
On the Emergence of Reasons in Inductive Logic

J.B. PARIS and M. WAFY¹, *Department of Mathematics, University of Manchester, Manchester M13 9PL, UK. E-mail: jeff@ma.man.ac.uk*

Abstract

We apply methods of abduction derived from propositional probabilistic reasoning to predicate probabilistic reasoning, in particular inductive logic, by treating *finite* predicate knowledge bases as potentially *infinite* propositional knowledge bases. It is shown that for a range of predicate knowledge bases (such as those typically associated with inductive reasoning) and several key propositional inference processes (in particular the Maximum Entropy Inference Process) this procedure is well defined, and furthermore yields an explanation for the validity of the induction in terms of ‘reasons’.

Keywords: Inductive Logic, Probabilistic Reasoning, Abduction, Maximum Entropy, Uncertain Reasoning.

1 Motivation

Consider the following situation. I am sitting by a bend in a road and I start to wonder how likely it is that the next car which passes will skid on this bend. I have some knowledge which seems relevant, for example I know that if there is ice on the road then there is a good chance of a skid, and similarly if the bend is unsigned, the camber adverse, etc.. I possibly also have some knowledge of how likely it is that there is ice on the road, how likely it is that the bend is unsigned (possibly conditioned on the iciness of the road) etc.. Notice that this is generic knowledge which applies equally to any potential passing car.

Armed with this knowledge base I may now form some opinion as to the likely outcome when the next car passes. Subsequently several cars pass by. I note the results and in consequence possibly revise my opinion as to the likelihood of the next car through skidding.

Clearly we are all capable of forming opinions, or beliefs, in this way, but is it possible to formalize this inductive process, this process of *uncertain reasoning* about a general population (of potential passing cars in this case) from basic generic knowledge (of ice etc.) and possibly some knowledge of a finite number of previous instances (of passing cars)?

In this paper we shall sketch such a formalization for a limited class of knowledge bases. It is based on extending ideas on abductive reasoning about *finite* propositional probabilistic knowledge bases to such predicate, and potentially *infinite*, knowledge bases. In order to do this we shall assume that we are working in a predicate language

¹Supported by an Egyptian Government Scholarship, File No.7083.

L with 0-ary predicates (i.e. propositional variables) Q_1, Q_2, \dots, Q_q (e.g. standing for ‘ice on the road’ etc.), a single unary predicate $P(x)$ (e.g. standing for car x skids) and a denumerable list of constants a_1, a_2, a_3, \dots (e.g. standing for the sequence of passing cars). Let SL denote the set of closed, quantifier free sentences of this language using, say, the connectives \neg, \wedge, \vee . Essentially then SL may be thought of as the set of sentences of the *propositional* language with the infinite set of propositional variables $Q_1, Q_2, \dots, Q_q, P(a_1), P(a_2), P(a_3), \dots$. For future use let $SL^{(r)} \subset SL$ be the set of sentences of the finite propositional language with propositional variables $Q_1, Q_2, \dots, Q_q, P(a_1), P(a_2), P(a_3), \dots, P(a_r)$.

We shall further assume that our (generic) knowledge base is of the form

$$\bigcup_{i=1}^{\infty} K(a_i),$$

where $K(a_1)$ consists of a (satisfiable) finite set of linear constraints (over the reals)

$$c_1 Bel(\theta_1) + c_2 Bel(\theta_2) + \dots + c_m Bel(\theta_m) = d,$$

on a subjective probability function $Bel : SL \rightarrow [0, 1]$, with $\theta_1, \theta_2, \dots, \theta_m$ sentences from $SL^{(1)}$ and $K(a_i)$ is the result of replacing a_1 everywhere in $K(a_1)$ by a_i . So, for example, with the above interpretation of $Q_1, P(a_1)$, etc. my knowledge that given the road is icy car a_1 will skid with (subjective) probability $1/5$ might be reformulated as the linear constraint

