# The Emergence of Reasons Conjecture

J. B. Paris & A. Vencovská*
Department of Mathematics
University of Manchester
Manchester M13 9PL
UK
jeff@maths.man.ac.uk, alena@maths.man.ac.uk

April 9, 2003

**Keywords** Inductive Logic, Maximum Entropy, Probability Logic, Uncertain Reasoning

**Abstract**

In this paper we show how the maxent paradigm may be used to produce an inductive method (in the sense of Carnap) applicable to a wide class of problems in inductive logic. We then state a conjecture about the answers given by this method and prove the conjecture in a particular case which extends all previous partial confirmations.

## 1 Introduction and Motivation

To motivate the inductive method studied in this paper[1] we first consider a hypothetical (i.e. contrived) example. Suppose I am interested in people's reactions to a cinema film they had seen, in particular whether or not they liked it. I know, or more precisely *believe*, that generally young people like action films, that about 60% of films are comedies, that action films are rarely also funny etc.. I might also be of the opinion that if a young person tells me they liked a film then there is a good chance that it will not have been enjoyed by someone older who saw it.

*Supported by the UK Engineering and Physical Sciences Research Council (EPSRC) research grant GR/M91396

[1]For a more detailed discussion of the ideas and background on which this paper is based we refer the reader to our earlier publications [17] (pages 197-8), [20], [21], [25] and [23].

In the spirit of probability logic (in which framework this paper is set) we might capture this knowledge as a set of linear constraints on the subjective probabilities, equivalently degrees of belief, that I am willing to assign to various events. So for example if $C, A, ...$ stand for 'comedy', 'action film',..., $P(x), Y(x), , ...$ stand for 'likes', 'young', ..., and $a_1, a_2, a_3, ...$ stand for the potentially infinite stream of cinema goers then I might formulate this knowledge as

$Bel(P(a_i) \mid Y(a_i) \wedge A) = 0.8$ *'generally young people like action films'*

$Bel(C) = 0.6$ *'about 60% of films are comedies'*

$Bel(C \mid A) = 0.1$ *'action films are rarely also funny'*

$Bel(P(a_i) \mid \neg Y(a_i) \wedge Y(a_j) \wedge P(a_j)) = 0.3$ *' if a young person tells me they liked a film then there is a good chance that it will not have been enjoyed by someone older who saw it'*

. . . . . . . . . . . etc.

where $i, j = 1, 2, 3, ..., i \neq j$ and, sufficient for the purposes of this paper[2], $Bel$ is a probability function on the set of sentences $SL$ of the *propositional language* $L$ with propositional variables $C, A, ..., P(a_1), P(a_2), P(a_3), ..., Y(a_1), Y(a_2), ...$ etc. That is, for $\theta, \phi \in SL$,

**(P1)** If $\models \theta$ then $Bel(\theta) = 1$.

**(P2)** If $\models \neg(\theta \wedge \phi)$ then $Bel(\theta \vee \phi) = Bel(\theta) + Bel(\phi)$.

Clearly the fact that there are infinitely many $a_i$ here means that this knowledge base is infinite. However we shall assume that it is generated from only finitely many schemata as above by running through the possible $a_i, a_j$ etc. In particular then there are only finitely many propositional variables $C, A, ...$ and finitely many predicates $P(x), Y(x), ...$

The question we are interested in is, on the basis of such a knowledge base, $K$ say, and knowing nothing more, what belief should be assigned to some sentence $\theta \in SL$? For example, what belief should be assigned to some three cinema goers $a_i, a_j, a_k$ liking a film. Or, to put it another way, what value should be assigned to $Bel(P(a_i) \wedge P(a_j) \wedge P(a_k))$? By 'should' here we mean 'on the basis of generally accepted desiderata or principles'.

Whilst it may seem at this point contentious whether or not in this case any such 'principles' exist there is a related situation in which in which we would

[2]By a theorem of Gaifman, [8], $Bel$ extends uniquely to a probability function on the corresponding predicate language.

claim they do. Namely, if we let $L^{(r)}$ be the *finite* sublanguage of $L$ with propositional variables $C, A, ..., P(a_1), P(a_2), ..., P(a_r), Y(a_1), Y(a_2), ..., Y(a_r)$, ... and let $K \upharpoonright L^{(r)}$ be the *finite* set of constraints from $K$ which only mention sentences from $SL^{(r)}$ then according to arguments in [17], [18], [19] there are principles of 'common sense' whose observance uniquely determines the values that may be assigned to $Bel(\theta)$ for $\theta \in SL^{(r)}$. Indeed these uniquely determined belief values turn out in this case to be precisely those given by the *maximum entropy* solution[3] , $ME(K \upharpoonright L^{(r)})$, of $K \upharpoonright L^{(r)}$, that is that (unique) probability function $Bel$ on $SL^{(r)}$ satisfying the constraints in $K \upharpoonright L^{(r)}$ for which the entropy,

$$-\sum_{\alpha \in At^{L^{(r)}}} Bel(\alpha) \log Bel(\alpha),$$

is maximal. Here $At^{L^{(r)}}$ are the *atoms* of $L^{(r)}$, that is the sentences of the form

$$q_1^{\epsilon_1} \wedge q_2^{\epsilon_2} \wedge ... \wedge q_J^{\epsilon_J}$$

where the $q_1, q_2, ..., q_J$ enumerate the (finitely many) propositional variables in $L^{(r)}$, $\epsilon_1, \epsilon_2, ..., \epsilon_J \in \{0, 1\}$ and $q^1 = q$, $q^0 = \neg q$.

Unfortunately these principles cannot immediately be applied to our problem because $L$, and $K$, are infinite. In reality however the number of cinema goers is of course not really infinite, just rather large, so it would seem very reasonable to replace $K$ in our original problem by $K \upharpoonright L^{(r)}$, *for some large* $r$. That is, for $\theta \in SL$ to take $Bel(\theta) = ME(K \upharpoonright L^{(r)})(\theta)$ for some large $r$ (certainly sufficiently large that $\theta \in SL^{(r)}$). The trouble is that this answer depends in general on the large $r$ we choose[4] and there appear to be no clear guidelines as to how exactly this choice should be made. This difficulty however could be avoided if it could be shown that

$$\lim_{r \to \infty} ME(K \upharpoonright L^{(r)})(\theta)$$

always exists, since then we could unambiguously assign this limit value to $Bel(\theta)$.

In [20] and [21] we showed that in certain special cases this limit does exist. That of course is hardly surprising, experience suggests that unless something very funny is going on, naturally arising (bounded) sequences usually do have proper limits. What was surprising however was the form that these limit values take. Namely they agree with the canonical values arising from a 'complete set of reasons'. To explain what precisely we mean by this it seems best at this point to move on from our specific cinema goer example and set up a formal framework.

---

[3]For another approach to predicate uncertain reasoning involving maximum entropy see the work of Halpern et al, [11].

[4]Since $ME$ is *language invariant*, see [17], enlarging the language of itself would not change the value $ME$ assigns. Here however additional constraints are admitted as $r$ increases.

## 2 A Formal Framework

From this point forwards let $q, n, s \in \mathbb{N}$ be fixed and let $L^{(r)}$ be the propositional language with propositional variables $Q_1, Q_2, ..., Q_q$ and $P_1(a_i), P_2(a_i), ..., P_n(a_i)$ for $i = 1, 2, ..., r$. Let $SL^{(r)}$ be the set of sentences of $L^{(r)}$ (using the connectives $\neg, \vee, \wedge$, say) and let $K = K(a_1, a_2, ..., a_s)$ be a finite set of linear constraints

$$\sum_{i=1}^{k} c_i Bel(\theta_i) = d, \tag{1}$$

on a probability function $Bel : SL^{(s)} \to [0,1]$ where the $c_i, d$ are reals and the $\theta_i \in SL^{(s)}$. Let $K^{(r)}$ be the combined set of constraints on a probability function $Bel : SL^{(r)} \to [0,1]$ formed by taking the union over all distinct $i_1, i_2, ..., i_s$ from $\{1, 2, ..., r\}$ of the $K(a_{i_1}, a_{i_2}, ..., a_{i_s})$ (i.e. the result of replacing each $a_j$ everywhere in $K(a_1, a_2, ..., a_s)$ by $a_{i_j}$ for $j = 1, 2, ..., s$). Let $\mathcal{C}_s$ be the set of such $K(a_1, a_2, ..., a_s)$ for which $K^{(r)}$ is satisfiable (by some probability function $Bel : SL^{(r)} \to [0,1]$) for each $r \geq s$.

Clearly the knowledge base we considered in the previous section is, after some trivial rearranging, of this form (we implicitly assumed satisfiability) once we identify a conditional constraint $Bel(\theta \mid \phi) = a$ with

$$Bel(\theta \wedge \phi) - aBel(\phi) = 0.$$

Certain knowledge bases in $\mathcal{C}_s$ will be of especial interest to us. These are the so called 'complete sets of reasons'. Precisely, we say that a knowledge base in $\mathcal{C}_s$ is *a complete set of reasons* if it has the form

$$\begin{array}{rl} \text{(i)} & Bel(P^{\vec{\epsilon}}(a_i) \wedge Q_j) - \beta_j^{\vec{\epsilon}} Bel(Q_j) = 0, \quad j = 1, 2, ..., q, \\ \text{(ii)} & Bel(Q_j \wedge Q_k) = 0, \quad 1 \leq j < k \leq q, \\ \text{(iii)} & Bel(Q_j) = \lambda_j, \quad j = 1, 2, ..., q, \end{array} \tag{2}$$

where $\sum_{j=1}^{q} \lambda_j = 1$, $\sum_{\vec{\epsilon}} \beta_j^{\vec{\epsilon}} = 1$ for $j = 1, 2, ..., q$,

$$P^{\vec{\epsilon}}(a_i) = P_1^{\epsilon_1}(a_i) \wedge P_2^{\epsilon_2}(a_i) \wedge ... \wedge P_n^{\epsilon_n}(a_i)$$

and the $\vec{\epsilon}$ range over $\{0,1\}^n$.

In this case there appears to be a very natural *canonical solution* to the problem of what values to assign to $Bel(\theta)$. Namely, on the basis of the knowledge there *seems* no reason for assuming any 'dependence' between the $P^{\vec{\epsilon}}(a_i) = P_1^{\epsilon_1}(a_i) \wedge P_2^{\epsilon_2}(a_i) \wedge ... \wedge P_n^{\epsilon_n}(a_i)$ given $Q_j$. Interpreting this apparent lack of dependence as *stochastic independence* leads one to set

$$Bel(\bigwedge_{i=1}^{m} P^{\vec{\epsilon^i}}(a_i) \mid Q_j) = \prod_{i=1}^{m} \beta_j^{\vec{\epsilon^i}}$$

and

$$Bel(\bigwedge_{i=1}^{m} P^{\vec{\epsilon_i}}(a_i)) = \sum_{j=1}^{q} \lambda_j \prod_{i=1}^{m} \beta_j^{\vec{\epsilon_i}},$$

where the $\vec{\epsilon_i} \in \{0,1\}^n$. We shall call this the *canonical solution* corresponding to this complete set of reasons $Q_j$.

# The Conjecture

We are now ready to state our conjecture.

**The Emergence of Reasons Conjecture**
Let $K = K(a_1, a_2, ..., a_s) \in \mathcal{C}_s$. Then the

$$\lim_{r\to\infty} ME(K^{(r)})(\bigwedge_{i=1}^{m} P^{\vec{\epsilon_i}}(a_i))$$

exist and agree with the values given by some complete set of reasons.

