Difficulties of Conditioning and Conditional Independence in Probabilistic Satisfiability

P. Hansen

GERAD and École des Hautes Études Commerciales de Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal, Canada, H3T 2A7
pierreh@crt.umontreal.ca

B. Jaumard

GERAD and École Polytechnique de Montréal
C.P. 6079, station Centre-ville
Montréal (Qc), Canada, H3C 3A7
brigitt@crt.umontreal.ca

A. D. Parreira

École Polytechnique de Montréal
C.P. 6079, station Centre-ville
Montréal (Qc), Canada, H3C 3A7
anderson@crt.umontreal.ca

M. Poggi de Aragão

Catholic University of Rio de Janeiro, Brazil
225, R. Marques São Vicente, Gávea
Rio de Janeiro (RJ), 22453-900
poggi@inf.puc-rio.br

February, 1999
Revised: March, 1999

Les Cahiers du GERAD

G–99–16

Abstract

The probabilistic satisfiability problem consists, given m logical sentences, with their probabilities to find if they are coherent, and if it is the case, to find new bounds on the probability of an additional sentence to be true. This problem can be extended to handle conditional probabilities. Criticism of the adequacy of this approach to modeling conditional uncertain information, due to Cohen, is answered by showing that additional conditions, e.g. facts, should be introduced in the objective and not the constraints.

Ways to express independence are then shown to correspond to two different situations according to presence or absence of *a priori* impossible worlds. Donkin's rule (similar to Dempster-Shafer correction) and Boole's approach which first computes *a priori* probabilities, respectively are the required solution methods.

Keywords: Probability, Logic, Probabilistic Satisfiability, Conditional Probability, Conditioning

Résumé

Le problème de la satisfiabilité probabiliste consiste, étant donné m phrases logiques, ainsi que leurs probabilités d'être vraies, à déterminer si ces probabilités sont cohérentes, ainsi que, dans ce cas à déterminer des bornes les meilleures possibles sur la probabilité d'être vraie d'une phrase additionnelle. Ce problème peut être étendu pour traiter les probabilités conditionnelles. Des critiques de l'adéquation de cette approche de la modélisation avec de l'information conditionnelle et incertaine, dues à Cohen, sont réfutées en montrant que les conditions additionnelles, par exemple des faits, doivent être introduites dans l'objectif et non dans le contraintes.

On montre ensuite qu'il y a deux manières d'exprimer l'indépendance, correspondant à deux situations différentes: l'absence ou la présence *à priori* de mondes impossibles. La règle de Donkin (semblable à correction de Dempster-Shafer) et l'approche de Boole, calculant d'abord des probabilités *à priori*, sont respectivement les méthodes de résolution requises.

Mots clés: Probabilité, Logique, Satisfiabilité Probabiliste, Probabilité Conditionnelle, Conditionnement.

1 Introduction

The *probabilistic satisfiability* problem (PSAT) can be expressed as follows. Let $x_1, x_2, \ldots, x_n$ be logical variables and $S_1, S_2, \ldots, S_m$ logical sentences built upon these variables with the usual operators $\vee$ (union, or Boolean sum), $\wedge$ (intersection, or Boolean product) and $^-$(negation). Assume probabilities $\pi_1, \pi_2, \ldots, \pi_m$ of being true are associated with $S_1, S_2, \ldots, S_m$. The *decision form* of PSAT is then to decide whether these probabilities are coherent or not. The *optimization form* of PSAT is, given an additional sentence S_{m+1}, to find best possible lower and upper bounds $\underline{\pi}_{m+1}$ and $\overline{\pi}_{m+1}$ on its probability π_{m+1} of being true. This problem dates from Boole (1954) and has been extensively studied by, among others, Hailperin (1986; 1996).

Consider all 2^n vectors of truth values for $x_1, x_2, \ldots, x_n$, or *constituents*, and let $A = (a_{ij})$ be the $m \times 2^n$ matrix where $a_{ij} = 1$ if S_i is true for the j^{th} such vector and $a_{ij} = 0$ otherwise. Associating probabilities p_j to all constituents, PSAT may be written (in optimization form):

$$\begin{array}{rcl} \min/\max\ \pi_{m+1} & = & \displaystyle\sum_{j=1}^{2^n} a_{m+1} p_j \\ \text{subject to:} & & \\ \mathbb{1}\, p & = & 1 \\ Ap & = & \pi \\ p & \geq & 0 \end{array} \tag{1}$$

i.e., as a linear program with an exponential number of variables. Observe that not all 2^m possible vectors of A need (or usually will) be present. Those which are will be called *possible worlds* and the others *impossible worlds*.

