Logique & Analyse **160** (1997), 375–384

PARACONSISTENT PROBABILITY THEORY AND PARACONSISTENT BAYESIANISM*

Edwin D. MARES

Abstract
This paper presents a theory of probability based on the paraconsistent logic D4. The resulting probability functions are then used to define two sorts of Bayesian updating. One sort of updating merely uses the simple rule of conditionalisation. The other sort adds a wrinkle to the simple rule so that agents' beliefs become more consistent as well as more complete through updating.

1. *Introduction*

Paraconsistent logic has become an entrenched part of logic and an active field of research. But, despite there being some interesting and important literature on the use of paraconsistent logic to do history and philosophy of science and mathematics, to develop a doxastic logic and to treat belief revision[1] there has been relatively little work done on integrating paraconsistent logic into mainstream epistemological theories. The present paper does just that. It attempts to meld paraconsistent logic and Bayesianism.

Before we launch into the formal development of a paraconsistent theory of probability, we need a few definitions. A paraconsistent logic is a logic that does not make valid the rule schema

$$\frac{\begin{array}{c} A \\ \neg A \end{array}}{\therefore B}.$$

This is the rule of ex falso quodlebet (EFQ). Different paraconsistent logicians have had different reasons for rejecting EFQ. Strong paraconsisten-

* Thanks go to Graham Priest for useful comments on an earlier draft.

[1] See, for example, Batens [1985], Jennings and Schotch [1989], Mortensen [1995] and Fuhrmann [1991].

tists hold that there are true contradictions. Weak paraconsistentists, on the other hand, believe that sometimes people have inconsistent beliefs, that we tell inconsistent stories, and so on, but there are no true contradictions. Weak paraconsistentists think that we need a paraconsistent logic to understand commitments made by people with inconsistent beliefs, made within inconsistent stories, etc.

Among weak paraconsistentists, some are "adjunctivists" and some are "non-adjunctivists". An adjunctivist holds that if someone states or believes that A and states or believes that B, then he or she is committed to the conjunction $A \wedge B$.[2] Typically, non-adjunctivists hold that a person cannot believe contradictions like $A \wedge \neg A$, but instead can only believe both A and $\neg A$. Thus, in most cases at least, non-adjunctivists think that one can have contradictory beliefs but not believe in contradictions.

Integrating non-adjunctivist paraconsistency into a Bayesian framework would seem to be reasonably simple. We would take an agent to have a set of probability functions. Where A and its negation are both believed, the agent has some probability functions that give A a high probability and others that give it a low probability. We then model his or her beliefs using these two sets of probability functions. Where the agent has many inconsistent beliefs this can become quite complicated, but the basic principle is not difficult.

It might seem odd to use an adjunctive paraconsistent logic as a basis for a probability theory. For, the probabilistic conception of belief and commitment is itself non-adjunctive. Just because A and B each have high probability does not mean that $A \wedge B$ also has high probability. Thus, belief (and assertibility) on the probabilistic conception is non-adjunctive. The virtue in using an adjunctive logic lies in the ability of giving contradictions such as $A \wedge \neg A$ high probability. Thus, we are able to model someone's believing in a contradiction. Using a purely non-adjunctive approach, we always give contradictions a null value.

In the present paper, I use an adjunctive paraconsistent logic as a basis for a theory of probability. I then put this theory of probability to work. I present two rules of conditionalisation and briefly develop two variations of Bayesianism. One of these versions of Bayesianism is an epistemological counterpart of a very strong paraconsistency and the other is an epistemology suited to the weak paraconsistentist or, with minor modifications, to a more mild form of strong paraconsistency.

[2]There are grades of adjunctivism. Jennings and Schotch hold a position inbetween adjunctivism and non-adjunctivism. See, e.g., Jennings and Schotch [1989].

2. *Paraconsistent Probability*

In this section we set out a theory of probability based on a paraconsistent logic. This is not the first time this has been done. In [1987], Priest presents a very similar theory.[3] What is novel about my approach is the way it the theory of probability is applied in section three below.