The short history of this conjecture is as follows. Around 1997 we, along with Maged Wafy, showed that these limits exist in the case $s = 1$, and noticed that furthermore the limit values correspond to those given by some complete set of reasons. Empirical data, courtesy of Gebriele Kern-Isberner in Hagen, subsequently suggested that this was also true for certain simple $s = 2$ cases and encouraged us to attempt to show the general conjecture. Unfortunately this has proved to be a considerably more demanding, and frustrating, challenge (for us at least) than we ever imagined it would be those five years ago. Prior to this paper the state of play was that the conjecture had been proved for the general $s = 1$ case, see [21], and in the 'base case' for $s = 2$, see [23]. Over the next two sections of this paper we shall extend this latter result to show that the conjecture holds in the $s = 2$ case when we only have a single unary predicate $P(x)$, though any finite number propositional variables $Q_1, Q_2, ..., Q_q$.

We apologize in advance for the tortuous nature of this proof, a particularly heart felt sentiment since we actually harbor lingering suspicions that there may already exist very general theorems within measure theory from which this conjecture follows as a special case. One reason for this feeling is that on the basis of a number of special cases this conjecture also seems credible for several other inference processes apart from $ME$, see [2], [20], [21]. [On the other hand, as observed in [25], there are 'reasonable' inference processes for which it fails.] However to date we have not found any such suitable general results in measure theory.

Before proceeding to the proof of this special case we make a few remarks related to the Conjecture. Firstly, whilst as stated it applies specifically to belief in 'state descriptions' $\bigwedge_{i=1}^m P^{\vec{\epsilon^i}}(a_i))$ it easily extends to general $\theta \in SL$[5]. Secondly, as remarked already in [23] these reasons that emerge do not in general appear to be expressible in the original language.

Finally we should comment on the relation of the answers obtained by this process vis-a-vis the classical induction problem of Carnap. The first thing to say is that, to our knowledge, in its mathematical formulation this induction problem has been studied almost solely in the case when $K = \emptyset$. In other words, what belief to give to $\bigwedge_{i=1}^m P^{\vec{\epsilon^i}}(a_i))$ in the absence of any prior knowledge whatsoever.

In its logico-mathematical formulation this problem came particularly to the fore in the work of Carnap, [3], [4] in the early 1950's, indeed Carnap coined the phrase 'Inductive Logic' to describe his formulation. Carnap's approach was to investigate certain natural, desirable, properties that one would expect of a sensible answer to the problem of assigning beliefs, or conditional beliefs (as probabilities) in, what amounts to, the case of our problem when $K = \emptyset$ and to consider what solutions these properties sanction. Most spectacularly, in [6], and independently [15], this approach was shown to lead to a certain family of belief, or probability, functions $Bel_\lambda$ $(0 \leq \lambda \leq \infty)$, called Carnap's $\lambda$-continuum and specified by

$$Bel(P^{\vec{\epsilon}}(a_{m+1}) \mid \bigwedge_{i=1}^m P^{\vec{\epsilon^i}}(a_i)) = \frac{m_{\vec{\epsilon}} + \lambda/2^n}{m + \lambda}$$

where $m_{\vec{\epsilon}} = |\{i | 1 \leq i \leq m,\ \vec{\epsilon} = \vec{\epsilon^i}\}|$. [For a full account of this endeavour and related results by others authors see [5], [13].] In fact, unbeknown to Carnap and his contemporaries this satisfying conclusion had actually already been discovered, and published, by W.E.Johnson nearly 20 years earlier, see [14].

Despite later criticisms of one of the key assumptions (now commonly referred to as Johnson's Sufficientness Principle) and indeed of the possibility of induction at all, see in particular [24], Carnap's continuum of inductive methods continues to have a powerful influence with, to our knowledge, few other even comparably justified alternatives (though see [16]).

Johnson's and Carnap's investigations and conclusions apply, as far as the slant of this paper is concerned, to the case of $K = \emptyset$ (though considerations of *conditional probabilities* certainly played an important role for them). In this $K = \emptyset$ case the method proposed in this paper gives limiting $ME$ values[6] which are just those given by the 'completely independent solution' (see [17]). That is, all the $a_i$ are treated as totally independent, $P^{\vec{\epsilon^i}}(a_i)$ is given belief $2^{-n}$ and $\bigwedge_{i=1}^m P^{\vec{\epsilon^i}}(a_i)$ is given belief $(2^{-n})^m = 2^{-nm}$. This is exactly the solution given

[5]For example, by treating the $Q_k$ as unary predicates and adding to $K$ the constraints $Bel(Q_k(a_i) \leftrightarrow Q_k(a_j)) = 1$. Indeed the proof we give in this paper actually covers the case of general $\theta$ directly without the need for this trick at all.

[6]Actually in this case these are constant up to the limit!

by Carnap's continuum when $\lambda = \infty$, and also by a counting satisfying models approach based on Fagin's Theorem, [7] (also proved independently by Glebskij et al, [9]). (For a considerable extension of this approach to predicate uncertain reasoning see [1], [11], [12].)

In the past we have been somewhat critical of this answer of the completely independent solution on the grounds that, for example, it would mean that if one had observed say $P_1(a_i)$ for $i = 1, 2, ..., 100$ then using conditional probability for prediction one should still give $P(a_{101})$ belief $1/2$, just as one would on the basis of no previous observations. So much for learning from experience!

However we are now of the opinion that this argument is flawed in this context[7]. The reason we feel the argument is flawed is that the justification for $ME$ as embodying common sense reasoning (see [19]) in this context is based on $K$ (as a family of schemata) summing up *all* the available knowledge and the $ME$ values being seen as subjective probabilities based on that knowledge. The point here is that the reason why the above 'unwillingness to learn' looks wrong is that we are *imagining* a situation in which the knowledge has been increased by at least

$$Bel(P(a_i)) = 1, \quad i = 1, 2, ..., 100.$$

But in that case the initial assumption that nothing is known is simply no longer valid, and nor are any beliefs inferred on the basis of the empty knowledge. Indeed it is clear that the new knowledge here is not simply $Bel(P(a_i)) = 1$, for $i = 1, 2, ..., 100$. The major component of the new knowledge is that in 100 supposedly random 'trials' $P(a_i)$ held each time. We have no idea how such knowledge could be incorporated into the present framework.

It should be pointed out however that considerations of *conditional probabilities* such as $Bel(P(a_{101})|P(a_1) \wedge ... \wedge P(a_{100}))$ where this evidence is understood in this way commonly does play an influential role in motivating, for example, Johnson's Sufficientness Principle. In comparing approaches therefore it would seem to be important to appreciate such differences in perspective, in particular that as far as this paper is concerned beliefs are intended as subjective probabilities, the creations of an agent without any pretensions to approximating 'objective truth'[8].

Of course the problem of induction has been studied, indeed generally is studied in philosophy, in the context of the real world, where in fact we already have an enormous quantity of relevant knowledge, just through living in the

[7]We are grateful to Colin Howson for conversations which forced us to re-evaluate this argument.

[8]Of course in contrived examples such as the one we give to motivate the main problem in this paper the difference between 'objective' and 'subjective' is scarcely black and white, so, like props on a film set, they should simply be seen as furnishing a context rather than anything capable of actually being leant on.

world, even when we might be inclined to throw up our hands and claim we know nothing. Whether the situation envisaged in the Conjecture at all corresponds to such a reality is a point we certainly would not wish to argue for. Nevertheless we feel that it is of some interest that in this highly idealized situation the world can be satisfactorily explained by 'inventing reasons' since we certainly seem to observe this same phenomenon being applied in the infinitely more complicate world wherein we lead our daily lives.

Having dallied long enough it is now time to roll up our sleeves and do some mathematics.

# Proof of the $s = 2$ base case of the Emergence of Reasons Conjecture

Our aim over this section and the next is to give a proof of the $s = 2$ single predicate special case of Conjecture, namely:

**Theorem 1** *Suppose that the language $L$ has a single predicate $P(x)$ and propositional variables $Q_1, Q_2, ..., Q_q$. Let $K = K(a_1, a_2) \in \mathcal{C}_2$. Then the*

$$\lim_{r \to \infty} ME(K^{(r)})(\bigwedge_{i=1}^{m} P^{\epsilon_i}(a_i))$$

*exist and agree with the values given by some complete set of reasons.*

It is a a straightforward corollary of this theorem that it holds equally well for any sentence $\theta(a_i)$ from $SL$ in place of $P(a_i)$. That is, that there are some

$$0 \leq \lambda_1, \lambda_2, ..., \lambda_t, \beta_1, \beta_2, ..., \beta_t \leq 1$$

such that

$$\lim_{r \to \infty} ME(K^{(r)})(\theta^{\epsilon_1}(a_1) \wedge \theta^{\epsilon_2}(a_2) \wedge ... \wedge \theta^{\epsilon_m}(a_m)) = \sum_{i=1}^{t} \lambda_i \beta_i^{\sum \epsilon_j}(1 - \beta_i)^{t - \sum \epsilon_j}.$$

The plan in this section is to first to prove the 'base' case[9] of Theorem 1 and then complete the proof of the full theorem in the following section.

The special base case we shall limit ourselves to considering in this section is when $q = 0$ (i.e. there are no $Q_i$) and $K = K(a_1, a_2)$ has the special form

$$\{ \; Bel(P(a_1)) = a, \;\; Bel(P(a_1) \wedge P(a_2)) = b \; \}. \tag{3}$$

[9]A proof of this case already appears in [23]. However for what follows in this paper we need a rather more detailed treatment than was required on that occasion.

In this case then

$$K^{(r)} = \{Bel(P(a_i)) = a,\ Bel(P(a_i) \wedge P(a_j)) = b \mid 1 \leq i, j \leq r,\ i \neq j\}. \quad (4)$$

In the later parts of the proof of Theorem 1 we shall work with knowledge bases which are consistent for all $r$, i.e. in $\mathcal{C}_2$. For the present however we want to consider the base case at a particular, albeit assumed large, $r$. It will be useful to have the following equivalent to $K^{(r)}$ being satisfiable for this base case.

**Lemma 2** *$K^{(r)}$ is satisfiable if and only if*

$$0, \frac{[ra](2ra - 1 - [ra])}{r(r-1)} \leq b \leq a \leq 1.$$

**Proof.** First assume that $K^{(r)}$ is consistent. The only non-trivial inequality here is the second lower bound on $b$. Let $Bel$ be some solution of $K^{(r)}$. Without loss of generality we may assume that $Bel$ is exchangeable (this just makes (5) below more obvious, though it actually holds in any case). For $0 \leq i \leq r$ let

$$X_i = Bel\left( \bigvee_{\substack{\vec{\epsilon} \in \{0,1\}^r \\ \sum \epsilon_j = i}} \bigwedge_{1 \leq t \leq r} P^{\epsilon_t}(a_t) \right).$$

Then since $Bel$ satisfies the constraints in $K^{(r)}$,

$$\sum_{i=0}^{r} X_i = 1, \quad \sum_{i=0}^{r} \frac{i}{r} X_i = a \ \text{ and } \ \sum_{i=0}^{r} \frac{i(i-1)}{r(r-1)} X_i = b. \quad (5)$$

Expanding the last of these identities shows that for a fixed $a$, $b$ takes its smallest possible value when $\sum_i i^2 X_i$ takes its smallest possible value, subject to the constraints (5). For this to happen it is not possible to have $X_j, X_k > 0$ with $j < p < k$ since then we could replace $X_j, X_p, X_k$ with

$$0,\ X_p + X_j\left(\frac{k-j}{k-p}\right), X_k - X_j\left(\frac{p-j}{k-p}\right),$$

if $X_k(k-p) \geq X_j(p-j)$ and

$$X_j - X_k\left(\frac{p-k}{j-p}\right),\ X_p + X_k\left(\frac{j-k}{j-p}\right), 0,$$

otherwise, and obtain a probability function with the same mean $a$ but smaller mean square. From this it follows that when $\sum_i i^2 X_i$ takes its minimum value at most $X_{[ra]}$ and $X_{[ra]+1}$ can be non-zero and from that we can solve to find the value as above for the minimum $b$.