$$Bel(P(a_1) \wedge Q_1) - 1/5 \cdot Bel(Q_1) = 0,$$

on my assigning subjective probability function Bel . Thus in this note we are identifying knowledge with a satisfiable set of linear constraints on a probability function Bel where, as usual (see [6]), a function $Bel : SL \rightarrow [0, 1]$ is a *probability function* if it satisfies that for all $\theta, \phi \in SL$,

$$(P1) \quad \text{If } \models \theta \text{ then } Bel(\theta) = 1,$$

$$(P2) \quad \text{If } \models \neg(\theta \wedge \phi) \text{ then } Bel(\theta \vee \phi) = Bel(\theta) + Bel(\phi).$$

[It is easy to check that if $K(a_1)$ is satisfiable then so is $\bigcup_{i=1}^{\infty} K(a_i)$, and, of course, conversely.]

Notice that the generic nature of the knowledge is captured by the fact that the knowledge base is invariant under renaming, i.e. permutation, of the a_i . Notice also that any constraint in this knowledge base only mentions at most one a_i . In other words we are assuming that any relation between the $P(a_i)$ is entirely accounted for by, or mediated through, their individual relationships to the Q_1, Q_2, \dots, Q_q . In terms of our skidding example this amounts to the assumption that the action of one car does not directly influence the actions of any other car.

The question of induction (\mathcal{Q}) that we are interested in here then is:-

Given my knowledge base $\bigcup_{i=1}^{\infty} K(a_i)$ what belief (i.e. subjective probability), $Bel(P(a_i))$, should I assign to $P(a_i)$? More generally what belief should I assign to $Bel(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i}))$ where the $\epsilon_i \in \{0, 1\}$, and $P^{\epsilon_i}(a_{n_i})$ is $P(a_{n_i})$ if $i = 1$ and $\neg P(a_{n_i})$ if $i = 0$?

[In such expressions we take it as read that the n_i are distinct.] It is important to point out here that in asking this question we assume that the knowledge base *sums up all my knowledge* (the so called *Watt's Assumption* of [6]).

Ideally (in our view) the answer to this question should follow from considerations of rationality and common sensicality. We shall consider that point shortly. For the moment however we notice that there is one situation in which the 'right' answer *seems* abundantly clear. Namely, suppose $K(a_1)$ consists of the (consistent) set of constraints:-

$$\begin{aligned} \text{(i)} \quad & Bel(P(a_1) \wedge Q_j) - \beta_j Bel(Q_j) = 0, \quad j = 1, 2, \dots, q, \\ \text{(ii)} \quad & Bel(Q_j \wedge Q_k) = 0, \quad 1 \leq j < k \leq q, \\ \text{(iii)} \quad & Bel(Q_j) = \lambda_j, \quad j = 1, 2, \dots, q, \quad \text{where } \sum_{j=1}^q \lambda_j = 1. \end{aligned} \quad (1.1)$$

In this case the Q_j form a *complete set of reasons*, in that they are (i) 'reasons' (for $P(a_1)$ if $\beta_j > 1/2$, *against* if $\beta_j < 1/2$), (ii) disjoint, and (iii) exhaustive. Given a knowledge base of this special form there is an evident solution based on the implicit assumption that the $P(a_i)$ are, modulo the knowledge base, independent of each other, namely:-

$$Bel(P(a_i)) = \sum_{j=1}^q \lambda_j \beta_j,$$

and more generally

$$Bel\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right) = \sum_{j=1}^q \lambda_j \beta_j^m (1 - \beta_j)^{r-m},$$

where $m = \sum \epsilon_i$. We shall call this the *canonical solution* based on this complete set of reasons. It is interesting to note at this point that if two reasons, Q_i and Q_j say, have the same strength, that is $\beta_i = \beta_j$, then as far as the canonical solution is concerned they may be combined into a single reason with this common strength and weight $\lambda = \lambda_i + \lambda_j$. From this point of view then reasons are characterized purely by their strengths.