Our language is a standard propositional language with the operators $\wedge$, $\vee$, $\neg$, parentheses and propositional variables p, q, r, ... The standard formation rules apply. We use upper case roman letters for metavariables.

Our logic is an elegant many-valued logic based on Dunn's four-valued semantics (see Dunn [1969] and [1976]). This semantics was originally used the model the first-degree fragment of the relevant logic E, but it has come into more general use in the literature on paraconsistent logic.[4] We call our logic 'D4'.

On Dunn's semantics, a valuation takes a wff to a subset of the set $\{t,f\}$. Thus, a wff can take the values $\{t\}$, $\{f\}$, $\{t,f\}$ and $\emptyset$. If t is in the value of a wff A, then A is said to be true. If f is in its value, then A is said to be false. Dunn gives the following truth and falsity conditions for the connectives:

- $A \wedge B$ is true if and only if A is true and B is true. $A \wedge B$ is false if and only if A is false or B is false.
- $A \vee B$ is true if and only if A is true or B is true. $A \vee B$ is false if and only if A is false and B is false.
- $\neg A$ is true if and only if A is false. $\neg A$ is false if and only if A is true.

Clearly, the D4 truth conditions generalise the classical truth conditions for the connectives.

Following standard practice in probability theory, we will construct a probability space from a set of points and an algebra of propositions. This algebra, however, will be not be a Boolean algebra as usual, but rather an algebra corresponding to D4. A *D4-set algebra* Γ is a structure $\langle Prop, \cup, \cap, F\rangle$ such that $Prop$ is a set of sets of points closed under union and intersection. $\cup$ and $\cap$ are standard set theoretic union and intersection and F is a DeMorgan complement on $Prop$. $Prop$ contains $\cup Prop$ and $\emptyset$. More-

[3] In addition, Leslie Roberts has been developing probabilistic semantics for paraconsistent logics. See, for example, her "Maybe, Maybe Not: Probabilistic Semantics for Two Paraconsistent Logics" presented to the Australasian Association for Logic conference, Sydney, 1998.

[4] The difference between our use of Dunn's four valued semantics and his original use of it is that we use a different semantic entailment relation.

over, the DeMorgan complement satisfies the following postulates. Where φ and ψ are propositions,

- $\varphi^{FF} = \varphi$ (period two);
- $(\cup Prop)^F = \emptyset$;
- $(\varphi^F \cap \psi)^F = (\varphi^F \cup \psi^F)$ (DeMorgan's Law).

A *paraconsistent probability space* (PPS) is a quadruple $\langle Sit, Prop, F, \Pr\rangle$ such that *Sit* is a non-empty set of points (called 'situations'), $\langle Prop, \cup, \cap\rangle$ is a D4-set algebra of subsets of Sit (such that $\cup Prop = Sit$), and Pr is a function from Prop into [0,1] satisfying the definitions and postulates below:

CP $\Pr(\psi/\varphi) = \frac{\Pr(\psi\cap\varphi)}{\Pr(\varphi)}$, where $\Pr(\varphi) \neq 0$;

Pr1 $\Pr(Sit) = 1$; $\Pr(\emptyset) = 0$;
Pr2 $\Pr(\varphi \cup \psi) = (\Pr(\varphi) + \Pr(\psi)) - \Pr(\varphi \cap \psi)$;
Pr3 If $\varphi \subseteq \psi$, $\Pr(\psi) - \Pr(\varphi) \geq 0$.

Pr3 is not redundant; since *Prop* is not closed under standard set-theoretic complementation, we cannot use the usual proof to derive Pr3 from Pr2.

This probability function is very similar to a finitely additive Kolmogorov function, but there are some interesting differences. Because negation here is paraconsistent and the set of propositions is not necessarily a sigma algebra, we cannot derive that $\Pr(\varphi^F) = 1 - \Pr(\varphi)$. Also, for similar reasons, for any given classical tautology, we cannot derive that the probability assigned to it is one.