In the other direction we simply reverse the process to find a solution $Bel_0$ from these, uniquely determined by (5), minimizing $X_{[ra]}$ and $X_{[ra]+1}$. Now it is straightforward to see that if we take $b = a$ then there is a solution $Bel_1$ to (5) and clearly this is the largest $b$ for which there is such a solution. From $Bel_0, Bel_1$ we can now form convex combinations with, necessarily, the same $a$ value and any required $b$ value between

$$\min\left\{0, \frac{[ra](2ra-1-[ra])}{r(r-1)}\right\}$$

and $a$, as required. ■

Henceforth for this section assume that is $K^{(r)}$ is satisfiable and $r$ is large, so in particular Lemma 2 gives that,

$$0, \frac{[ra](2ra-1-[ra])}{r(r-1)} \leq b \leq a \leq 1. \tag{6}$$

Let

$$W_i = ME(K^{(r)})\left(\bigvee_{\substack{\vec{\epsilon}\in\{0,1\}^r \\ \sum \epsilon_j = i}} \bigwedge_{1\leq t\leq r} P^{\epsilon_t}(a_t)\right).$$

Since $ME(K^{(r)})$ satisfies the constraints in $K^{(r)}$,

$$\sum_{i=0}^{r} W_i = 1, \quad \sum_{i=0}^{r} \frac{i}{r} W_i = a \quad \text{and} \quad \sum_{i=0}^{r} \frac{i(i-1)}{r(r-1)} W_i = b. \tag{7}$$

Now assume until further mention that all the inequalities in (6) are strict, that is

$$0, \frac{[ra](2ra-1-[ra])}{r(r-1)} < b < a < 1. \tag{8}$$

In this case $K^{(r)}$ has a strictly positive solution. To see this notice that for a fixed value of $a$ there are solutions giving $b$ values $[ra](2ra-1-[ra])r^{-1}(r-1)^{-1}$ and $a$ and a *positive solution* giving $b$ value $a^2$. From these one can form a non zero convex combination with, necessarily, the same $a$ value and and required $b$ value strictly between

$$\min\left\{0, \frac{[ra](2ra-1-[ra])}{r(r-1)}\right\}$$

and $a$.

It now follows by the Open Mindedness Principle (see [17]) that $ME(K^{(r)})$ is also positive and this ensures that the method of Lagrange Multipliers may be validly employed to give that,

$$ME(K^{(r)})\left(\bigwedge_{j=1}^{i} P(a_j) \wedge \bigwedge_{j=i+1}^{r} \neg P(a_j)\right) = xy^i z^{i^2} \tag{9}$$

for some $x, y, z > 0$[10]. To see this notice that by the Renaming Principle (see [17] or [18]), $ME(K^{(r)})\left(\bigwedge_{j=1}^{r} P^{\epsilon_j}(a_j)\right)$ depends only on $\sum_{j=1}^{r} \epsilon_j$, so our system of constraints corresponds to

$$\sum_{i=0}^{r} \binom{r}{i} u_i = 1, \quad \sum_{i=0}^{r} \binom{r}{i} \frac{i}{r} u_i = a \text{ and } \sum_{i=0}^{r} \binom{r}{i} \frac{i(i-1)}{r(r-1)} u_i = b \tag{10}$$

where $u_i$ denotes $Bel\left(\bigwedge_{j=1}^{r} P^{\epsilon_j}(a_j)\right)$ for $\sum_{j=1}^{r} \epsilon_j = i$. The Lagrange Multiplier method can correctly be applied to show that

$$-1 - \log u_i - \lambda - \frac{i}{r}\mu - \frac{i(i-1)}{r(r-1)}\nu = 0$$

for some Lagrange Multipliers $\lambda, \mu, \nu$. Hence

$$u_i = e^{-1-\lambda}\left(e^{-\frac{\mu}{r}} e^{\frac{\nu}{r(r-1)}}\right)^i \left(e^{-\frac{\nu}{r(r-1)}}\right)^{i^2}$$

which gives the above identity (9).

For $x, y, z$ as specified in (9) and $i = 0, 1, \ldots r$, define

$$Z_i = \binom{r}{i} y^i z^{i^2}. \tag{11}$$

We shall now spend some time investigating the properties of these $Z_i$ for large $r$.

By Stirling's formula, for $0 < \alpha < 1$, and $r\alpha$ a natural number,

$$\begin{aligned}
\log Z_{r\alpha} \;=\; & -r \log \alpha^{\alpha}(1-\alpha)^{1-\alpha} + r\alpha \log y + r^2\alpha^2 \log z \\
& -1/2 \log(\alpha(1-\alpha)) - 1/2 \log 2\pi r \\
& -r^{-1}h(r) + (r\alpha)^{-1}h(r\alpha) + (r(1-\alpha))^{-1}h(r(1-\alpha))
\end{aligned} \tag{12}$$

where $h(x)$ is a power series in $x^{-1}$. Now for $0 < t < 1$ let

$$H_r(t) = -r \log t^t (1-t)^{1-t} + rt \log y + r^2 t^2 \log z - 1/2 \log(t(1-t))$$

[10] These $x, y, z$ are of course functions of $r, a, b$. However to keep the notation simple we shall, throughout, avoid explicit mention of such dependencies whenever possible.

so that

$$\begin{aligned}
\frac{dH_r(t)}{dt} &= -r\log t + r\log(1-t) + r\log y + 2r^2 t\log z - 1/2t + 1/2(1-t),\\
\frac{d^2H_r(t)}{dt^2} &= -r/t - r/(1-t) + 2r^2\log z + 1/2t^2 + 1/2(1-t)^2,\\
\frac{d^3H_r(t)}{dt^3} &= r/t^2 - r/(1-t)^2 - 1/t^3 + 1/(1-t)^3.
\end{aligned}$$

Then $d^3H_r(t)/dt^3$ has a zero at $1/2$ (at which point it is decreasing) and two others close to 0 and 1. From this it follows that $H_r(t)$ can have at most two maxima in $[0,1]$, one on each side of $1/2$.

**Lemma 3** *Let $k$ be a natural number. Then there is an $r_k$ such that if $0<\alpha<\alpha+1/k<\beta<1$, $r\geq r_k$ and the $Z_i$ are decreasing (increasing) for $r\alpha\leq i\leq r\beta$ then $Z_{[r\alpha]}>r^3Z_{[r\beta]}$ ($Z_{[r\beta]}>r^3Z_{[r\alpha]}$).*

**Proof.** Since there is a finite bound on the number of zeros that $d^2H_r/dt^2$, $d^3H_r/dt^3$, $-1/t-1/(1-t)$ and $1/t^2+1/(1-t)^2$ can have on $[0,1]$ *independent* of the values of $x,y,z$ there is a fixed $j$, independent of $x,y,z,\alpha,\beta,r$ such that there is some subinterval $[l/jk,(l+1)/jk]$ of $[\alpha,\alpha+1/k]$ on which $d^2H_r/dt^2$, $d^3H_r/dt^3$, $-1/t-1/(1-t)$ and $1/t^2+1/(1-t)^2$ all have constant sign. Let $c_k$ be the minimum of the moduli of the difference of the values of

$$-1/t-1/(1-t)$$

at the endpoints of an interval of length $1/jk$ on which it and its derivative have constant sign. Then for large $r$ the values of $d^2H_r/dt^2$ at successive endpoints of subintervals $[s/2jk,(s+1)/2jk]$ of $[l/jk,(l+1)/jk]$ must disagree by at least $rc_k/2$, so on at least one such interval $|d^2H_r/dt^2|$ must be at least $rc_k/2$. Since $d^2H_r/dt^2$ has constant sign on this interval $|dH_r/dt|$ must be at least $rc_k/2\times 1/6jk$ on some subinterval $[q/6jk,(q+1)/6jk]$ of $[s/2jk,(s+1)/2jk]$.

From this it follows that (in the case of the $Z_i$ decreasing),

$$\begin{aligned}
\log Z_{[r\alpha]} - \log Z_{[r\beta]} &\geq 1/3(H_r((p+1)/36jk) - H_r(p/36jk))\\
&\geq rc_k/6^3j^2k^2\\
&\geq \log r^3
\end{aligned}$$

since $r$ is large, as required. ■

In [23] an alternate, longer, but more elementary proof was given of this result. We favour the version given here for this paper because of the expectation that it will more easily generalize.

Let $m_1$ and $m_2$ be the two points at which the maxima of the $W_i$ are attained (if there is only one, take that for $m_1$ and leave $m_2$ undefined). Let $e_1 = m_1/r$ and $e_2 = m_2/r$. As shown in [23] if both exist then $e_1 \leq 1/2 \leq e_2$. [It follows from our earlier considerations of the function $H_r(t)$ that these maxima cannot both be far from 1/2 and on the same side of it. In fact this is all we need, though for convenience we shall use the stronger result proved in [23].]

**Corollary 4** *There is a function $\delta(r)$, independent of $e_1, e_2, a, b$, such that $\delta(r) \to 0$ as $r \to \infty$ and $Z_{r\alpha} \leq 1/r^3 Z_{r\beta}$ whenever the $Z_i$ are increasing on $[r\alpha, r\beta]$ (or decreasing on $[r\beta, r\alpha]$) and $|\alpha - \beta| \geq \delta(r)$.*

Clearly in future applications we may assume that $\delta(r) > 1/\sqrt{r}$ for all $r$. Since we shall also be needing various powers of $\delta(r)$ in what follows we introduce the notation $\delta_n(r)$ to stand for $\delta(r)^{1/n}$.

Let

$$Y_1 = \sum_{|i-re_1| \leq r\delta(r)} W_i, \qquad Y_2 = \sum_{\substack{|i-re_2| \leq r\delta(r) \\ i-re_1 > r\delta(r)}} W_i.$$

From (7) we obtain

$$Y_1 + Y_2 = 1 + O(1/r), \tag{13}$$
$$e_1 Y_1 + e_2 Y_2 = a + O(\delta(r)), \tag{14}$$
$$e_1^2 Y_1 + e_2^2 Y_2 = b + O(\delta(r)), \tag{15}$$

with the obvious omissions if $e_2$ is not defined.