2 The Maximum Entropy solution

We now turn to considering solutions Bel to $\bigcup_{i=1}^{\infty} K(a_i)$ based on principles of common sense. Common sense principles were introduced explicitly in [8] (several of these had appeared earlier, especially in [12], a paper drawing similar conclusions albeit from rather stronger initial assumptions) as *constraints on the process of assigning beliefs* from (finite, linear) probabilistic knowledge bases. In this paper (subsequently improved in [6], [10] and [7]) it was shown that the *Maximum Entropy Inference Process*, ME , is the only inference process which satisfies all these common sense principles.

To expand on this result and its context, in [8] we define an *inference process* N to be a function which for any finite, linear, satisfiable, set K of constraints on

a probability function Bel on the sentences $S\mathcal{L}$ of a finite propositional language \mathcal{L} , selects a particular probability function $Bel = N(K)$ satisfying K^2 . Thus N corresponds to a process for (consistently) assigning probabilities on the basis of such knowledge bases K . The common sense principles referred to above arise by considerations of the consistency (in its informal, everyday, sense) of this process (see in particular [7]).

As far as the inference process ME is concerned, whilst it could be defined as the unique solution to these principles it has an alternate, older, and much more practical characterization. Namely $ME(K)$ is that solution Bel to K for which the entropy

$$-\sum_{i=1}^{2^n} Bel(\alpha_i) \log(Bel(\alpha_i))$$

is maximal, where the α_i run over the *atoms* of $S\mathcal{L}$, that is the sentences

$$p_1^{\epsilon_1} \wedge p_2^{\epsilon_2} \wedge \dots \wedge p_n^{\epsilon_n}, \quad \epsilon_1, \epsilon_2, \dots, \epsilon_n \in \{0, 1\},$$

where p_1, p_2, \dots, p_n enumerate the propositional variables of \mathcal{L} .

Given this privileged status of ME it would seem natural to argue that the answer to our question \mathcal{Q} should be that provided by ME . Indeed, it could be claimed that to do otherwise would be to contradict common sense. Unfortunately however, we cannot mechanically apply ME here because our knowledge base $\bigcup_{i=1}^{\infty} K(a_i)$ and the overlying language are infinite. Nevertheless, the nature of the original problem points clearly to the direction we should take. To illustrate this in the example of the passing cars, we remark that the idea that there are actually infinitely many of them queuing up to negotiate this bend is (despite the daily impression left by the rush hour!) clearly an idealization³. In truth there is only a ‘potential infinity’, and this being the case we would argue that the correct application of common sense in \mathcal{Q} would be to assign

$$Bel\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right)$$

the value

$$\text{Lim}_{n \rightarrow \infty} ME\left(\bigcup_{j=1}^n K(a_j)\right)\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right), \quad (2.1)$$

assuming that this limit exists. Notice that if all such limits do exist (indeed this applies to any inference process, not just ME) then the property that they satisfy $(P1)$, $(P2)$ is also preserved in the limit. In other words these limiting values determine a probability function on sentences of the language with propositional variables $P(a_1), P(a_2), P(a_3), \dots$.

²Strictly we should also include the language \mathcal{L} as an argument of N . However for this paper we shall only consider *language invariant* inference processes, that is inference processes which are independent of the overlying language insofar as assigning probabilities to a particular sentence is concerned. For a further explanation of this point see

[6]

³Though such a word can hardly be considered appropriate in this case!

It is rather straightforward to show that in the case where $K(a_1)$ is a complete set of reasons as in (1.1) the limit in (2.1) exists and equals the canonical solution. [For a proof of this result and the other main results in this paper see [13] or [11].] Indeed, in this case the situation is particularly simple because each sequence

$$ME\left(\bigcup_{j=1}^n K(a_j)\right)\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right),$$

is eventually constant.