A *paraconsistent probability model* (PPM) is a quintuple $\langle Sit, Prop, F, \Pr, V\rangle$ such that $\langle Sit, Prop, F, \Pr\rangle$ is a PPS and V is a function from propositional variables to propositions. V in turn determines a homomorphism, $|\ |$, from formulae to propositions such that $|A \wedge B| = |A| \cap |B|$, $|A \vee B| = |A| \cup |B|$ and $|\neg A| = |A|^F$.

We also define a semantic entailment relation $\models$ between sequences of wffs and wffs such that $A_1, ..., A_n \models B$ if and only if $|A_1| \cap ... \cap |A_n| \subseteq |B|$. From this definition and Pr3, we immediately obtain the following fact:

Fact 1 If $A_1, ..., A_n \models B$, *then* $\Pr(|A_1| \cap ... \cap |A_n|) \leq \Pr(B)$.

If we interpret our probability functions as agents' subjective probability functions, then we take agents' beliefs to be closed under D4 entailment. That is, if an (ideal) paraconsistent agent gives high probability to A and A semantically entails B, then she gives high probability to B as well.

One might wonder why I did not use the (slightly) simpler logic LP, which has the same truth conditions as D4, but only three truth values, $\{t\}$, $\{f\}$ and $\{t,f\}$. The reason why I did not is two-fold. First, an LP algebra is a particular case of a D4 algebra. An LP algebra is a D4 algebra that satisfies the additional postulate that $\varphi \cup \varphi^F = \cup Prop$. Thus the present theory is a bit more general than one that uses LP as a base. Second, agents on the present view are not committed to believing in all classical tautologies. If we were to use LP, then we would have to accept this additional commitment.

In what follows, where no confusion will result, we will write 'Pr(A)' instead of 'Pr($|A|$)'.

3. *Bayesian Updating*

Now we come to the application of our theory of probability. As I say in the introduction above, my aim here is to integrate paraconsistency into Bayesianism. In particular, I want to have a theory of belief updating (or, rather, a theory of updating one's degrees of belief) based on paraconsistent probability.

First, we should interpret our probability functions as representing distributions of agents' degrees of belief, as in standard Bayesianism. Then, we need a rule of conditionalisation to govern changes of degrees of belief in response to evidence. The orthodox Bayesian rule for updating degrees of belief is simple conditionalisation, that is,

(RC) $\Pr_A(B)$ - $\Pr(B/A)$,

where $\Pr(A) \neq 0$ and where A is a piece of evidence and Pr is the agent's probability function prior to learning that A.

We can apply (RC) to paraconsistent probability without really changing anything. We need only interpret Pr as ranging over paraconsistent probability functions.

This all seems quite straightforward, but I think there may be a better updating rule. I developed this second rule with weak paraconsistency in mind, so let me explain it in that context. A weak paraconsistentist might hold that many people believe in contradictions, but that they should try to eliminate these beliefs. It would seem reasonable if a weak paraconsistentist were to hold that a person should try to make his or her beliefs more consistent while updating. Thus, in updating one's beliefs, on this view, an agent should develop not only a more complete but also a more consistent picture of the world. With this in mind, I developed the second rule of conditionalisation, (RC*):

$$\text{(RC*)} \quad \Pr^{*}_{A}(B) = \frac{\Pr(B \wedge A) - \Pr(B \wedge A \wedge \neg A)}{\Pr(A) - (\Pr(A \wedge \neg A)},$$

where $\Pr(A) > \Pr(A \wedge \neg A)$. In updating, as in orthodox Bayesianism., the agent sets to the probability of the new evidence to one and adjusts her other degrees of belief accordingly. At the same time, she sets the conjunction of the evidence and its negation to zero and adjusts the other degrees of belief accordingly.