**Lemma 5** *If $e_1, e_2$ are both defined then either $Y_1 > r^2 Y_2$ or $Y_2 > r^2 Y_1$ or ,*

$$e_2 = 1 - e_1 + O(\delta(r)). \tag{16}$$

**Proof.** The proof turns on looking at the $W_i$ around $r/2$. At this point the $W_i$ may be increasing, decreasing or 'level', by which we mean that the relation $\circ$ below is $<, >$ or $=$ respectively:

$$W_{\frac{r-1}{2}} \circ W_{\left(\frac{r-1}{2}+1\right)} \qquad \text{if } r \text{ is odd}$$

$$W_{\left(\frac{r}{2}-1\right)} \circ W_{\left(\frac{r}{2}+1\right)} \qquad \text{if } r \text{ is even.}$$

Note that the above implies $1 \circ yz^r$ (whether $r$ is odd or even). First suppose that the $W_i$ are increasing (around $r/2$) for arbitrarily large $r$. That means that (for such $r$)

$$\frac{W_{((r-i)+1)}}{W_{(r-i)}} = y^2 z^{2r} \frac{W_{(i-1)}}{W_i} > \frac{W_{(i-1)}}{W_i} \qquad i = 1, 2, ..., r/2(\text{ or } (r-1)/2) \tag{17}$$

so $e_2 \geq 1 - e_1$ (the $W_{(\frac{r}{2}+i)}$, or $W_{(\frac{r+1}{2}+i)}$ if $r$ is odd, cannot start descending (again) before the $W_{(\frac{r}{2}-i)}$, or $W_{(\frac{r-1}{2}-i)}$, start descending). On the other hand, if $e_2 > 1 - e_1 + 2\delta(r)$ then

$$m_2 > r - m_1 + 2r\delta(r). \tag{18}$$

Now by induction on $i \leq r/2$ from (17) we obtain that $W_{\frac{r}{2}+i} \geq W_{\frac{r}{2}-i}$ for $r$ even (and the corresponding result for $r$ odd) since if this holds for $i$ then

$$W_{\frac{r}{2}+i+1} = y^2 z^{2r} W_{\frac{r}{2}-i-1} \frac{W_{\frac{r}{2}+i}}{W_{\frac{r}{2}-i}} \geq y^2 z^{2r} W_{\frac{r}{2}-i-1} \geq W_{\frac{r}{2}-i-1}.$$

In particular then $W_{(r-m_1)} \geq W_{m_1}$. Together with (18) and the definitions of $\delta(r)$, $m_1$, $m_2$ and $Y_1$, this gives that

$$Y_2 \geq W_{m_2} > r^3 W_{r-m_1} \geq r^3 W_{m_1} \geq r^2 Y_1$$

for arbitrarily large $r$. The same obtains similarly when the $W_i$ are decreasing and $e_2$ is not within $2\delta(r)$ of $1 - e_1$. (When the $W_i$ are level, they are entirely symmetric around $r/2$ so $e_2$ actually equals $1 - e_1$.) ■

Consider now the equations (13), (14), (15). If $e_1, e_2$ are both defined, $e_2 - e_1 > \delta_3(r)$ and $Y_1 \leq r^2 Y_2, Y_2 \leq r^2 Y_1$ then from these equations

$$\begin{aligned}
e_1 &= \frac{1}{2}(1 - \sqrt{1 - 4(a-b) + O(\delta(r))}) + O(\delta(r)), &(19)\\
e_2 &= \frac{1}{2}(1 + \sqrt{1 - 4(a-b) + O(\delta(r))}) + O(\delta(r)), &(20)\\
Y_1 &= \frac{1}{2}\left(1 - \frac{2a-1}{\sqrt{1 - 4(a-b) + O(\delta(r))}}\right) + O(\delta(r)), &(21)\\
Y_2 &= \frac{1}{2}\left(1 + \frac{2a-1}{\sqrt{1 - 4(a-b) + O(\delta(r))}}\right) + O(\delta(r)). &(22)
\end{aligned}$$

Furthermore since $e_2 - e_1 > \delta_3(r)$ it must be the case that $1 - 4(a-b) > \delta_3(r)^2$ so

$$\begin{aligned}
e_1 &= \frac{1}{2}(1 - \sqrt{1 - 4(a-b)}) + O(\delta_2(r)), &(23)\\
e_2 &= \frac{1}{2}(1 + \sqrt{1 - 4(a-b)}) + O(\delta_2(r)), &(24)\\
Y_1 &= \frac{1}{2}\left(1 - \frac{2a-1}{\sqrt{1 - 4(a-b)}}\right) + O(\delta_2(r)), &(25)\\
Y_2 &= \frac{1}{2}\left(1 + \frac{2a-1}{\sqrt{1 - 4(a-b)}}\right) + O(\delta_2(r)). &(26)
\end{aligned}$$

Now suppose in addition to the above assumptions that $b \leq a^2 + \delta_2(r)$. Then, using Lemma 2,

$$\sqrt{1-4(a-b)+O(\delta(r))} = |1-2a| + O(\delta_4(r)),$$

so either $Y_1$ or $Y_2$ (say $Y_2$) is $O(\delta_4(r))$ and (13), (14), (15) reduce to

$$Y_1 = 1 + O(\delta_4(r)), \tag{27}$$
$$e_1 Y_1 = a + O(\delta_4(r)), \tag{28}$$
$$e_1^2 Y_1 = b + O(\delta_4(r)). \tag{29}$$

Furthermore a similar conclusion follows when $Y_1 > r^2 Y_2$ or $Y_2 > r^2 Y_1$ or $e_2 - e_1 \leq \delta_3(r)$ or $e_2$ does not exist.

In summary then we see that there are essentially two, not necessarily exclusive, situations. Either $Y_1 \leq r^2 Y_2, Y_2 \leq r^2 Y_1$, $e_1$, $e_2$ both exist, $e_2 - e_1 > \delta_3(r)$, $1 - 4(a-b) > \delta_3(r)^2$, $b > a^2 + \delta_2(r)$ and (23), (24),(25),(26), hold, or

$$\sum_{|i-ar| \leq r\delta_8(r)} W_i \geq 1 - \delta_5(r), \tag{30}$$

in which case it must be that $b \leq a^2 + \delta_8(r)$ and

$$a = \frac{1}{2}(1 \pm \sqrt{|1-4(a-b)|}) + O(\delta_{16}(r)). \tag{31}$$

**Corollary 6** *For fixed $m < r$*

$$ME(K^{(r)})\left(\bigwedge_{j=1}^{i} P(a_j) \wedge \bigwedge_{j=i+1}^{m} \neg P(a_j)\right) = \lambda_1 \beta_1^i (1-\beta_1)^{m-i} + \lambda_2 \beta_2^i (1-\beta_2)^{m-i} + O(\delta_{16}(r)),$$

*where*

$$\beta_1 = \frac{1}{2}(1 - \sqrt{1-4(a-b)}),$$
$$\beta_2 = \frac{1}{2}(1 + \sqrt{1-4(a-b)}),$$
$$\lambda_1 = \frac{1}{2}\left(1 - \frac{2a-1}{\sqrt{1-4(a-b)}}\right),$$
$$\lambda_2 = \frac{1}{2}\left(1 + \frac{2a-1}{\sqrt{1-4(a-b)}}\right),$$

*if* $b > a^2 + \delta_8(r)$ *and*

$$\beta_1 = a, \quad \beta_2 = 0, \quad \lambda_1 = 1, \quad \lambda_2 = 0,$$

*if* $b \leq a^2 + \delta_8(r)$.

**Proof.** As in (10),

$$\begin{aligned}
&ME(K^r)\left(\bigwedge_{j=1}^{i} P(a_j) \wedge \bigwedge_{j=i+1}^{m} \neg P(a_j)\right) \\
&= \sum_{j=0}^{r-m} \binom{r-m}{j} xy^{(j+i)} z^{(j+i)^2} \\
&= \sum_{j=i}^{r-m+i} \binom{r-m}{j-i} xy^j z^{j^2} \\
&= \sum_{j=0}^{r} \binom{r}{j} \frac{j(j-1)...(j-i+1)}{r(r-1)...(r-i+1)} \frac{(r-j)(r-j-1)...(r-j-m+i+1)}{(r-i)(r-i-1)...(r-m+1)} xy^j z^{j^2} \\
&= \sum_{j=0}^{r} \frac{j(j-1)...(j-i+1)}{r(r-1)...(r-i+1)} \frac{(r-j)(r-j-1)...(r-j-m+i+1)}{(r-i)(r-i-1)...(r-m+1)} W_j.
\end{aligned}$$

In the case $b \leq a^2 + \delta_8(r)$ (30) holds and the result follows by restricting the summation to $|i - ar| \leq r\delta_8(r)$. In the case $b > a^2 + \delta_8(r)$ the result follows by splitting the summation into the ranges $|i - re_1| \leq r\delta(r)$ and $|i - re_2| \leq r\delta(r)$. ■

This concludes the proof of the base case. However, before turning to the next case, where we have just one predicate, $P(x)$, but a general knowledge base, it will be useful to consider the actual entropy for the set of constraints $K^{(r)}$ of this section (strangely that has not been necessary up to this point).

Using (9) the entropy associated with the solution $ME(K^{(r)})$ of $K^{(r)}$ is given by

$$E(a, b, r) = -\sum_{i=0}^{r} \binom{r}{i} xy^i z^{i^2} \log xy^i z^{i^2}. \tag{32}$$

Using (7) and (11) we can rewrite this as

$$-\sum_{i=0}^{r} \frac{Z_i}{\sum_{j=0}^{r} Z_j} \left\{ \log\left(\frac{Z_i}{\sum_{j=0}^{r} Z_j}\right) - \log\binom{r}{i} \right\}. \tag{33}$$

**Lemma 7** *Under the assumption that* $b > a^2 + \delta_4(r)$,

$$E(a, b, r) = -r(e_1 \log e_1 + (1 - e_1)\log(1 - e_1) + O(\delta_4(r))).$$

**Proof.** Under this assumption $e_1, e_2$ are both defined and $Y_1 > r^2 Y_2$ etc. and we can express (33) as the sum of six terms, $B_1, B_2, B_3, B_4, B_5, B_6$ where

$$\begin{aligned}
B_1 &= \sum_{|i-e_1r|\le r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \binom{r}{i}, \\
B_2 &= \sum_{\substack{|i-e_2r|\le r\delta(r)\\ i>e_1r+r\delta(r)}} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \binom{r}{i}, \\
B_3 &= -\sum_{|i-e_1r|\le r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \left( \frac{Z_i}{\sum_{j=0}^r Z_j} \right), \\
B_4 &= -\sum_{\substack{|i-e_2r|\le r\delta(r)\\ i>e_1r+r\delta(r)}} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \left( \frac{Z_i}{\sum_{j=0}^r Z_j} \right), \\
B_5 &= \sum_{|i-e_1r|,|i-e_2r|>r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \binom{r}{i}, \\
B_6 &= -\sum_{|i-e_1r|,|i-e_2r|>r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log \left( \frac{Z_i}{\sum_{j=0}^r Z_j} \right).
\end{aligned}$$

We shall show that $B_3, B_4, B_5, B_6$ are all small compared with $B_1 + B_2$. .

For $i$ in the range of summation of $B_5$

$$\frac{Z_i}{\sum_{j=0}^r Z_j} = W_i \le \frac{1}{r^2}$$

(by choice of $\delta(r)$) whilst the factors

$$\log \binom{r}{i} \le r \log 2,$$

since $\binom{r}{i} \le 2^r$ so $|B_5| \le r^{-1} \log 2$.

Turning to $B_3$, since the maximum of the function $-x \log x$ on $[0, 1]$ is taken at $x = e^{-1}$ this sum is clearly no more than

$$-\sum_{|i-e_1r|\le r\delta(r)} e^{-1} \log e^{-1} = e^{-1}(2r+1)\delta(r). \tag{34}$$

A similar bound, of course, also applies to $B_4$.