As noted by Boole (1954), introducing logical variables $s_i = S_i$ and constraints $r_i = \bar{s}_i S_i \vee s_i \bar{S}_i = 0$, PSAT may be expressed as "minimize or maximize π_{m+1} subject to $p(s) = \pi$ and $R = (r_i) = 0$".

In this form PSAT consists in a set of variables with given probabilities (which may be viewed as assumptions) and a set of rules which restrict the values these variables can take. Clearly, this is nothing else than a Probabilistic Assumption-Based Truth Maintenance System (probabilistic ATMS) without independence assumptions. As the converse is obviously true one has (in modern terms):

Theorem 1.1 *[Boole, 1854] Without independence assumptions, PSAT and probabilistic ATMS are equivalent.*

Now, both Boole (1854) and several recent authors studying probabilistic ATMS [de Kleer and Williams, 1987; Laskey and Lehner, 1989; Kohlas *et al.*, 1998] make independence assumptions, but not always in the same way. In the present paper we show that there are basically two situations, according to whether probabilities have been observed without or with some impossible worlds. The relevant ways to handle PSAT are then Donkin's rule (1851), similar to the Dempster-Shafer [Dempster, 1967; Shafer, 1976] correction and Boole's method of *absolute* (or *a priori*) *probabilities*. Before discussing this in detail we pave the way by refuting some criticism of Cohen (1989) about Ramsey's (1931) treatment of conditional probabilities and related work of de Finetti (1937; 1974), which are prefigurations of PSAT.

2 Cohen's Critique

As mentioned above PSAT can be extended to handle conditional probabilities, i.e., conditions of the form

$$p(S_1|S_2) = p(S_1S_2)/p(S_2) = \pi_{1|2}.$$

This was already done for particular examples (and sometimes approximately) by Boole (1854), briefly evoked by Ramsey (1931) and de Finetti (1937; 1974), and systematically explored by Hailperin (1986; 1996) and Jaumard *et al.* (1991), as well as, in a more complicated way, by members of the Italian school of subjective probabilities.

Ramsey (1931) and de Finetti (1937) prove the following result. Assume (subjective) probabilities π_i express the odds at which a decision-maker would accept to bet for or against the occurrence of the event corresponding to S_i (e.g., $\pi_i = 1/4$ this decision-maker would accept a bet of 3 to 1 that the event will occur). A *Dutch book* is a combined bet in which the bettor will lose *regardless of the outcome* (so if he accepts both a bet on S_1 at 1 to 2 and on $\bar{S}_1$ at 1 to 2, he would, in total, lose 1 regardless of whether S_1 is true of false, i.e., a Dutch book would have been made against him).

Theorem 2.1 *[Ramsey, 1931; de Finetti, 1937] Subjective probabilities do not lead to a Dutch book if and only if they satisfy the axioms of finite probability theory (or, in other words, if the set of constraints of problem (1) is not contradictory).*

It is well known that for this result to hold payments should be expressed in utility unit and utility functions be linear.

In his book *An Introduction to the Philosophy of Induction and Probability*, Cohen (1989), notes that

> "Ramsey held that a person's judgment of probability measured his degree of rational belief, where degree of belief is manifested by the lowest betting-odds on A that he would accept within a coherent systems of wagers".

He then adds

> "But Ramsey was mistaken in thinking that $p(A|B)$ can then be interpreted as expressing the lowest odds at which the person would now bet on A, with the bet only being valid if B is true".

and

> "Ramsey here seems to have been confounding statements of the form $p(A|B) = n$ with statements of the form *If and only if B is true, then* $p(A) = n$".

But consider again the definition of a conditional probability, under the form

$$p(S_1S_2) - p(S_2)\pi_{1|2} = 0$$

or

$$a'_{1\wedge 2}p = (a_{1\wedge 2} - \pi_{1|2}a_2)p = 0$$

with $a_{1|2,j} = 1$ (resp. $a_{2j} = 1$ if S_1 and S_2 (resp. S_2) are true and 0 otherwise. Then add $\pi_{1|2}p = \pi_{1|2}$ thus obtaining

$$a''_{1\wedge 2}p = \pi_{1|2}$$

where $a''_{1\wedge 2,j} = 1$ if S_1 and S_2 are true, $a''_{1\wedge 2,j} = 0$ is S_1 is false and S_2 true and $a''_{1\wedge 2,j} = \pi_{1|2}$for S_2 false in world j. So, one indeed gets from the classical definition of conditional probability payments 1 in case of gain and 0 in case of loss when the condition holds and $\pi_{1|2}$ in case the condition does not hold, i.e., the bet is cancelled. If it thus clear that Ramsey (1931) does not make the confusion mentioned above.