This view of updating makes sense, as we have said, from the point of view of weak paraconsistency. If we assume as, in effect, do orthodox Bayesians that our means of collecting data is perfectly reliable, then it would seem right to treat our updating process as a means of resolving our contradictory beliefs. If we weaken the assumption that we have perfectly reliable means for obtaining evidence, then we should modify our updating function or our background theory of probability. We could, for example, try to develop a paraconsistent version of Jeffrey conditionalisation to treat evidence of which we are not certain. Or we might attempt to create a paraconsistent version of Popperian conditional probability which allows us to "revive" propositions to which we had previous given zero probability.

(RC*) can also be integrated into a strong paraconsistent epistemology. The strong paraconsistent holds, as we have said, that there are true contradictions. A strong paraconsistentist might demarcate the sorts of propositions φ of which it is reasonable to believe both φ and its negation. This paraconsistentist could then restrict (RC*) for use with propositions not in this class. Or (RC*) can be used to update a proposition the negation of which has been rejected. Philosophers such as Priest (Priest [1987]; see also Parsons [1984]) distinguish between asserting the negation of a proposition and rejecting that propositions. In rejecting a proposition, perhaps among other acts, one refuses to assert it. It would seem reasonable to use (RC*) to update propositions that are accepted and the negation of which are rejected; the agent sets the newly accepted evidence to one and its rejected negation (and the intersection of the two) to zero. Thus, although (RC*) was created with weak paraconsistency in mind, it can with restrictions be integrated into a strong paraconsistent epistemology as well.

In the remainder of the paper we will use the following more general version of (RC*):

$$\text{(RC*)} \quad \Pr^{*}_{\alpha}(\beta) =_{df} \frac{\Pr(\alpha \cap \beta) - \Pr(\alpha \cap \beta \cap \alpha^{F})}{\Pr(\alpha) - \Pr(\alpha \cap \alpha^{F})},$$

where $\Pr(\alpha) > \Pr(\alpha \cap \alpha^{F})$. We shall also use the following definition:

Definition 2 A proposition α *is* Pr-normal *if* $\Pr(\alpha) - \Pr(\alpha \cap \alpha^F) > 0$.

Fact 3 If Pr *is a probability function and* α *is* Pr-*normal*, $\Pr^*_\alpha(\alpha^F) = 0$.

Proof. Suppose that α is Pr-normal.

By definition, $\Pr^*_\alpha(\alpha^F) = \frac{\Pr(\alpha\cap\alpha^F)-\Pr(\alpha\cap\alpha^F\cap\alpha^F)}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)}$.

Since α is Pr-normal, $\Pr(\alpha) - \Pr(\alpha \cap \alpha^F) > 0$.
Moreover, $\Pr(\alpha \cap \alpha^F \cap \alpha^F) = \Pr(\alpha \cap \alpha^F)$.

So, $\Pr^*_\alpha(\alpha^F) = \frac{\Pr(\alpha\cap\alpha^F)-\Pr(\alpha\cap\alpha^F\cap\alpha^F)}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)} = \frac{\Pr(\alpha\cap\alpha^F)-\Pr(\alpha\cap\alpha^F)}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)} = 0$. ■

We now show that updating with Pr-normal evidence yields a probability function.

Lemma 4 If Pr *is a probability function and* α *is* Pr-*normal, then* $\Pr^*_\alpha(Sit) = 1$ *and* $\Pr^*_\alpha(\emptyset) = 0$.

Proof. Easy. ■

Lemma 5 If Pr *is a probability function and* α *is* Pr-*normal, then* $\Pr^*_\alpha(\beta \cup \gamma) = (\Pr^*_\alpha(\beta) + \Pr^*_\alpha(\gamma)) - \Pr^*_\alpha(\beta \cap \gamma)$.

Proof. Suppose that Pr is a probability function and α is Pr-normal.