Finally turning to $B_6$, for $i$ in the range of summation

$$r^2 Z_i \le \sum_{j=0}^r Z_j$$

so the sum $B_6$ is no more than

$$-\sum_{|i-e_1r|,|i-e_2r|>\delta(r)} 1/r^2\log(1/r^2) \leq \frac{2}{r}\log r < \delta(r).$$

We now consider the terms $B_1$ and $B_2$. By the mean value theorem

$$B_1 = \sum_{|i-e_1r|\leq r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j}\log\binom{r}{u},$$

for some *real* $u$ with $|u-e_1r| \leq r\delta(r)$. By Stirling's formula

$$\begin{aligned}\log\binom{r}{u} &= -r\log(v^v(1-v)^{1-v}) - 1/2\log(2\pi rv(1-v)) \\ &\quad -u^{-1}h(u) - (r-u)^{-1}h(r-u) + r^{-1}h(r)\end{aligned}$$

where $v = u/r$ and $h(x)$ is a power series in $x^{-1}$. Hence

$$\begin{aligned}B_1 &= \sum_{|i-e_1r|<r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j}\big(-r\log(v^v(1-v)^{1-v}) - 1/2\log(2\pi rv(1-v)) \\ &\quad -u^{-1}h(u) - (r-u)^{-1}h(r-u) + r^{-1}h(r)\big). \qquad (35)\end{aligned}$$

Since $|u-e_1r| \leq r\delta(r)$ and $1/r < \delta(r)$ this gives in turn that provided $e_1$ is sufficiently far from 0 or 1, say at least $\delta_2(r)$, then,

$$\log(v^v(1-v)^{1-v}) = \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(\delta_2(r)).$$

Using this and suitable estimates for the other (relatively small) terms in (35) gives

$$B_1 = -\sum_{|i-e_1r|\leq r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} r\log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_2(r)). \qquad (36)$$

On the other hand if $e_1 < \delta_2(r)$ then both $-\log(e_1^{e_1}(1-e_1)^{1-e_1})$ and $-\log(v^v(1-v)^{1-v})$ are less than $-2\delta_2(r)\log\delta_2(r)$ which in turn is itself less than $\delta_4(r)$. Hence

$$B_1 = -\sum_{|i-e_1r|\leq r\delta(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} r\log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_4(r)). \qquad (37)$$

Turning to $B_2$, we can similarly obtain

$$B_2 = -\sum_{\substack{|i-e_2r|\leq r\delta(r)\\ i>e_1r+r\delta(r)}} \frac{Z_i}{\sum_{j=0}^r Z_j} r\log(e_2^{e_2}(1-e_2)^{1-e_2}) + O(r\delta_4(r)). \qquad (38)$$

In this case $e_2 = 1 - e_1 + O(\delta(r))$ so again provided $e_1 \geq \delta_2(r)$,

$$\log(e_2^{e_2}(1-e_2)^{1-e_2}) = \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(\delta_2(r)) \tag{39}$$

and substituting this into (38) gives

$$B_2 = -\sum_{\substack{|i-e_2 r|\leq r\delta(r)\\ i>e_1 r + r\delta(r)}} \frac{Z_i}{\sum_{j=0}^r Z_j} r \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_4(r)). \tag{40}$$

On the other hand if $e_1 < \delta_2(r)$ then $1 - e_2 < 2\delta_2(r)$ and as in (37) $B_2 = O(r\delta_4(r))$. Either way then we obtain that

$$\begin{aligned}
\sum_{j=1}^{6} B_j &= B_1 + B_2 + O(r\delta_4(r)) \\
&= -\sum_{\substack{|i-e_1 r|\leq r\delta(r)\\ \text{or } |i-e_2 r|\leq r\delta(r)}} \frac{Z_i}{\sum_{j=0}^r Z_j} r \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_4(r)) \\
&= -(Y_1 + Y_2) r \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_4(r)) \\
&= -r \log(e_1^{e_1}(1-e_1)^{1-e_1}) + O(r\delta_4(r))
\end{aligned}$$

by (13). ■

**Lemma 8** *Under the assumption that $b \leq a^2 + \delta_8(r)$,*

$$E(a,b,r) = -r(a\log a + (1-a)\log(1-a)) + O(r\delta_{64}(r)).$$

**Proof.** The proof is very similar to the proof of Lemma 7 but using (30). This time we consider $B_1, B_2, B_3, B_4$ where

$$\begin{aligned}
B_1 &= \sum_{|i-ar|\leq r\delta_8(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log\left\{\binom{r}{i}\right\}, \\
B_2 &= -\sum_{|i-ar|\leq r\delta_8(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log\left(\frac{Z_i}{\sum_{j=0}^r Z_j}\right), \\
B_3 &= \sum_{|i-ar|> r\delta_8(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log\left\{\binom{r}{i}\right\}, \\
B_4 &= -\sum_{|i-ar|> r\delta_8(r)} \frac{Z_i}{\sum_{j=0}^r Z_j} \log\left(\frac{Z_i}{\sum_{j=0}^r Z_j}\right).
\end{aligned}$$

Using the same tricks as in that previous proof we can again show that $B_2, B_3, B_4$ are small $(O(\delta_8(r)))$ whilst

$$B_1 = -r(a\log a + (1-a)\log(1-a)) + O(r\delta_{64}(r)).$$

We leave the details to the reader! ■

**Corollary 9**

$$E(a,b,r) = rT(a-b) + O(r\delta_{64}(r)).$$

*where $T(c)$ is the function*

$$-\frac{1}{2}(1-\sqrt{|1-4c|})\log(\frac{1}{2}(1-\sqrt{|1-4c|}))-\frac{1}{2}(1+\sqrt{|1-4c|})\log(\frac{1}{2}(1+\sqrt{|1-4c|})).$$

Note that the modulus signs in this expression are only required in the case that $c > 1/4$. A necessary condition for this to happen with $c = a-b$ is that $a$ is close to $1/2$ and $b < a^2$, in which case $e_1$ will be within $O(\delta_4(r))$ of $1/2$. Thus if this condition does not hold we can safely remove the modulus signs here.

**Proof.** In the case that (30) holds this follows from Lemma 8 and (31). Otherwise the result follows from Lemma 7. ■

In what follows let $c = a - b$.

**Lemma 10** *Suppose that $e_2$ is defined, $Y_1 \leq r^2 Y_2$, $Y_2 \leq r^2 Y_1$, $e_2 - 1/2 \geq \delta_4(r)$ and $e_1 > \delta_2(r)$. Then*

$$\begin{aligned} E(a,b,r) &= rT(c) + rG(c,r) + \\ & -\frac{1}{2}(1+\frac{1-2a}{\sqrt{1-4c}})\log\left(\frac{1}{2}(1+\frac{1-2a}{\sqrt{1-4c}})\right) + \\ & -\frac{1}{2}(1-\frac{1-2a}{\sqrt{1-4c}})\log\left(\frac{1}{2}(1-\frac{1-2a}{\sqrt{1-4c}})\right) \end{aligned}$$

*where $G(c,r) \to 0$ as $r \to \infty$.*

**Proof.** From (9) and (7)

$$\sum_{i=0}^{r} \binom{r}{i} xy^i z^{i^2} = 1.$$

Differentiating with respect to $a$ with $c$ fixed[11] gives, with (7),

$$\frac{1}{x}\frac{\partial x}{\partial a} + \frac{ra}{y}\frac{\partial y}{\partial a} + \frac{r(r-1)b + ra}{z}\frac{\partial z}{\partial a} = 0. \tag{41}$$

Again using (7) with (32) we have that

$$E(a,b,r) = -\log x - ra\log y - (r(r-1)b + ra)\log z, \tag{42}$$

[11] The Lagrange Multipliers are indeed differentiable in this case by the implicit function theorem, see for example [10], p45.

and differentiating this with respect to $a$ with $c$ fixed[12] gives

$$\begin{aligned}\frac{\partial E}{\partial a} &= -\frac{1}{x}\frac{\partial x}{\partial a} - r\log y - \frac{ra}{y}\frac{\partial y}{\partial a} \\ &\quad -r^2\log z - \frac{r(r-1)b + ra}{z}\frac{\partial z}{\partial a} \\ &= -r\log y - r^2\log z. \end{aligned} \tag{43}$$

We now derive a good approximation for this last expression. By replacing the variable of summation, $i$, by $r-i$ in the expressions for $Y_1, Y_2$ in (13), (14), (15) we obtain

$$\sum_{j=1,2}\sum_{|r(1-e_j)-i|\le r\delta(r)} \binom{r}{i} xy^i z^{i^2}(yz^r)^{r-2i} = 1 + O(1/r),$$

$$\sum_{j=1,2}\sum_{|r(1-e_j)-i|\le r\delta(r)} \binom{r}{i} \frac{r-i}{r} xy^i z^{i^2}(yz^r)^{r-2i} = a + O(\delta(r)),$$

$$\sum_{j=1,2}\sum_{|r(1-e_j)-i|\le r\delta(r)} \binom{r}{i} \frac{(r-i)(r-i-1)}{r(r-1)} xy^i z^{i^2}(yz^r)^{r-2i} = b + O(\delta(r)).$$

Clearly we have the same equations with $\delta_2(r)$ in place of $\delta(r)$ here so using Lemma 5 (that $e_2$ and $1-e_1$ are within $\delta_2(r)/2$ of each other) and the fact that all terms on the left hand side are positive we obtain

$$\sum_{j=1,2}\sum_{|re_j-i|\le r\delta_2(r)} \binom{r}{i} xy^i z^{i^2}(yz^r)^{r-2i} = 1 + O(1/r) \tag{44}$$

$$\sum_{j=1,2}\sum_{|re_j-i|\le r\delta_2(r)} \binom{r}{i} \frac{r-i}{r} xy^i z^{i^2}(yz^r)^{r-2i} = a + O(\delta_2(r)),$$

$$\sum_{j=1,2}\sum_{|re_j-i|\le r\delta_2(r)} \binom{r}{i} \frac{(r-i)(r-i-1)}{r(r-1)} xy^i z^{i^2}(yz^r)^{r-2i} = b + O(\delta_2(r)).$$

Since the ranges of summation here are of length just $2\delta_2(r)$ and $e_j = 1 - e_{3-j} + O(\delta(r))$ for $j = 1,2$ these last two equations give

$$\sum_{j=1,2}\sum_{|re_j-i|\le r\delta_2(r)} \binom{r}{i} e_{3-j} xy^i z^{i^2}(yz^r)^{r-2i} = a + O(\delta_2(r)), \tag{45}$$

$$\sum_{j=1,2}\sum_{|re_j-i|\le r\delta_2(r)} \binom{r}{i} e_{3-j}^2 xy^i z^{i^2}(yz^r)^{r-2i} = b + O(\delta_2(r)). \tag{46}$$

[12] $E$ is differentiable since the Lagrange Multipliers are.

Let

$$V_1 = \sum_{|re_2 - i| \le r\delta_2(r)} \binom{r}{i} xy^i z^{i^2} (yz^r)^{r-2i}, \qquad V_2 = \sum_{|re_1 - i| \le r\delta_2(r)} \binom{r}{i} xy^i z^{i^2} (yz^r)^{r-2i}.$$

By the intermediate value theorem there are $\theta_1, \theta_2$ with $|\theta_1|, |\theta_2| \le \delta_2(r)$ such that

$$V_1 = (yz^r)^{r-2re_2+r\theta_1} \sum_{|re_2 - i| \le r\delta_2(r)} \binom{r}{i} xy^i z^{i^2}, \tag{47}$$

$$V_2 = (yz^r)^{r-2re_1+r\theta_2} \sum_{|re_1 - i| \le r\delta_2(r)} \binom{r}{i} xy^i z^{i^2}. \tag{48}$$

Substituting $V_1, V_2$ into (44),(45),(46) now gives

$$\begin{aligned} V_1 + V_2 &= 1 + O(\delta_2(r)), \\ e_1 V_1 + e_2 V_2 &= a + O(\delta_2(r)), \\ e_1^2 V_1 + e_2^2 V_2 &= b + O(\delta_2(r)). \end{aligned}$$

But as we have already seen for $Y_1$ and $Y_2$, together with the fact that $|e_2 - (1 - e_1)| \le \delta_2(r)$, these equations determine $V_1$ and $V_2$ within $O(\delta_2(r))$, indeed we must have that

$$V_1 = Y_1 + O(\delta_2(r)), \qquad V_2 = Y_2 + O(\delta_2(r)).$$

But by (47), (48),

$$\begin{aligned} V_1 &= (yz^r)^{r-2re_2+r\theta_1}(Y_2 + O(1/r)), \\ V_2 &= (yz^r)^{r-2re_1+r\theta_2}(Y_1 + O(1/r)). \end{aligned}$$

Without loss of generality we may assume that $Y_2 > 1/4$, in which case the first of these equations gives

$$(yz^r)^{r-2re_2+r\theta_1} = Y_1/Y_2 + O(\delta_2(r)),$$

and hence

$$r(1 - 2e_2 + \theta_1)\log(yz^r) = \log(Y_1/Y_2 + O(\delta_2(r))).$$

Using our assumption that $e_2 - 1/2 \ge \delta_4(r)$ this gives, with (24), (25), (26) (noticing that the interior $O(\delta(r))$ can be extricated under this assumption) that

$$\begin{aligned} r\log(yz^r) &= \frac{\log(Y_1/Y_2)}{1 - 2e_2} + O(\delta_4(r)) \\ &= -\frac{1}{\sqrt{1-4c}} \log\left(\frac{\sqrt{1-4c} - 2a + 1}{\sqrt{1-4c} + 2a - 1}\right) + O(\delta_4(r)). \end{aligned} \tag{49}$$

This is the estimate for $r\log y + r^2 \log z$ that we have been looking for.