Cohen (1989) has another, more precise, objection. He states

> "An obvious trouble unnoticed by Ramsey, is that the interpretation in terms of a conditional bet would allow a Dutch book to he made against certain otherwise unexceptionable assignments of equal values to $p(A|B)$ and $p(notA|C)$ if B and C both turn out to be true. Consider for example a case in which A is "It will rain tomorrow", B is "The barometer is falling", and C is "The wind is in the east" where the bets would be 1 to 2 on A if B is true and 1 to 2 on $notA$ is C is true".

Note that here the so-called *Dutch book* is subject to the condition that B and C are both true, and not a loss in all cases; but the difficulty is a real one: why should one accept to have a contradiction in that case? The possible worlds are, before stating that B and C are true

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
A	1	1	1	1	0	0	0	0
B	1	1	0	0	1	1	0	0
C	1	0	1	0	1	0	1	0

and the PSAT model:

$$\begin{array}{ll} \text{min/max} & p_1 + p_2 + p_3 + p_4 \\ \text{subject to:} & \\ & p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 = 1 \\ & p_1 + p_2 + \frac{1}{3}p_3 + \frac{1}{3}p_4 + \frac{1}{3}p_7 + \frac{1}{3}p_8 = \frac{1}{3} \\ & \frac{1}{3}p_2 + \frac{1}{3}p_4 + p_5 + \frac{1}{3}p_6 + p_7 + \frac{1}{3}p_8 = \frac{1}{3} \\ & p_j \geq 0 \end{array} \qquad (2)$$

and $\underline{\pi}_{m+1} = 0$, $\overline{\pi}_{m+1} = 1$ (with $p_8 = 1$ and $p_4 = 1$ respectively). Now, if one would add constraints stating that $p(B) = 1$ and $p(C) = 1$, i.e.,

$$\begin{array}{ccccccccccccccc} p_1 & + & p_2 & & & + & p_5 & + & p_6 & & & = & 1 \\ p_1 & & & + & p_3 & + & p_5 & & & + & p_7 & = & 1 \end{array} \tag{3}$$

this would indeed imply $p_2 = p_3 = p_6 = p_7 = 0, p_1 = p_5 = 1/3$, and $p_1 + p_5 = 1 > 1/3 + 1/3$, a contradiction. But such an addition is incorrect modeling. Indeed, the first constraints describe the observed *general situation* (i.e., for rainy and not rainy days, high and low barometers values, wind from the east or another direction). Fixing B and C corresponds to a *particular situation*, worthy of study, but different. It is not true that one always has B and C both true in the general situation and that is what constraints (3) express. To study the case when they are, one must *condition the objective function*, i.e., seek bounds on the probability of A given B and C within the general model consisting of the given constraints. The variation of PSAT so obtained leads to the hyperbolic program [Hailperin, 1986; Jaumard *et al.*, 1991]

$$\min/\max \ \frac{p_1 + p_2 + p_3 + p_4}{p_1 + p_5} \tag{4}$$

subject to the constraints of (2). As noted in [Hailperin, 1986], the technique proposed by [Charnes and Cooper, 1962] may then be used, to reduce this problem to a linear program: an additional variable t is used as a parameter to obtain a value 1 in the denominator of the objective function. Both numerator and denominator of the objective function are multiplied by t as well as all constraints. Then setting $y_i = tp_i$, one obtains

$$\begin{array}{rl} \min/\max & y_1 + y_2 + y_3 + y_4 \\ \text{subject to:} & \\ & y_1 + y_5 = 1 \\ & y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 = t \\ & y_1 + y_2 + \frac{1}{3}y_3 + \frac{1}{3}y_4 + \frac{1}{3}y_7 + \frac{1}{3}y_8 - \frac{1}{3}t = 0 \\ & \frac{1}{3}y_2 + \frac{1}{3}y_4 + y_5 + \frac{1}{3}y_6 + y_7 + \frac{1}{3}y_8 - \frac{1}{3}t = 0 \\ & y_j \geq 0 \end{array} \tag{5}$$