1. $\Pr^*_\alpha(\beta \cup \gamma) = \frac{\Pr((\beta\cup\gamma)\cap\alpha)-\Pr((\beta\cup\gamma)\cap\alpha\cap\alpha^F)}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)}$, def $\Pr^*_\alpha$
2. $\Pr((\beta \cup \gamma) \cap \alpha) = \Pr((\beta \cap \alpha) \cup (\gamma \cap \alpha))$
3. $\Pr((\beta \cap \alpha) \cup (\gamma \cap \alpha)) = (\Pr(\beta \cap \alpha) + \Pr(\gamma \cap \alpha))$
   $- \Pr(\alpha \cap \beta \cap \gamma)$, Pr2
4. $\Pr((\beta \cup \gamma) \cap \alpha \cap \alpha^F) = (\Pr(\beta \cap \alpha \cap \alpha^F) + \Pr(\gamma \cap \alpha \cap \alpha^F))$
   $- \Pr(\beta \cap \gamma \cap \alpha \cap \alpha^F)$, Pr2
5. $\Pr^*_\alpha(\beta \cup \gamma) =$

$$\frac{((\Pr(\beta\cap\alpha)+\Pr(\gamma\cap\alpha))-\Pr(\alpha\cap\beta\cap\gamma))-((\Pr(\beta\cap\alpha\cap\alpha^F)+\Pr(\gamma\cap\alpha\cap\alpha^F))-\Pr(\beta\cap\gamma\cap\alpha\cap\alpha^F))}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)},$$

1,3,4

6. $\mathrm{Pr}^*_\alpha(\beta \cup \gamma) =$

$$\frac{((\mathrm{Pr}(\beta\cap\alpha)+\mathrm{Pr}(\gamma\cap\alpha))-(\mathrm{Pr}(\beta\cap\alpha\cap\alpha^F)+\mathrm{Pr}(\gamma\cap\alpha\cap\alpha^F))-(\mathrm{Pr}(\alpha\cap\beta\cap\gamma)-\mathrm{Pr}(\beta\cap\gamma\cap\alpha\cap\alpha^F)))}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)},$$

5, rearranging

7. $\mathrm{Pr}^*_\alpha(\beta \cup \gamma) =$

$$\frac{((\mathrm{Pr}(\beta\cap\alpha)-\mathrm{Pr}(\beta\cap\alpha\cap\alpha^F))+(\mathrm{Pr}(\gamma\cap\alpha)-\mathrm{Pr}(\gamma\cap\alpha\cap\alpha^F)))}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)} - \frac{\mathrm{Pr}(\alpha\cap\beta\cap\gamma)-\mathrm{Pr}(\beta\cap\gamma\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)},$$

6, rearranging

8. $\mathrm{Pr}^*_\alpha(\beta \cup \gamma) =$

$$\left(\frac{\mathrm{Pr}(\beta\cap\alpha)-\mathrm{Pr}(\beta\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)}+\frac{\mathrm{Pr}(\gamma\cap\alpha)-\mathrm{Pr}(\gamma\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)}\right)-\frac{\mathrm{Pr}(\alpha\cap\beta\cap\gamma)-\mathrm{Pr}(\beta\cap\gamma\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha)-\mathrm{Pr}(\alpha\cap\alpha^F)},$$

7, rearranging

9. $\mathrm{Pr}^*_\alpha(\beta \cup \gamma) = (\mathrm{Pr}^*_\alpha(\beta) + \mathrm{Pr}^*_\alpha(\gamma))$ - $\mathrm{Pr}^*_\alpha(\beta \cap \gamma)$, 8, def. Pr^*_α.

■

Lemma 6 If Pr *is a probability function, α is* Pr*-normal, α is* Pr*-normal and $\beta \subseteq \gamma$, then* $\mathrm{Pr}^*_\alpha(\gamma)$ - $\mathrm{Pr}^*_\alpha(\beta) \geq 0$.

Proof. Suppose that Pr is a probability function, α is Pr-normal, α is Pr-normal and $\beta \subseteq \gamma$.