Returning to the main proof we now have by (43) that

$$\frac{\partial E}{\partial a} = \frac{1}{\sqrt{1-4c}} \log\left(\frac{\sqrt{1-4c}-2a+1}{\sqrt{1-4c}+2a-1}\right) + O(\delta_4(r)).$$

Hence

$$\begin{aligned} E(a,b,r) &= -\frac{1}{2}\left(1+\frac{1-2a}{\sqrt{1-4c}}\right)\log\left(\frac{1}{2}(1+\frac{1-2a}{\sqrt{1-4c}})\right) + \\ &\quad -\frac{1}{2}\left(1-\frac{1-2a}{\sqrt{1-4c}}\right)\log\left(\frac{1}{2}(1-\frac{1-2a}{\sqrt{1-4c}})\right) + K(c,r) \\ &\quad +O(\delta_4(r)) \end{aligned} \tag{50}$$

for some function $K$. The lemma now follows by combining (50) with Corollary 9.

■

**Corollary 11** *Under the assumption that $b > a^2 + \delta_8(r)$*

$$\begin{aligned} E(a,b,r) &= rT(c) + rG(c,r) + \\ &\quad -\frac{1}{2}(1+\frac{1-2a}{\sqrt{1-4c}})\log\left(\frac{1}{2}(1+\frac{1-2a}{\sqrt{1-4c}})\right) + \\ &\quad -\frac{1}{2}(1-\frac{1-2a}{\sqrt{1-4c}})\log\left(\frac{1}{2}(1-\frac{1-2a}{\sqrt{1-4c}})\right) \end{aligned}$$

*where $G(c,r) \to 0$ as $r \to \infty$.*

**Proof.** It is enough to notice that if any of the conditions given in Lemma 10 fail then so does the requirement that $b > a^2 + \delta_8(r)$. ■

The results proved so far have assumed (8) in order to ensure that the method of Lagrange Multipliers was valid. However in the case that $a = 0, 1$ or $a = b$ Lemma 10 and Corollary 11 still hold, as can be easily seen by direct calculation. As for the lower bounds on $b$, if these were to be taken for $K^{(r)}$ then by the Open Mindedness Principle $K^{(r)}$ could have no other solutions, in which case for suitably large $r' > r$ $K^{(r')}$ would not even be consistent. Since we are ruling out that possibility in what follows we can assume that the lower bounds in (8) are satisfied.

With these results we complete our investigations into the base case. In the next section we shall consider the case of a single predicate $P(x)$.

# 3 The general case for one predicate

Before embarking on the main theorem of this paper we need to introduce a little notation. Recall that in the this case we have propositional variables $Q_1, ..., Q_q$

in addition to the $P(a_i)$. Let $\rho_1, ..., \rho_{2^q}$ denote the atoms $\pm Q_1 \wedge \pm Q_2 \wedge ... \wedge \pm Q_q$. For this section let $K$ now be a general (finite) set of linear constraints involving the $Bel(\rho_j \wedge \pm P(a_1) \wedge \pm P(a_2))$ and let $K^{(r)}$ be as usual the set of constraints obtained by taking the union of all $K$ with $P(a_1)$ and $P(a_2)$ replaced by $P(a_i)$ and $P(a_j)$ respectively, ($i \neq j$, $1 \leq i, j \leq r$). We assume throughout that $K^{(r)}$ is consistent for all $r$.

For $\vec{s} = \langle s_1, ..., s_{2^q} \rangle$, $\vec{a} = \langle a_1, ..., a_{2^q} \rangle$, and $\vec{b} = \langle b_1, ..., b_{2^q} \rangle$ such that $0 \leq s_j$, $\sum_{j=1}^{2^q} s_j = 1$, $0 \leq b_j \leq a_j \leq s_j$ ($j = 1, ..., 2^q$) let[13]

$$\begin{aligned} K(\vec{s}, \vec{a}, \vec{b}, r) \;\; = \;\; & \{ \; Bel(\rho_j \wedge P(a_i) \wedge P(a_k)) = s_j b_j, \; Bel(\rho_j \wedge P(a_i)) = s_j a_j, \\ & Bel(\rho_j) = s_j \mid 1 \leq j \leq 2^q, \; 1 \leq i, k \leq r, \; i \neq k \; \}. \end{aligned} \tag{51}$$

To simplify the notation we shall use $ME(\vec{s}, \vec{a}, \vec{b}, r)$ to denote the maximum entropy solution of $K(\vec{s}, \vec{a}, \vec{b}, r)$ (i.e. $ME(K(\vec{s}, \vec{a}, \vec{b}, r))$), assuming this set of constraints is consistent.

**Lemma 12** *$ME(K^{(r)}) = ME(\vec{s}, \vec{a}, \vec{b}, r)$ for some $\vec{s}, \vec{a}, \vec{b}$ (depending on $r$ and with $K(\vec{s}, \vec{a}, \vec{b}, r)$ consistent).*

**Proof.** Let $\widetilde{K}$ denote the constraints (51) with $s_j, s_j a_j$ and $s_j b_j$ replaced by $ME(K^{(r)})(\rho_j)$, $ME(K^{(r)})(\rho_j \wedge P(a_1))$ and $ME(K^{(r)})(\rho_j \wedge P(a_1) \wedge P(a_2)))$ respectively. By the Principles of Renaming and Obstinacy (see [17]), $ME(K^{(r)}) = ME(K^{(r)} \cup \widetilde{K})$ and by the Equivalence Principle, since any solution of $\widetilde{K}$ satisfies $K^{(r)}$, $ME(K^{(r)} \cup \widetilde{K}) = ME(\widetilde{K})$, as required. ■

Henceforth when referring to $ME(\vec{s}, \vec{a}, \vec{b}, r)$ etc. it will be assumed without mention that $K(\vec{s}, \vec{a}, \vec{b}, r)$ is consistent. In the next lemma $E(\vec{s}, \vec{a}, \vec{b}, r)$ is the entropy of the probability function $ME(\vec{s}, \vec{a}, \vec{b}, r)$ and $\vec{s}\vec{a} = \langle s_1 a_1, s_2 a_2, ..., s_{2^q} a_{2^q} \rangle$ etc..

**Lemma 13**

$$E(\vec{s}, \vec{a}, \vec{b}, r) = r(U(\vec{s}, \vec{s}\vec{a} - \vec{s}\vec{b}) + \delta_{32}(r))$$

*where for $\vec{0} \leq \vec{c} \leq \vec{s}$*

$$U(\vec{s}, \vec{c}) = \sum_{j=1}^{2^q} s_j T(c_j / s_j).$$

**Proof.** Let $\alpha_1, ..., \alpha_{2^r}$ denote the atoms $\pm P(a_1) \wedge ... \wedge \pm P(a_r)$. By the Relativization Principle,

$$ME(\vec{s}, \vec{a}, \vec{b}, r)(\rho_j \wedge \alpha_i) = s_j ME(K_j^{(r)})(\alpha_i),$$

[13]Hopefully our use of subscripted $a$'s for both the constants of the language and values appearing in the constraints will not cause any confusion. It should be clear from the context which is intended in any instance.

where

$$K_j = \{Bel(P(a_1)) = Bel(P(a_2)) = a_j,\ Bel(P(a_1) \wedge P(a_2)) = b_j\}.$$

Consequently, by Corollary 9

$$\begin{aligned} E(\vec{s}, \vec{a}, \vec{b}, r) &= -\sum_{j=1}^{2^q} \sum_{i=1}^{2^r} \left(s_j ME(K_j^{(r)})(\alpha_i\right) \log \left(s_j ME(K_j^{(r)})(\alpha_i)\right) \\ &= -\sum_{j=1}^{2^q} s_j \log s_j + \sum_{j=1}^{2^q} s_j E(a_j, b_j, r) \\ &= r\left(\sum_{j=1}^{2^q} s_j T(a_j - b_j) + O(\delta_{32}(r))\right). \end{aligned}$$

Putting

$$U(\vec{s}, \vec{s}\vec{a} - \vec{s}\vec{b}) = \sum_{j=1}^{2^q} s_j T(a_j - b_j)$$

the lemma follows. ■

**Lemma 14** *$U(\vec{s}, \vec{c})$ is strictly concave for $\vec{0} \leq \vec{c} \leq \vec{s}, \vec{\frac{1}{4}}$.*

**Proof.** To show this first notice that $T$ is strictly concave, as can be seen by taking second derivatives. Hence for $0 < t < 1$,

$$\begin{aligned} &U(t\vec{s} + (1-t)\vec{s'}, t(\vec{s}\vec{a} - \vec{s}\vec{b}) + (1-t)(\vec{s'}\vec{a'} - \vec{s'}\vec{b'})) = \\ &= \sum_{j=1}^{2^q} (ts_j + (1-t)s'_j) T\left(\frac{t(s_j a_j - s_j b_j) + (1-t)(s'_j a'_j - s'_j b'_j)}{ts_j + (1-t)s'_j}\right) \\ &> \sum_{j=1}^{2^q} (ts_j + (1-t)s'_j)\left(\frac{ts_j}{ts_j + (1-t)s'_j} T(a_j - b_j) + \frac{(1-t)s'_j}{ts_j + (1-t)s'_j} T(a'_j - b'_j)\right) \\ &= tU(\vec{s}, \vec{s}\vec{a} - \vec{s}\vec{b}) + (1-t)U(\vec{s'}, \vec{s'}\vec{a'} - \vec{s'}\vec{b'}), \end{aligned}$$

as required. ■

Let $H^{(r)}$ be the set of $\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b} \rangle$ such that $ME(\vec{s}, \vec{a}, \vec{b}, r)$ is a solution of $K^{(r)}$, or equivalently such that

$$\begin{aligned} \{\ &Bel(\rho_j \wedge P(a_i)) = s_j a_j,\ Bel(\rho_j \wedge P(a_i) \wedge P(a_k)) = s_j b_j \\ &Bel(\rho_j) = s_j \mid 1 \leq j \leq 2^q,\ 1 \leq i, k \leq r,\ i \neq k\ \} \end{aligned} \tag{52}$$

is consistent with $K^{(r)}$. Let $H = \bigcap_{r \geq 2} H^{(r)}$ and let $K_j$ be as in the proof of Lemma 13.