Again one finds $\underline{\pi}_{m+1} = 0$, $\overline{\pi}_{m+1} = 1$ (as can easily be checked with a Venn diagram). Cohen (1969) further discusses a way out of this difficulty, referring to de Finetti (1974):

> “Someone might try to defend the Ramsey-de Finetti interpretation of dyadic probability-functions by pointing out that it will work quite satisfactorily in practice if the proposition upon which the validity of the bet is made conditional is required not only to be true, but also to state the *only* additional fact that must becomes known, if the condition's fulfillment is to be warranted to be assertable. This manœuvre will certainly preclude generating the kind of incoherence or contradiction to which Ramsey's analysis was shown to be exposed”.

But again, this is not the case. Indeed consider the above mentioned situation, i.e., $p(A|B) = 1/3, p(\bar{A}|C) = 1/3$ and assume it is also known that $p(\bar{B}\bar{C}) = 1/2$, i.e., there is a 1/2 probability not to have both a barometer not falling and no wind from the east. Then adding the condition $p(B) = 1$ (or the condition $p(C) = 1$) in the constraints yields a contradiction, while there is no substantive reason to have one. Once more, conditioning the objective function does not do so.

3 Donkin's Rule

The bounds $\underline{\pi}_{m+1}$and $\overline{\pi}_{m+1}$given by PSAT may be loose. Therefore further assumptions have been made in a variety of ways from Boole (1954) to the present, in order to get better estimates of π_{m+1}. The easiest way is to make some independence assumption on the probabilities of the sentences. But this is tricky, as these sentences are subject to constraints.

We first discuss the case where there are none *a priori*, and consider an example close to one of Boole (1854).

Consider an urn with marble and wooden balls which are colored white or black. Assume their numbers n_{ij}, where i refers to material and j to color, are given in the following table

i \ j	White	Black
Marble	10	3
Wood	5	7

The probability that a ball drawn at random is in marble (event A) is $p(A) = 13/25$, that it is in white (event B) is $p(B) = 15/25$, etc. The four constituents AB, $A\bar{B}$, $\bar{A}B$, and $\bar{A}\bar{B}$ have probabilities $p_1 = 10/25 = 0.4$, $p_2 = 3/25 = 0.12$, $p_3 = 5/25 = 0.2$, and $p_4 = 7/25 = 0.28$.

Consider an iterated random drawing of balls, with replacement in the case the ball is in wood and black. What is then the probability that the first ball not replaced in the urn be white?

An answer can be found, using Donkin's rule [Donkin, 1851] which was known to Boole:

> "If there be n hypotheses $h_1, h_2, \ldots, h_n$ which in a certain state of information are believed to be *exhaustive*, and of which the probabilities are then $p_1, p_2, \ldots, p_n$; and if it is afterwards discovered either that some of them must be rejected or that others must be admitted, or both, without any further information as to those of the original set which are retained, then these latter have the same mutual ratios *after* the new information that they had before".

So the *a priori* (or in Boole's terms *absolute*) probabilities p_1, p_2 and p_3 of the first three constituents must be divided by $1 - p_4$ in order to obtain new values p_1', p_2' and p_3'

for the probabilities of these possible worlds which sum to 1, once the replacement rule is enforced.

Donkin (1851) viewed his principle as axiomatic. Boole (1854) observes that, in the finite case, it follows from the classical (Pascalian) definition of probabilities. In our example,

$$p'_1 = \frac{\dfrac{n_{11}}{n_{11}+n_{12}+n_{21}+n_{22}}}{1-\dfrac{n_{22}}{n_{11}+n_{12}+n_{21}+n_{22}}} = \frac{n_{11}}{n_{11}+n_{12}+n_{21}},$$

etc, or numerically $p'_1 = 0.4/(1-0.28) = 0.555, p'_2 = 0.12/(1-0.28) = 0.166$, and $p'_3 = 0.2/(1-0.28) = 0.277$. Moreover, $p'(A) = p'_1 + p'_3 = 0.555 + 0.277 = 0.832$.