1. $\beta \subseteq \gamma$, hypothesis
2. $\mathrm{Pr}^*_\alpha(\gamma) - \mathrm{Pr}^*_\alpha(\beta) = \frac{\mathrm{Pr}(\gamma\cap\alpha)-\mathrm{Pr}(\gamma\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha\cap\alpha^F)} - \frac{\mathrm{Pr}(\beta\cap\alpha)-\mathrm{Pr}(\beta\cap\alpha\cap\alpha^F)}{\mathrm{Pr}(\alpha\cap\alpha^F)}$
3. $(\mathrm{Pr}(\gamma \cap \alpha) - \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F)) - (\mathrm{Pr}(\beta \cap \alpha) - \mathrm{Pr}(\beta \cap \alpha \cap \alpha^F))$
   $= (\mathrm{Pr}(\gamma \cap \alpha) + \mathrm{Pr}(\beta \cap \alpha \cap \alpha^F)) - (\mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F) + \mathrm{Pr}(\beta \cap \alpha))$
4. $\mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F) + \mathrm{Pr}(\beta \cap \alpha)$
   $= \mathrm{Pr}((\gamma \cap \alpha \cap \alpha^F) \cup (\beta \cap \alpha)) + \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F \cap \beta \cap \alpha)$
5. $\mathrm{Pr}((\gamma \cap \alpha \cap \alpha^F) \cup (\beta \cap \alpha)) + \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F \cap \beta \cap \alpha)$
   $= \mathrm{Pr}(((\beta \cap \alpha) \cup \gamma) \cap ((\beta \cap \alpha) \cup \alpha) \cap ((\beta \cap \alpha) \cup \alpha^F))$
   $+ \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F \cap \beta \cap \alpha)$
6. $\mathrm{Pr}(((\beta \cap \alpha) \cup \gamma) \cap ((\beta \cap \alpha) \cup \alpha) \cap ((\beta \cap \alpha) \cup \alpha^F))$
   $+ \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F \cap \beta \cap \alpha)$
   $= \mathrm{Pr}(\gamma \cap \alpha \cap ((\beta \cap \alpha) \cup \alpha^F)) + \mathrm{Pr}(\gamma \cap \alpha \cap \alpha^F \cap \beta \cap \alpha)$
7. $\mathrm{Pr}(\gamma \cap \alpha) \geq \mathrm{Pr}(\gamma \cap \alpha \cap ((\beta \cap \alpha) \cup \alpha^F))$, Pr3
8. $(\mathrm{Pr}(\gamma \cap \alpha) + \mathrm{Pr}(\alpha \cap \alpha^F \cap \beta)) - (\mathrm{Pr}(\gamma \cap \alpha \cap ((\beta \cap \alpha) \cup \alpha^F))$
   $+ \mathrm{Pr}(\alpha \cap \alpha^F \cap \beta)) \geq 0$, 7

9. $\Pr(\alpha) - \Pr(\alpha \cap \alpha^F) > 0$, α is Pr-normal

10. $\frac{(\Pr(\gamma\cap\alpha)+\Pr(\alpha\cap\alpha^F\cap\beta))-(\Pr(\gamma\cap\alpha\cap((\beta\cap\alpha)\cup\alpha^F))+\Pr(\alpha\cap\alpha^F\cap\beta))}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)} \geq 0$, 8,9

11. $\Pr^*_\alpha(\gamma) - \Pr^*_\alpha(\beta) =$

$\frac{(\Pr(\gamma\cap\alpha)+\Pr(\alpha\cap\alpha^F\cap\beta))-(\Pr(\gamma\cap\alpha\cap((\beta\cap\alpha)\cup\alpha^F))+\Pr(\alpha\cap\alpha^F\cap\beta))}{\Pr(\alpha)-\Pr(\alpha\cap\alpha^F)}$, 2–6

12. $\Pr^*_\alpha(\gamma) - \Pr^*_\alpha(\beta) \geq 0$, 10,11

■

Lemma 7 If Pr *is a probability function and* α *is* Pr-*normal,* $0 \leq \Pr^*_\alpha(\beta) \leq 1$.

Proof. Suppose that Pr is a probability function and α is Pr-normal.