**Lemma 15** *$\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle$ is in $H^{(r)}$ iff $\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle \in H^{(2)}$ and for each $j = 1, 2, ..., 2^q$ with $s_j \neq 0$, $K_j^{(r)}$ is consistent.*

**Proof.** If $\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle$ is in $H^{(r)}$ and $s_j \neq 0$ then the conditional probability function $ME(\vec{s}, \vec{a}, \vec{b}, r)(-|\rho_j)$ satisfies $K_j^{(r)}$. Conversely if $\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle \in H^{(2)}$ and the probability function $Bel_j$ on the language with propositional variables $P(a_1), ..., P(a_r)$ satisfies $K_j^{(r)}$ whenever $s_j \neq 0$ then the probability function $Bel$ on $L^{(r)}$ defined by

$$Bel(\rho_j \wedge \bigwedge_{i=1}^{r} P^{\epsilon_i}(a_i)) = s_j Bel_j(\bigwedge_{i=1}^{r} P^{\epsilon_i}(a_i))$$

satisfies $K(\vec{s}, \vec{a}, \vec{b}, r)$. ■

**Lemma 16** *$\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle$ is in $H$ iff*
$\{\ Bel(\rho_j) = s_j,\ \ Bel(\rho_j \wedge P(a_1)) = Bel(\rho_j \wedge P(a_2)) = s_j a_j,$
$Bel(\rho_j \wedge P(a_1) \wedge P(a_2)) = s_j b_j\ \}$
*is consistent with $K^{(2)}$, and $a_j^2 \leq b_j \leq a_j$ for $j = 1, 2, ..., 2^q$.*

**Proof.** The result follows by Lemmas 2 and 15. ■

**Lemma 17** *$H$ is convex.*

**Proof.** This follows since the constraints in $K(\vec{s}, \vec{a}, \vec{b}, r)$ are linear in $s_1, s_1a_1, s_1b_1$, ..., $s_{2^q}, s_{2^q}a_{2^q}, s_{2^q}b_{2^q}$ so if $0 < t < 1$ and $Bel, Bel'$ are solutions of $K(\vec{s}, \vec{a}, \vec{b}, r)$ and $K(\vec{s'}, \vec{a'}, \vec{b'}, r)$ respectively then $tBel + (1-t)Bel'$ is a solution of $K(\vec{s''}, \vec{a''}, \vec{b''}, r)$ where

$$\begin{aligned}
\vec{s''} &= t\vec{s} + (1-t)\vec{s'}, \\
\vec{s''}\vec{a''} &= t\vec{s}\vec{a} + (1-t)\vec{s'}\vec{a'}, \\
\vec{s''}\vec{b''} &= t\vec{s}\vec{b} + (1-t)\vec{s'}\vec{b'}.
\end{aligned}$$

In other words, if $\langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle, \langle \vec{s'}, \vec{s'}\vec{a'}, \vec{s'}\vec{b'}\rangle \in H$ then

$$\langle t\vec{s} + (1-t)\vec{s'}, t\vec{s}\vec{a} + (1-t)\vec{s'}\vec{a'}, t\vec{s}\vec{b} + (1-t)\vec{s'}\vec{b'}\rangle \in H.$$

■

Now for $j = 1, 2, ..., 2^q$ let

$$\begin{aligned}
s_j(r) &= ME(K^{(r)})(\rho_j), \\
s_j(r)a_j(r) &= ME(K^{(r)})(\rho_j \wedge P(a_1)), \\
s_j(r)b_j(r) &= ME(K^{(r)})(\rho_j \wedge P(a_1) \wedge P(a_2)).
\end{aligned}$$

[If $s_j(r) = 0$ we may, arbitrarily, give $a_j(r), b_j(r)$ value 0 too.]

The whole of the remaining proof hinges essentially on showing that this sequence has a single limit.

**Lemma 18** *Any convergent subsequence $\langle \vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k), \vec{s}(r_k)\vec{b}(r_k)\rangle$ of $\langle \vec{s}(r), \vec{s}(r)\vec{a}(r), \vec{s}(r)\vec{b}(r)\rangle$ converges to a point in $H$.*

**Proof.** Let $m$ be a fixed natural number. By compactness there is a subsequence

$$ME(\vec{s}(r_{k'}), \vec{s}(r_{k'})\vec{a}(r_{k'}), \vec{s}(r_{k'})\vec{b}(r_{k'}), r_{k'})$$

of the sequence

$$ME(\vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k), \vec{s}(r_k)\vec{b}(r_k), r_k)$$

for which the sequences

$$ME(\vec{s}(r_{k'}), \vec{s}(r_{k'})\vec{a}(r_{k'}), \vec{s}(r_{k'})\vec{b}(r_{k'}), r_{k'})(\rho_j \wedge \pm P(a_1) \wedge ... \wedge \pm P(a_m)) \qquad (53)$$

converge for any fixed $\rho_j \wedge \pm P(a_1) \wedge ... \wedge \pm P(a_m)$. Furthermore, since all the $ME(\vec{s}(r_{k'}), \vec{s}(r_{k'})\vec{a}(r_{k'}), \vec{s}(r_{k'})\vec{b}(r_{k'}), r_{k'})$ for $r_{k'} \geq m$ satisfy $K^{(m)} (\subseteq K^{(r_{k'})})$, and the set of solutions to $K^{(m)}$ is closed, the limit values of (53) also satisfy $K^{(m)}$. Since $\langle \vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k), \vec{s}(r_k)\vec{b}(r_k)\rangle$ and $\langle \vec{s}(r_{k'}), \vec{s}(r_{k'})\vec{a}(r_{k'}), \vec{s}(r_{k'})\vec{b}(r_{k'})\rangle$ obviously have the same limit point the result now follows. ■

By Lemmas 17 and 18 $H$ is closed convex. It follows that the set

$$H_* = \{\langle \vec{s}, \vec{s}\vec{a} - \vec{s}\vec{b}\rangle \mid \langle \vec{s}, \vec{s}\vec{a}, \vec{s}\vec{b}\rangle \in H\}$$

is also closed convex. Hence, by Lemma 13, the function $U(\vec{s}, \vec{s}\vec{c})$ has a unique maximum point, $\langle \vec{s}_*, \vec{s}_*\vec{c}_*\rangle$ say, on $H_*$.

**Lemma 19** *If $\langle \vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k), \vec{s}(r_k)\vec{b}(r_k)\rangle$ is a convergent subsequence of $\langle \vec{s}(r), \vec{s}(r)\vec{a}(r), \vec{s}(r)\vec{b}(r)\rangle$ then $\langle \vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k) - \vec{s}(r_k)\vec{b}(r_k)\rangle$ converges to $\langle \vec{s}_*, \vec{s}_*\vec{c}_*\rangle$.*

**Proof.** Suppose on the contrary the limit of $\langle \vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k), \vec{s}(r_k)\vec{b}(r_k)\rangle$ is $\langle \vec{s}_\circ, \vec{s}_\circ\vec{a}_\circ, \vec{s}_\circ\vec{b}_\circ\rangle$ and $\langle \vec{s}_\circ, \vec{s}_\circ\vec{a}_\circ - \vec{s}_\circ\vec{b}_\circ\rangle \neq \langle \vec{s}_*, \vec{s}_*\vec{c}_*\rangle$. Let $\langle \vec{s}_*, \vec{s}_*\vec{a}_*, \vec{s}_*\vec{b}_*\rangle \in H$ be such that $\vec{s}_*\vec{a}_* - \vec{s}_*\vec{b}_* = \vec{s}_*\vec{c}_*$. [In general there will be many possible choices here for $\vec{a}^*, \vec{b}^*$.]

Then

$$U(\vec{s}_*, \vec{s}_*\vec{a}_* - \vec{s}_*\vec{b}_*) > U(\vec{s}_\circ, \vec{s}_\circ\vec{a}_\circ - \vec{s}_\circ\vec{b}_\circ) + \eta$$

for some fixed $\eta > 0$. Hence, for large $k$

$$U(\vec{s}_*, \vec{s}_*\vec{a}_* - \vec{s}_*\vec{b}_*) > U(\vec{s}(r_k), \vec{s}(r_k)\vec{a}(r_k) - \vec{s}(r_k)\vec{b}(r_k)) + \eta/2.$$

But then, by Lemma 13,

$$E(\vec{s}_*, \vec{a}_*, \vec{b}_*) > E(\vec{s}(r_k), \vec{a}(r_k), \vec{b}(r_k)),$$

for large $k$, which is clearly a contradiction since $ME(\vec{s}_*, \vec{a}_*, \vec{b}_*, r)$ is a solution of $K^{(r)}$ for every $r$, so certainly for the subsequence of $r_k$. ■

Our next lemma effectively completes the proof of the Main Theorem in the case of a single predicate $P(x)$.

**Lemma 20** *The sequence $\langle \vec{s}(r), \vec{s}(r)\vec{a}(r), \vec{s}(r)\vec{b}(r)\rangle$ converges.*

**Proof.** Let $\langle \vec{s}(r_k^+), \vec{s}(r_k^+)\vec{a}(r_k^+), \vec{s}(r_k^+)\vec{b}(r_k^+)\rangle$ be some subsequence of $\langle \vec{s}(r), \vec{s}(r)\vec{a}(r), \vec{s}(r)\vec{b}(r)\rangle$ converging to $\langle \vec{s}^+, \vec{s}^+\vec{a}^+, \vec{s}^+\vec{b}^+\rangle$. Since, by Lemma 18, $\langle \vec{s}^+, \vec{s}^+\vec{a}^+, \vec{s}^+\vec{b}^+\rangle \in H \subseteq H^{(r_k^+)}$, it follows by the Open Mindedness Principle that $s(r_k^+)_j$ can only be zero if also $s_j^+ = 0$.

Set $\vec{c} = \vec{a}^+ - \vec{b}^+$ and let $\langle \vec{s}_*, \vec{s}_*\vec{a}_*, \vec{s}_*\vec{b}_*\rangle$ be that (now unique) point in $H$ for which $\vec{s}_* = \vec{s}^+$, $\vec{s}_*\vec{a}_* - \vec{s}_*\vec{b}_* = \vec{s}^+\vec{a}^+ - \vec{s}^+\vec{b}^+ = \vec{s}^+\vec{c}$ and the concave function function

$$-\sum_{j=1}^{2^q}\Bigg[s_j \log s_j + \frac{s_j}{2}\left(1+\frac{1-2a_j}{\sqrt{1-4c_j}}\right)\log\frac{1}{2}\left(1+\frac{1-2a_j}{\sqrt{1-4c_j}}\right) + \frac{s_j}{2}\left(1-\frac{1-2a_j}{\sqrt{1-4c_j}}\right)\log\frac{1}{2}\left(1-\frac{1-2a_j}{\sqrt{1-4c_j}}\right)\Bigg] \tag{54}$$

is maximal.

We now show that the vector

$$\langle \vec{s}(r_k^+), \vec{s}(r_k^+)\vec{a}(r_k^+) + \vec{s}_*\vec{a}_* - \vec{s}^+\vec{a}^+, \vec{s}(r_k^+)\vec{b}(r_k^+) + \vec{s}_*\vec{b}_* - \vec{s}^+\vec{b}^+\rangle \tag{55}$$

is in $H^{(r_k^+)}$, that is that it satisfies the conditions specified in Lemma 15, for large $k$ and each $j = 1, 2, ..., 2^q$. Clearly if $s_{*j} = 0$ then we have what we want so assume $s_{*j} \neq 0$. By Lemma 16, $a_{*j}^2 \leq b_{*j} \leq a_{*j}$. If $a_{*j}^2 < b_{*j} < a_{*j}$ then the same inequalities will hold for (55) for large $k$ since this vector converges to $\langle \vec{s}_*, \vec{s}_*\vec{a}_*, \vec{s}_*\vec{b}_*\rangle$ in the limit.