In the general case consider again model (1) and assume that all worlds are possible *a priori*, but then that additional information shows that some of them are in fact impossible. Let V denote the event "the world is possible". Then *a priori* for any event A

$$p(A) = p(A|V)p(V) + p(A|\bar{V})p(\bar{V})$$

and $p(V)$ may be less than one. Updating means that one must find new probabilities such that $p'(A) = p(A|V)p'(V) + p(A|\overline{V})p'(\overline{V})$, $p'(V) = 1$ and $p'(\bar{V}) = 0$. They are given by

$$p'(A) = \frac{p(A)}{p(V)} = \frac{p(A)}{1-p(\bar{V})} = p(A|V) = \frac{p(AV)}{p(V)}$$

i.e., they are the probabilities of events conditioned by V. The computations are the counterpart for given values of the *a priori* probabilities of the constituents of the computations done by hyperbolic programming with model (4) for unknown ones satisfying the constraints.

Let us now assume independence of the events corresponding to sentences of (1) and that all worlds are *a priori* feasible, i.e., there are no rules. In other words, all sentences are variables s_i with probabilities π_i. Then the probability of world j is

$$p_j = \prod_{i=1}^{m} \pi_i^{a_{ij}} (1-\pi_i)^{1-a_{ij}}.$$

After worlds have been excluded one gets for any event A

$$p'(A) = \frac{\sum_{j\in A\cap V} p_j}{1-\sum_{j\in\overline{V}} p_j} = \frac{\sum_{j\in A\cap V}\prod_{i=1}^{m} \pi_i^{a_{ij}}(1-\pi_i)^{1-a_{ij}}}{1-\sum_{j\in\overline{V}}\prod_{i=1}^{m} \pi_i^{a_{ij}}(1-\pi_i)^{1-a_{ij}}}$$

(where the index j designates the corresponding constituent).

So, while computation may be tedious, they are easy in this case. Moreover, simpler boolean expressions for V and V' than sums of constituents can be found by various algorithms and the values of the π_i substituted for the corresponding s_i in such formulae to compute $p(V)$ or $(1-p(\overline{V}))$. This has been explored in depth by Kohlas and Money (1995) and Anrig *et al* (1998).

4 Boole's Method

Correctness of the application of Donkin's or Dempster-Shafer's rule to PSAT with an independence assumption is dependent on the further assumption that there are *a priori* no impossible worlds. When this in not true one should not apply this rule anymore (at least in isolation). To see why, consider the following example. Let s_1 denote the sentence "the sky is blue", s_2 the sentence "there are heavy clouds" and s_3 the additional sentence "the sun shines". Moreover, consider the rules "if the sky is blue, then the sun shines", (i.e., $s_1 \wedge \bar{s}_3 = 0$), "if there are heavy clouds then the sun does not shine", (i.e., $s_2 \wedge s_3 = 0$) and "there cannot be both a blue sky and heavy clouds", (i.e.,$s_1 \wedge s_2$). If the sky is not blue nor are there heavy clouds the sun may shine or not according to the fact that the light clouds present hide it or not.

In California, one might assume that $\pi_1 = p(s_1) = 0.8$ and $\pi_2 = p(s_2) = 0.1$. There are four worlds: $s_1s_2, s_1\bar{s}_2, \bar{s}_1s_2$ and $\bar{s}_1\bar{s}_2$. The rules imply that the first one is impossible and the three others possible. Moreover, s_3 is true in world $s_1\bar{s}_2$, and possibly in world $\bar{s}_1\bar{s}_2$, false in world $\bar{s}_1s_2$. To find new possible bounds, without independence assumptions, one solves the problem

$$\underline{\pi}_3 = \min \quad p_2, \qquad\qquad \overline{\pi}_3 = \max \quad p_2 + p_4$$
$$\text{subject to:}$$
$$\begin{array}{ccccccccc} p_1 & + & p_2 & + & p_3 & + & p_4 & = & 1 \\ p_1 & + & p_2 & & & & & = & 0.8 \\ p_1 & & & + & p_3 & & & = & 0.1 \\ p_1 & & & & & & & = & 0 \end{array}$$
$$p_2, p_3, p_4 \geq 0$$

(alternately, one might use eight worlds instead of two different objective functions). The unique solution is $p_1 = 0, p_2 = 0.8, p_3 = 0.1$ and $p_4 = 0.1$ with $\underline{\pi}_3 = 0.8$ and $\overline{\pi}_3 = 0.9$.