More interesting is the case of a general (finite, satisfiable, of course) $K(a_1)$ as the next theorem shows:-

Theorem 2.1 The limits

$$\text{Lim}_{n \rightarrow \infty} ME\left(\bigcup_{j=1}^n K(a_j)\right)\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right)$$

exist and agree with the canonical solution for some complete set of reasons.

This is perhaps initially a rather surprising result. It says that no matter what our generic knowledge $K(a_1)$ is, if we assign values to the $(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i}))$ according to $ME(\bigcup_{j=1}^n K(a_j))$ (i.e. according to common sense on the basis of knowledge $\bigcup_{j=1}^n K(a_j)$) then in the limit these assignments look as if they have been based, canonically, on some complete set of reasons. In other words, in the limit ‘reasons’ have *emerged* to explain our answers!

Of course the proof itself provides some de-mystification, these ‘reasons’ in fact correspond in the limit to the atoms of the language with propositional variables Q_1, Q_2, \dots, Q_q , so in particular there are just 2^q of them. Thinking of these atoms as specifying the background world, or state of the world, in which the experiments are to be conducted leads then in turn to identifying these ‘reasons’ with the ‘possible worlds’.

To give a specific, albeit completely contrived, example suppose you are sitting in the waiting room of a driving test centre waiting for your nephew to return from taking his test. You know nothing at all about your nephew’s prowess at the wheel. However you do have some fragments of knowledge about the arrangements for the test itself. Namely; on any one day all tests are carried out by the same examiner (one of L or S) around the same circuit (one of A or B); on average 30% of drivers pass; a driver who passes is twice as likely to have been tested by L than by S ; if circuit A is chosen then the driver has only a 20% chance of passing; S prefers circuit B 70% of the time. The secretary now tells you with a malicious grin that all 3 previous tests that morning have resulted in failure. Based on this limited information what belief, as subjective probability, should you give to your nephew breaking the pattern?

Denoting the event of a successful test for driver a_i by $P(a_i)$ etc. and your personal probability function by Bel your knowledge might reasonably be captured by the following $K(a_1)$:

$$\begin{aligned} Bel(P(a_1)) &= 3/10 \\ Bel(P(a_1) \wedge \neg S) &= 2Bel(P(a_1) \wedge S) \end{aligned}$$

$$\begin{aligned} Bel(P(a_1)|A) &= 1/5 \\ Bel(\neg A|S) &= 7/10 \end{aligned}$$

In this case the probability function,

$$Bel\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right) = \text{Lim}_{n \rightarrow \infty} ME\left(\bigcup_{j=1}^n K(a_j)\right)\left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right)$$

given by Theorem 2.1 is the canonical solution of the complete set of reasons given by (1.1) when $q = 4$ (corresponding to the 4 atoms $A \wedge S, A \wedge \neg S, \neg A \wedge S, \neg A \wedge \neg S$) and the β_i, λ_i are given by

i	1	2	3	4
β_i	0.554	0.666	0.282	0.275
λ_i	0.0824	0.000868	0.192	0.7243

In particular then, on learning of the 3 previous failures that morning *Bel* dictates that you should only give probability

$$Bel(P(a_4)|\neg P(a_1) \wedge \neg P(a_2) \wedge \neg P(a_3)) = \frac{\sum_{i=1}^4 \lambda_i \beta_i (1 - \beta_i)^3}{\sum_{i=1}^4 \lambda_i (1 - \beta_i)^3} = 0.167$$

to your nephew passing.

It is important to emphasize in this example that, as with the rest of this paper, we are dealing here with beliefs as *subjective probabilities*. Were we to consider the above example in terms of objective probabilities, say as given by the long term frequencies of the various combinations of circuits, examiners and outcomes, then surely most of us would be reasonably happy to assign a probability to $\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})$ of

$$\sum RF(\pm A \wedge \pm S) \cdot RF(P(a_1)|\pm A \wedge \pm S)^m (1 - RF(P(a_1)|\pm A \wedge \pm S))^{r-m},$$

where as usual $m = \sum \epsilon_i$ and $RF(A \wedge S)$ is the relative frequency of $A \wedge S$ etc.. In other words, to accept these $\pm A \wedge \pm S$ as 'reasons' with the λ_i, β_i given by the corresponding relative frequencies. What we show in this paper is that, if we accept the arguments for using *ME* in this way then complete sets of reasons emerge naturally also in the case where probabilities are subjective degrees of belief (and, as far as this paper is concerned, $K(a_1)$ has this rather restricted form).

