Machine Studies* 10, 1978, pp. 395-460.
[Zadeh, 1985] L.A. Zadeh, "Syllogistic reasoning as a basis for combination of evidence in expert systems", *Proceedings International Joint Conference on Artificial Intelligence (IJCAI'85)*, UCLA, 1985, vol. 1, pp. 417-419.