1. $\Pr^*_\alpha(\beta) = \Pr^*_\alpha(\beta \cap \alpha)$, def. $\Pr^*_\alpha$
2. $\Pr^*_\alpha(\beta \cap \alpha) \leq \Pr^*_\alpha(\alpha)$, lemma 6 above
3. $\Pr^*_\alpha(\alpha) = 1$, α is normal
4. $\Pr^*_\alpha(\beta) \leq 1$, 1,2,3
5. $\Pr^*_\alpha(\beta \cap \alpha) - \Pr^*_\alpha(\beta \cap \alpha \cap \alpha^F) \geq 0$, lemma 6
6. $\Pr^*_\alpha(\alpha) - \Pr^*_\alpha(\alpha \cap \alpha^F) \geq 0$, lemma 6
7. $\Pr^*_\alpha(\beta) = \frac{\Pr^*_\alpha(\beta\cap\alpha)-\Pr^*_\alpha(\beta\cap\alpha\cap\alpha^F)}{\Pr^*_\alpha(\alpha)-\Pr^*_\alpha(\alpha\cap\alpha^F)} \geq 1$, 5,6

■

Theorem 8 If Pr *is a probability function and* α *is* Pr-*normal, then* $\Pr^*_\alpha$ *is a probability function.*

Proof. Follows directly from lemmas 4,5,6 and 7. ■

4. *Concluding Remarks*

In sum, we have set out a theory of probability that takes D4 as its base logic. We put this theory of probability to work by using it to define two versions of Bayesian updating. Then we showed that the latter, more interesting version, has one property that it will need if it is to be a useful epistemological tool. We showed that when updating with a Pr-normal proposition, we obtain a probability function.

Clearly we have just scratched the surface of the topic of the relationship between probability, paraconsistency and epistemology. For example, there are more interesting variations on Bayesianism. that we should explore. A paraconsistent version of Richard Jeffrey's theory (see Jeffrey [1983]) is an obvious choice for the next step in this project. Moreover, we should examine paraconsistent versions of related theories, like Bas van Fraassen's logic of full belief (van Fraassen [1995]) and Wolfgang Spohn's ordinal conditional functions (Spohn [1988]). To modify van Fraassen's view to accommodate paraconsistency would require the construction of a paraconsistent theory of conditional probability. All in all, there is a lot of work to do, and there are many more theorems to be proved.

Victoria University of Wellington

REFERENCES

[1985] Batens, D. "Meaning, Acceptance and Dialectics" in J. C. Pitt (ed.), *Change and Progress in Modern Science*, Dordrecht: Reidel, 1985, pp. 333–360.

[1969] Dunn, J. M. "Natural Language versus Formal Language" talk presented at the joint ASL/APA symposium, New York, 1969.

[1976] Dunn, J. M. "Intuitive Semantics for First-Degree Entailments and 'Coupled Trees'" *Philosophical Studies* 29 (1976) pp. 149–168.

[1995] van Fraassen, B. C. "Fine-Grained Opinion and the Logic of Full Belief" *Journal of Philosophical Logic* 24 (1995) pp. 349–377.

[1991] Fuhrmann, A. "Theory Contraction Through Base Contraction" *Journal of Philosophical Logic* 20 (1991) pp. 175–203.

[1983] Jeffrey, R. C., *The Logic of Decision*, New York: McGraw-Hill, 1983, second edition.

[1989] Jennings, R. and P. Schotch "On Detonating" in G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic*, Munich: Philosophia Verlag, 1989.

[1995] Mortensen, C., *Inconsistent Mathematics*, Dordrecht: Reidel, 1995.

[1984] Parsons, T. "Assertion, Denial and the Liar Paradox" *Journal of Philosophical Logic* 13 (1984) pp. 137–152.

[1987] Priest, G., *In Contradiction*, The Hague: Nejhoff, 1987.

[1988] Spohn, W. "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States" in W. L. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change and Statistics II*, Dordrecht: Reidel, 1988.