To those familiar with the popular image of *ME* the conclusion that the atoms are the 'reasons' may, after some brief consideration, appear a not unexpected *artifice* of *ME*. After all, a common view of *ME* is that it tries to avoid introducing unnecessary dependencies. In this case (i.e. $K(a_1)$) it seems that the only way knowledge of the outcome of one experiment can provide information about the outcome of another experiment is through the mediation of its effect on the possible worlds. Conditioning on a fixed world then should leave *ME* free to treat the experiments as entirely independent.

Attractive as this explanation may appear in this simple case further investigations would seem to suggest that this is not quite the whole story. Firstly, as we shall shortly

see, this behavior is not simply an artifice of ME . Secondly, as will be shown in a forthcoming paper, this behavior continues to be manifest in more general cases (than the simple $K(a_1)$ we looked at here), in particular where finitely many predicates $P(x)$ are allowed and where the above mentioned conditional independence is lacking.

3 The Minimum Distance and CM^∞ solutions

In this section we consider the corresponding situation with two other choices of inference process, the minimum distance inference process, MD , and the limiting centre of mass inference process CM^∞ .

To begin with MD , this can be defined analogously to ME but with the alternate measure of ‘information content’,

$$\sum_{i=1}^{2^n} (Bel(\alpha_i) - 1/2^n)^2 \quad (3.1)$$

replacing the Shannon information content

$$\sum_{i=1}^{2^n} Bel(\alpha_i) \log(Bel(\alpha_i)).$$

In other words $MD(K)$ is that probability function Bel satisfying K for which the expression (3.1) is *minimal*. [See [6] for further motivation and properties of this inference process. In particular it is shown there that MD is, like ME , language invariant.]

In this case rather more work is required (see [13] or [11]) to prove a result which for ME was rather straightforward, namely:-

Theorem 3.1 If $K(a_1)$ is such that the Q_j form a complete set of reasons then the limits

$$\text{Lim}_{n \rightarrow \infty} MD\left(\bigcup_{j=1}^n K(a_j)\right) \left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})\right)$$

exist and equal the canonical solution.

We would conjecture that an analogous result to Theorem 2.1 also holds for MD (and even gives the same answer as ME). This has already been proved in a number of cases although confirming it in full generality remains a topic for future investigation.

Turning now to CM^∞ , its motivation (which was first explained in [9], see also [6]) is rather different from that of MD or ME . Briefly, given a set of constraints K as above (on a probability function Bel on sentences of the language with propositional variables p_1, p_2, \dots, p_n) an initially perhaps rather obvious choice of a particular ‘assigning’ probability function satisfying K might be the ‘most average’ solution to K , or more formally the centre of mass of the polytope of solutions of K (assuming uniform density). Attractive as this choice might appear, based as it is on some idea of indifference, it actually has a serious flaw. Namely language invariance fails. That is, if we instead had considered K as a set of constraints on a probability function

Bel defined on the sentences of some other overlying language, for example the larger language with $n + 1$ propositional variable $p_1, p_2, \dots, p_n, p_{n+1}$, then this centre of mass solution may well not agree with the centre of mass solution for the smaller language on arguments common to both of them. In other words assigned beliefs depend on the chosen overlying language, despite the fact that in the real world we apparently do not consider this to be relevant. This clearly calls into question the intuition behind selecting the centre of mass point in the first place. However some reconciliation is possible by noticing that if we continue to enlarge our language then these centre of mass probability functions do settle down i.e. converge, on their common arguments. The inference process selecting these limiting probability functions, which *does* satisfy language invariance, is denote CM^∞ (the ‘centre of mass as the language size tends to infinity’).