Assume now *a priori* independence of s_1 and s_2, neglecting the rules in a first step. This gives probabilities $p_1 = 0.08, p_2 = 0.72, p_3 = 0.02$ and $p_4 = 0.18$. As the first world is impossible, Donkin's rule gives $p'_2 = 0.72/(1-0.08) = 0.7826, p'_3 = 0.02/(1-0.08) = 0.0217, p'_4 = 0.18/(1-0.08) = 0.1957, \underline{\pi}'_3 = 0.7826$ and $\overline{\pi}'_3 = 0.9783$. But such results do not make sense: why should the probability of having sun be bounded below by a value smaller than the probability of blue sky and why should this same probability be bounded above by a number much larger than 1 minus the probability of heavy clouds?

What has gone wrong is that the values observed here are not *a priori* probabilities (the π above), but probabilities conditioned on the non-occurrence of impossible worlds (the π'_i above). So, according to Boole (1854), one must first compute the *a priori* probabilities π_i form the conditional ones π'_i. Then one can apply Donkin's rule with these π_i. Garbolino (1998) argues in a rational reconstruction of this reasoning, based on Hailperin (1986), that Boole probably discovered his method by thus reversing the order of computations of Donkin's rule.

It is the *a priori* probabilities which should in this case be assumed to be independent. Boole (1854) thought that such an assumption could always be made, a position sharply criticized by Wilbraham (1854). Nevertheless, his concept of independence, which differ both from logical and from stochastic independence appears to be an interesting, if neglected, generalization of the latter.

For our example one must solve the equations

$$\pi_1' = \frac{\pi_1 \times (1-\pi_2)}{1-(\pi_1 \times \pi_2)} \qquad \text{and} \qquad \pi_2' = \frac{(1-\pi_1)}{1-(\pi_1 \times \pi_2)}$$

which have for solution (neglecting the solution $\pi_1 = \pi_2 = 1$ which gives undefined values to π_1' and π_2'):

$$\pi_1 = \frac{\pi_1'}{1-\pi_2'} = \frac{0.8}{1-0.1} = \frac{8}{9} \qquad \text{and} \qquad \pi_2 = \frac{\pi_2'}{1-\pi_1'} = \frac{0.1}{1-0.8} = \frac{1}{2}$$

Note that the *a priori* probabilities may differ considerably form the observed conditional ones. Then $\underline{\pi}_3' = 0.8$ and $\overline{\pi}_3' = 0.9$, which agrees with the unique solution obtained without independence assumptions. Of course for more complex examples this coincidence might not be observed anymore, i.e., one might have $\underline{\pi}_{m+1} < \underline{\pi}'_{m+1} \leq \overline{\pi}'_{m+1} < \overline{\pi}_{m+1}$.

In the general case one must solve the system of equations:

$$\pi_i = \frac{\sum_{j \in V} a_{ij} \prod_{k=1}^{m} \pi_k'^{a_{kj}} (1-\pi_k')^{(1-a_{kj})}}{1 - \sum_{j \in \overline{V}} \prod_{k=1}^{m} \pi_k'^{a_{kj}} (1-\pi_k')^{(1-a_{kj})}} \tag{6}$$

for the π_i given the π_k'. Such equations are, after multiplying both sides by the denominators, multilinear in the π_k''s. In view of Donkin's rule, Boole (1854) first thought that it was obvious that if the PSAT problem (1) had a solution, the corresponding constraints (6) had a single solution with $0 < \pi_i < 1$ for $i = 1, 2, \ldots, m$. He later realized that a proof was needed, worked hard on if and provided one in Boole (1862). Hailperin (1986) discusses Boole's proof, from the standpoint of modern rigor and finds one step of it not to be conclusive, but without providing a counter-example.

5 Conclusion

We have shown that when a general situation is described by a PSAT model, new observations such as facts, corresponding to particular situations, should be introduced into the model by conditioning the objective function. If they are added in the constraints, substantively unjustified contradictions may occur.

Moreover, we have made clear that introducing independence assumptions into PSAT should be done in two distinct ways. If there are no *a priori* impossible worlds, Donkin's rule (or Dempster-Shafer's correction) can be applied to the observed or given *a priori* probabilities. If there are some *a priori* impossible worlds, one must, following Boole, first deduce from the observed conditional probabilities the corresponding *a priori* probabilities and then perform this correction.

These remarks have implications for the correctness of many proposed ways to treat uncertainty in knowledge-based system, which will be studied in future work.