On the other hand, if $a_{*j}^2 = b_{*j}$ then we must also have that $a_j^{+2} = b_j^+$ and $a_j^+ \neq 1 - a_{*j}$ (unless $a_{*j} = 1/2$). The reason for this is that otherwise, as in the proof of the Open Mindedness Principle given in [17], for small $\epsilon > 0$ the value of the function (54) on $\langle \vec{s}_*, \vec{s}_*((1-\epsilon)\vec{a}_* + \epsilon\vec{a^+}), \vec{s}_*((1-\epsilon)\vec{b}_* + \epsilon\vec{b^+})\rangle$ will exceed its value on $\langle \vec{s}_*, \vec{s}_*\vec{a}_*, \vec{s}_*\vec{b}_*\rangle$, which is a contradiction by Lemma 17. Furthermore since

$\vec{a_*} - \vec{b_*} = \vec{a^+} - \vec{b^+}$ this forces that $a_{*j} = a_j^+$. Consequently the $j$th requirement (according to Lemma 15) on (55) just reduces to the requirement that

$$\frac{[r_k^+ a(r_k^+)_j](2r_k^+ a(r_k^+)_j - 1 - [r_k^+ a(r_k^+)_j])}{r_k^+ - 1} \leq b(r_k^+)_j \leq a(r_k^+)_j$$

which of course does hold since $\langle \vec{s}(r_k^+), \vec{s}(r_k^+)\vec{a}(r_k^+), \vec{s}(r_k^+)\vec{b}(r_k^+)\rangle \in H^{(r_k^+)}$.

Finally if $a_{*j} = b_{*j}$ then it must have been the case that also $a_j^+ = b_j^+$ (since the $c$ values are the same) so since by Lemma 15 $b(r_k^+)_j \leq a(r_k^+)_j$ the corresponding $j$th requirement will also hold for (55).

Having shown that the vector (55) is in $H^{(r_k^+)}$ we claim that for large $k$ the entropy associated with this vector will exceed that for

$$\langle \vec{s}(r_k^+), \vec{s}(r_k^+)\vec{a}(r_k^+), \vec{s}(r_k^+)\vec{b}(r_k^+)\rangle, \tag{56}$$

thus providing the required contradiction.

First notice that (55) can be rewritten as

$$\langle \vec{s}(r_k^+), \vec{s}(r_k^+)(\vec{a}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+)), \vec{s}(r_k^+)(\vec{b}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{b}_* - \vec{b}^+))\rangle$$

(where we take $s_{*j}/s_j(r_k^+)$ to be 0 if $s_{*j} = 0$) and (55) and (56) will have the same $\vec{s}$ and $\vec{c}$ so any difference in their entropies will be as a result of the contributions due to their $\vec{a}$. Namely, by Corollary 11,

$$E(\vec{s}(r_k^+), \vec{a}(r_k^+) + \tfrac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+), \vec{b}(r_k^+) + \tfrac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{b}_* - \vec{b}^+)) - E(\vec{s}(r_k^+), \vec{a}(r_k^+), \vec{b}(r_k^+))$$

$$= -\sum_{j=1}^{2^q} \frac{s_j(r_k^+)}{2}\{$$

$$\left(1 + \frac{1 - 2(\vec{a}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+))}{\sqrt{1 - 4c_j(r_k^+)}}\right) \log\left(\frac{1}{2} + \frac{1 - 2(\vec{a}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+))}{2\sqrt{1 - 4c_j(r_k^+)}}\right)$$

$$+ \left(1 - \frac{1 - 2(\vec{a}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+))}{\sqrt{1 - 4c_j(r_k^+)}}\right) \log\left(\frac{1}{2} - \frac{1 - 2(\vec{a}(r_k^+) + \frac{\vec{s}_*}{\vec{s}(r_k^+)}(\vec{a}_* - \vec{a}^+))}{2\sqrt{1 - 4c_j(r_k^+)}}\right)$$

$$- \left(1 + \frac{1 - 2a_j(r_k^+)}{\sqrt{1 - 4c_j(r_k^+)}}\right) \log\frac{1}{2}\left(1 + \frac{1 - 2a_j(r_k^+)}{\sqrt{1 - 4c_j(r_k^+)}}\right)$$
$$- \left(1 - \frac{1 - 2a_j(r_k^+)}{\sqrt{1 - 4c_j(r_k^+)}}\right) \log\frac{1}{2}\left(1 - \frac{1 - 2a_j(r_k^+)}{\sqrt{1 - 4c_j(r_k^+)}}\right)\Bigg\}. \tag{57}$$

Of course this expression *appears* to assume that the conditions of Corollary 11, namely that $b-a^2 > \delta_8(r)$, holds for all coordinates of (55) and (56). However if this does not hold at a coordinate for sufficiently large $r_k^+$ then, as we have already seen, it must be that the corresponding $a_{*j} = a_j^+$, $b_{*j} = b_j^+$ so this $j$th term disappears from the sum in any case!

Focusing then on the coordinates that matter note that for these the sum (54) with $a_j = (1-t)a_{*j} + ta_j^+$ *decreases* as we move up from $t = 0$. Fix small $t$, so that Corollary 11 *is valid* for these $a_j = (1-t)a_{*j} + ta_j^+$ (and the running $c_j$) for large $r$ on the coordinates for which $a_{*j}^2 < b_{*j} < a_{*j}$. Then for large $k$ the value of the function (54) with the sum restricted to these coordinates on

$$\langle \vec{s}(r_k^+), \vec{s}(r_k^+)\vec{a}(r_k^+) + (1-t)\vec{s}_*(\vec{a}_* - \vec{a}^+), \vec{s}(r_k^+)\vec{b}(r_k^+) + (1-t)\vec{s}_*(\vec{b}_* - \vec{b}^+)\rangle, \quad (58)$$

is less than that on (55), since it is in the limit. However since, with this restriction, the function (54) is concave its value at (56) will be yet smaller, contradiction!! ■

This concludes the proof of the main theorem of this paper, apart from showing that this limit solution guaranteed by Lemma 20 corresponds to a complete set of reasons. But this is now clear since, in the notation of that Lemma,

$$\begin{aligned}
&\lim_{r\to\infty} ME(K^{(r)})(P^{\epsilon_1}(a_1) \wedge P^{\epsilon_2}(a_2) \wedge \ldots \wedge P^{\epsilon_m}(a_m)) \\
&= \lim_{r\to\infty} \sum_{j=1}^{2^q} s_j(r) ME(K_j(r))(\bigwedge_{i=1}^{m} P^{\epsilon_i}(a_i)) \\
&= \lim_{r\to\infty} \sum_{j=1}^{2^q} s_{j*} ME(K_{*j}(r))(\bigwedge_{i=1}^{m} P^{\epsilon_i}(a_i)) \\
&= \sum_{j=1}^{2^q} s_{*j} \lim_{r\to\infty} ME(K_{*j}(r))(\bigwedge_{i=1}^{m} P^{\epsilon_i}(a_i)),
\end{aligned}$$

where

$$K_j(r) = \{Bel(P(a_i)) = a_j(r),\ Bel(P(a_i) \wedge P(a_k)) = b_j(r) \mid 1 \le i,k \le r,\ i \ne k\},$$

$$K_{*j}(r) = \{Bel(P(a_i)) = a_{*j},\ Bel(P(a_i) \wedge P(a_k)) = b_{*j} \mid 1 \le i,k \le r,\ i \ne k\},$$

and these final summands are of the required form by the base case.

# 4 Conclusion

In this note we have, by a somewhat torturous route, proved the $s = 2$ induction theorem for a single predicate. There are two obvious next steps. Prove this theorem for multiple predicates as was done in [21] for the $s = 1$ case, and extend this result to the $s > 2$ cases.

The former of these, despite the initial feeling that what it amounts to is but a trivial extension of the one predicate case, would nevertheless seem to present formidable technical difficulties, at least if one is to prove it by directly generalizing the results presented in this paper. Notwithstanding however we have no hesitation in conjecturing its truth, and indeed its truth for $s > 2$ too.

As for the single predicate, or base cases, for $s > 2$, the methods of this paper would seem to be somewhat more easily generalizable to those, well certainly the base cases, and we would see that as the natural next step.

# References


[1] Bacchus, F., Grove, A., Halpern, J.Y. & Koller, D., From statistics to beliefs. In *Proc. of National Conference on Artificial Intelligence (AAAI-92)*, 1992, pp 602-608. http://citeseer.nj.nec.com/article/bacchus92from.html.

[2] Barnett, O.W. & Paris, J.B., A proof of the base cases of Emergence of Reason's Conjecture for $CM^\infty$, *Technical Report of the Manchester Centre for Pure Mathematics*, University of Manchester, UK., *to appear.*

[3] Carnap, R., *Logical Foundations of Probability,* Routledge & Kegan Paul Ltd., 1951.

[4] Carnap, R., *The Continuum of Inductive Methods,* University of Chicago Press, 1952.

[5] Carnap, R. & Jeffrey, R.C. (eds.), *Studies in Inductive Logic and Probability*, University of California Press, 1971.

[6] Carnap, R. & Stegmüller, W., *Induktive Logik und Wahrscheinlichkeit*, Springer-Verlag, Vienna, 1958, pp 243-249.

[7] Fagin, R., Probabilities on finite models, *Journal of Symbolic Logic,* **41**:1, 1976, pp 50-58.

[8] Gaifman, H., Concerning measures in first order calculi, *Israel Journal of Mathematics,* **24**, 1964, pp 1-18.

[9] Glebskij, Y.V., Kogan, D.I., Liogon'kij, M.I. & Talanov, V.A., Range and degree of realizability of formulas in the restricted predicate calculus, *Kibernetika*, **2**, 1969, pp 17-28.

[10] Goursat, E., *A Course in Mathematical Analysis, Vol.I*, Ginn & Company, Boston, 1904.

[11] Grove, A.J., Halpern, J.Y. & Koller, K., Random worlds and maximum entropy, *Journal of Artificial Intelligence Research*, **2**, 1994, pp 33-88.

[12] Halpern, Y.H., An analysis of first-order logics of probability, *Artificial Intelligence*, **46**, 1990, pp 311-350.

[13] Jeffrey, R.C. (ed.), *Studies in Inductive Logic and Probability (Volume II)*, University of California Press, 1980.

[14] Johnson, W.E., Probability: The deductive and inductive problems, *Mind,* **49**, 1932, pp 409-423.

[15] Kemeny, J.G., A contribution to inductive logic, *Philos. Phenomenological Res.* **13**, 1953, pp 371-374.

[16] Nix, C.J. & Paris, J.B., A continuum of inductive methods arising from a generalized principle of instantial relevance, submitted to the *Journal of Philosophical Logic.*

[17] Paris, J.B., *The Uncertain Reasoner's Companion: A Mathematical Perspective*, Cambridge University Press, 1994.

[18] Paris, J.B. & Vencovská, A., A note on the inevitability of maximum entropy, *International Journal of Approximate Reasoning,* **4**(3), 1990, pp 183-224.

[19] Paris, J.B., Common sense and maximum entropy, *Synthese,* **117**, 1999, pp 75-93.

[20] Paris, J.B. & Wafy, M., On the emergence of reasons, *Journal of the IGPL*, **9(2)**, 2001, pp207-216. [In electronic form at http://www3.oup.co.uk/igpl/Volume_09/Issue_02/#Paris.]

[21] Paris, J.B., Vencovská, A. & Wafy, M., Some limit theorems for $ME$, $CM^\infty$ and $MD$, *Technical Report of the Manchester Centre for Pure Mathematics*, University of Manchester, UK., no. 2001/9, ISSN 1472-9210, 2001.

[22] Paris, J.B. & Vencovská, A., Common sense and stochastic independence. *Foundations of Bayesianism*, eds. D.Corfield & J.Williamson, Kluwer Academic Press, 2002, pp 203-240.

[23] Paris, J.B. & Vencovská, A., There is a reason for everything (probably): On the application of maxent to induction, in *Proceedings of the Conditionals, Information and Inference Workshop*, eds. G.Kern-Isberner & W.Rodder, FernUniverstät, Hagen, Germany, 2002. pp 211-229. [To appear in the Springer Lecture Notes in Computer Science series.]

[24] Popper, K. & Miller, D.W., A proof of the impossibility of inductive probability, *Nature,* **302**, 1983, pp 687-688.

[25] Wafy, M., *A study of an inductive problem using inference processes*, Ph.D. Thesis, Manchester University, 2000.