Fortunately this inference process has, as shown in [9], an alternative characterization much more akin to those of ME and MD . Namely, $CM^\infty(K)$ is that solution Bel to K for which the sum

$$\sum_{i \notin I} \log(Bel(\alpha_i))$$

is maximal, where $I = \{ i \mid \text{for all } Bel \text{ satisfying } K, Bel(\alpha_i) = 0 \}$. [It is easy to show, see for example [6] page 74, that this maximum is not $-\infty$, that is that there is a solution Bel to K such that $Bel(\alpha_i) > 0$ for all $i \notin I$.]

For this inference process we can prove the following result:-

Theorem 3.2 If $K(a_1)$ is such that the Q_j form a complete set of reasons then the limits

$$\text{Lim}_{n \rightarrow \infty} CM^\infty \left(\bigcup_{j=1}^n K(a_j) \right) \left(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i}) \right)$$

exist and equal the canonical solution corresponding to a complete set of 3 reasons, Q'_1, Q'_2, Q'_3 , with $\beta_1 = 1, \beta_2 = 1/2, \beta_3 = 0$.

In other words a complete set of reasons again emerges, only in this case, unlike ME and MD , it is not necessarily the complete set $K(a_1)$!

We would conjecture that an analogous result to Theorem 2.1 also holds for CM^∞ (again with a limiting set of 3 complete reasons as in Theorem 3.1), a result we have already proved in a number of cases (see [11]) though not yet in complete generality. Clearly such an answer seems open to criticism in the car/skid scenario since it would imply that once examples of both skidding and non-skidding cars had been observed the conditional probability of any particular future car skidding should be $1/2$, independent of whatever other patterns or propensities had been observed.

4 De Finetti’s Theorem

The fact that each of the inference processes ME, MD, CM^∞ gives limit probabilities corresponding to canonical solutions of complete sets of reasons (at least for the cases so far proven) is intriguing. Is there some common reason for the emergence of ‘reasons’ like this? It is certainly not the case that this behavior is exhibited by all

inference processes. For example, for

$$\alpha = \bigwedge_{i=1}^n p_i^{\epsilon_i} \quad \text{let} \quad \sigma(\alpha) = \frac{1}{(n+1) \binom{n}{t}}$$

where $t = \sum \epsilon_i$, and define the (language invariant) inference process D_2 to select from the solutions to K that probability function Bel which minimizes the ‘cross-entropy’

$$\sum_{i=1}^{2^n} (Bel(\alpha_i) \log(Bel(\alpha_i)/\sigma(\alpha_i))).$$

In this case then, for $K(a_1)$ the empty set of constraints,

$$\text{Lim}_{n \rightarrow \infty} D_2(\bigcup_{j=1}^n K(a_j)) (\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})) = \frac{1}{(r+1) \binom{r}{m}}, \quad (4.1)$$

where $m = \sum \epsilon_i$, and these values do not correspond to the canonical solution of any complete set of reasons (as will be apparent shortly).

There is another way of saying that the limit solutions for ME , MD , CM^∞ correspond to canonical solutions of complete sets of reasons. According to the celebrated theorem of de Finetti, [3], (see also [4]), if B is an *exchangeable* probability function on the sentences of the language with propositional variables $P(a_1), P(a_2), \dots$, that is

$$B(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i}))$$

depends only on r and $\sum \epsilon_i$, then

$$B(\bigwedge_{i=1}^r P^{\epsilon_i}(a_{n_i})) = \int x^{\sum \epsilon_i} (1-x)^{r-\sum \epsilon_i} dF(x)$$

for some normalized measure F on $[0, 1]$. Furthermore the values of B will correspond to the canonical solution of a complete set of reasons just if F is finite discrete, that is if all the measure in F (the λ 's) is concentrated on a finite number of discrete points (the β 's).