Acknowledgments

Research of the first three authors was supported by Defense Research Establishment (DREV) Valcartier contract. Work of the first and second authors was supported by FCAR grants 92-EQ-1048 and 95-ER-1048. Work of the first has also been supported by NSERC grant to HEC and NSERC grant GP0105574. Work of the second author was also supported by NSERC grants GP0036426 and EQP0157431. Work of the third author has also been supported by CNPq (Conselho Nacional de Desevolvimento Científico e Tecnológico), Brazil, grant 200722/95-6.

References

[Anrig *et al.*, 1997] B. Anrig, R. Haenni, J. Kohlas, and N. Lehmann. Assumption-based Modeling Using ABEL. In D. Gabbay, R. Kruse, A. Nonnengart, and H.J. Ohlbach, editors, *First Int. Joint Conf. on Qualitative and Quantitative Practical Reasoning.* ECSQARU–FAPR'97, Springer 1997.

[Boole, 1854] G. Boole. *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities.* London: Walton and Maberley, 1854.

[Boole, 1862] G. Boole. On the Theory of Probabilities. *Philosophical Transactions of the Royal Society*, 152: 225–252, 1862.

[Charnes and Cooper, 1962] A. Charnes and W.W. Cooper. Programming with Linear Fractional Functionals. *Naval Research Logistics Quarterly*, 9: 181–186, 1962.

[Cohen, 1989] L. J. Cohen. *An Introduction to The Philosophy of Inductions and Probability.* Clarendon Press, Oxford, 1989.

[de Finetti, 1937] B. de Finetti. La Prévision: Ses Lois Logiques, Ses Sources Subjectives. *Annales de l'Institut Poincaré*, 7: 1–68, 1937. Tr. eng. in H.E. Kyburg and H.E. Smokler (eds.), *Foresight: Its Logical laws, Its Subjective Sources* in *Studies in Subjective Probability* (New York: John Wiley and Sons, 1964), pp. 99-158.

[de Finetti, 1974] B. de Finetti. *Theory of Probability - A Critical Introductory Treatment*, volume 1. Wiley, New York, 1974.

[de Kleer and Williams, 1987] J. de Kleer and B. C. Williams. Diagnosing Multiple Faults. *Artificial Intelligence*, 32(1): 97–130, 1987.

[Dempster, 1967] A.P. Dempster. Upper and Lower Probabilities Generated by a Random Closed Interval. *Annals of Math. Stat.*, 35: 325–339, 1967.

[Donkin, 1851] W. Donkin. On Certain Questions Relating to the Theory of Probabilities. *Philosophical Magazine*, 1: 55–60, 1851.

[Garbolino, 1998] P. Garbolino. George Boole e lo specttro di Laplace. In E. Agazzi and N. Vassalo, editors, *George Boole Filosofia, Logica, Matematica*, pages 97–110. Franco Angeli, Milan, 1998.

[Hailperin, 1986] T. Hailperin. Boole's Logic and Probability. In *Studies in Logic and the Foundations of Mathematics 85*. North Holland, Amsterdam, 2nd enlarged edition, 1986.

[Jaumard *et al.*, 1991] B. Jaumard, P. Hansen, and M. V. Poggi de Aragão. Column Generation Methods for Probabilistic Logic. *ORSA Journal on Computing*, 3: 135–148, 1991.

[Kohlas and Monney, 1995] J. Kohlas and P. A. Monney. Probabilistic Assumption-based Reasoning. In *A Mathematical Theory of Hints*, volume 425 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin Heidelberg, 1995.

[Kohlas *et al.*, 1998] J. Kohlas, B. Anrig, R. Haenni, and P.A. Monney. Model-Based Diagnostics and Probabilistic Assumption-Based Reasoning. *Artificial Intelligence*, 104: 71–106, 1998.

[Laskey and Lehner, 1989] K.B. Laskey and P.E. Lehner. Assumptions, Beliefs and Probabilities. *Artificial Intelligence*, 41: 65–77, 1989.

[Ramsey, 1931] F.P. Ramsey. *The Foundations of Mathematics and other Logical Essays.* Routledge and Kegan Paul, London, 1931.

[Shafer, 1976] G. Shafer. *A Mathematical Theory of Evidence.* Princeton Univ. Press, 1976.

[Wilbraham, 1854] H. Wilbraham. On the Theory of Chances developped in Professor Boole's "Laws of Thought". *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, 7: 465–476, 1854.