Now it is easy to check that in the cases of ME , MD , CM^∞ discussed in this paper the limiting probability functions $\text{Lim}_{n \rightarrow \infty} ME(\bigcup_{j=1}^n K(a_j))$ are exchangeable (essentially because these inferences processes satisfy the *Principle of Renaming*, see [6]) so to say that the solution agrees with a canonical solution of a complete set of reasons is equivalently saying that the corresponding de Finetti measure is finite discrete. [This explains the above example. The values given in (4.1) correspond to the standard (uniform) Lebesgue measure which, of course, is not discrete.]

5 Conclusion

In this paper we have provided a framework and methodology for inductive reasoning as a limiting case of probabilistic propositional uncertain reasoning. We have shown

that in a number of important cases the limit is well defined and furthermore corresponds to the canonical solution based on a complete set of reasons. The problem of fully explaining this phenomenon and ascertaining its persistence for more general knowledge bases remains a subject of research.

The emergence of ‘reasons’ in this fashion is particularly intriguing in the case of the maximum entropy inference process which, we have previously argued, for example in [7], corresponds to the idealization of common sense, and so one might argue, should be normative for intelligent agents like ourselves. Such a conclusion for induction would stand squarely opposed to the conventional Carnapian approach (see for example [5], [1], [2]) based on considerations of symmetry etc. which yield continuous de Finetti measures.

6 Acknowledgements

We would like to thank Graham Little for his considerable help in the proof of Theorem 3.2. We would also like to thank the Egyptian Government for funding the second author during the period of this research (Egyptian Government Scholarship, File No.7083).

References

- [1] Carnap, R., ‘A basic system for inductive logic, part 1’, in *Studies in Inductive Logic and Probability, Volume I*, Eds. R.Carnap & R.C.Jeffrey, University of California Press, Berkeley and Los Angeles, 1971.
- [2] Carnap, R., ‘A basic system for inductive logic, part 2’, in *Studies in Inductive Logic and Probability, Volume II*, Ed. R.C.Jeffrey, University of California Press, Berkeley and Los Angeles, 1980.
- [3] De Finetti, B., ‘Su significato soggettivo della probabilità’, *Fund. Math.*, **17**, 298-329, 1931.
- [4] Hewit, E. & Savage, L.J., ‘Symmetric measures on Cartesian products’, *TAMS*, **8**, 484-489, 1955.
- [5] Johnson, W.E., ‘Probability: The deductive and inductive problems’, *Mind*, **49**, 409-423, 1932.
- [6] Paris, J.B., *The Uncertain Reasoner’s Companion - A Mathematical Perspective*, Cambridge University Press, 1994.
- [7] Paris, J.B., ‘Common sense and maximum entropy’, *Synthese*, **117**, 75-93, 1999.
- [8] Paris, J.B. & Vencovská, A., ‘A note on the inevitability of maximum entropy’, *International Journal of Approximate Reasoning*, **4**, 183-224, 1990.
- [9] Paris, J.B. & Vencovská, A., ‘A method of updating that justifies minimum cross-entropy’, *International Journal of Approximate Reasoning*, **7**, 1-8, 1992.
- [10] Paris, J.B. & Vencovská, A., ‘In defence of the maximum entropy inference process’, *International Journal of Approximate Reasoning*, **17**, 77-103, 1997.
- [11] Paris, J.B., Vencovská, A. & Wafy, M., ‘Some limit theorems for ME , MD and CM^∞ ’, *Technical Report, Manchester Centre for Pure Mathematics*, to appear.
- [12] Shore, J.E. & Johnson, R.W., ‘Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy’, *IEEE Transactions on Information Theory*, **IT-26**(1), 26-37, 1980.
- [13] Wafy, M., *A study of an inductive problem using inference processes*, Ph.D. Thesis, Manchester University, 2000.

Received 3 October, 2000