An Algorithm for Computing the Maximum Entropy Ranking for Variable Strength Defaults

Rachel A Bourne and Simon Parsons
Department of Electronic Engineering
Queen Mary & Westfield College
University of London
London E1 4NS, UK
r.a.bourne,s.d.parsons@elec.qmw.ac.uk

Abstract

A new algorithm for computing the maximum entropy ranking (me-ranking) over a set of variable strength defaults is given. Although this requires more information than its predecessor in terms of strengths for defaults, it has a much wider applicability, and allows for greater expressiveness in encoding default knowledge. The algorithm is shown to be sound and complete to the extent that it always produces an me-ranking when one exists. The reasons for multiple rankings and for no ranking are explored. A sufficient condition for uniqueness of the me-ranking is given.

1 Background

Defaults like $bird \Rightarrow fly$ are natural rules which can be used to describe the normal behaviour of some domain. Default knowledge can be used to make inferences about a given situation so that, for example, an arbitrary bird would be assumed to fly unless there were information to the contrary.

The question arises of what *should* be inferred from any given set of defaults and research into this has led to the development of many very different nonmonotonic reasoning formalisms [8, 12, 13, 14]. A probabilistic model of default reasoning has proven to capture the most basic requirements of nonmonotonic behaviour [1, 7, 11]. A natural extension to this model is to apply the principle of maximum entropy to find a distinguished consequence relation which captures the default knowledge but makes no additional assumptions, i.e., is the least biased [6, 10]. In this paper, the original work of [5] on applying maximum entropy to default reasoning is extended to cater for arbitrary sets of variable strength defaults. In doing so, it is hoped both to clarify the assumptions required by the me-approach, and to demonstrate that this flexible method can help to explain why there has been disagreement among researchers regarding the nature of default inference.

2 Deriving the me-ranking

First, it is necessary to look at what a default represents and how defaults are connected to ranking functions and consequence relations.

Definition 2.1 *A* default *is a natural rule of the form $a \Rightarrow b$ such that the antecedent, a, and the consequent, b, are formulæ of a finite propositional language, $\mathcal{L}$.*

The symbol, $\Rightarrow$, is a new default connective the semantics of which will be described shortly; it should not be confused with the material implication symbol, $\rightarrow$. The connectives $\wedge$, $\vee$, $\neg$, $\rightarrow$, have their usual meaning. The semantics of $\mathcal{L}$ are given in terms of the set of its models, $\mathcal{M}$. A model, $m \in \mathcal{M}$, is said to *verify* a default, $a \Rightarrow b$, if $m \models a \wedge b$. Conversely, a model, m, is said to *falsify* a default, $a \Rightarrow b$, if $m \models a \wedge \neg b$.

Definition 2.2 *A* ranking function*, κ, is a mapping from $\mathcal{M}$ to the nonnegative integers for which at least one model, m, has $\kappa(m) = 0$. This determines a preference ordering models, so that*

$$\kappa(m) < \kappa(m')$$

means that m is preferred to, or more normal than, m'.

This function, κ, in turn determines a preference ordering over the formulas of $\mathcal{L}$, where a formula is as preferred as its most preferred model, so that

$$\kappa(a) = \min_{m \models a}[\kappa(m)] \tag{1}$$

Equivalently, $\kappa(a) < \kappa(b)$ means that there exists an m such that $m \models a$ and for all m' such that $m' \models b$, $\kappa(m) < \kappa(m')$.

A default constrains a ranking function so that it is more normal to verify the default than to falsify it. Such a ranking function is said to be admissible with respect to that default. More formally,

Definition 2.3 *A ranking function, κ,* satisfies *a default, $a \Rightarrow b$, or is* admissible *with respect to it, iff*

$$\kappa(a \wedge b) < \kappa(a \wedge \neg b) \tag{2}$$

A ranking function over models is equivalent to a rational consequence relation [8]. To determine whether some consequent, b, is a consequence of some antecedent, a, with respect to a ranking function, κ, it is necessary to check whether κ satisfies $a \Rightarrow b$. If $\mid\!\sim_\kappa$ represents the consequence relation, then

$$a \mid\!\sim_\kappa b \quad \text{iff} \quad \kappa(a \wedge b) < \kappa(a \wedge \neg b)$$

Defaults have also been given a semantics in terms of parameterised probability distributions [1, 5], and ranking functions can be viewed as abstractions of these. The parameterised probability semantics defines a default, $a \Rightarrow b$, as a constraint on the conditional probability to be 'almost' certain, $P(b|a) \geq 1 - \varepsilon$, where ε is a parameter close to zero. Now, if constraints are expressed in terms of powers of ε, the corresponding defaults can be viewed as having priorities, or strengths, which correspond to the exponent of ε. This gives a revised notion of satisfaction and admissibility so that

$$a \stackrel{s}{\Rightarrow} b \quad \text{means that} \quad \kappa(a \wedge b) + s \leq \kappa(a \wedge \neg b) \tag{3}$$

The default represents the constraint $P(b|a) \geq 1 - \varepsilon^s$.

The equivalence of the probabilistic and ranking function representations for defaults is useful. It is well known that when a set of constraints determines some class of probability distributions, there exists a (usually) unique member which corresponds to the least biased estimate, or that which makes the least assumptions about the rest of the data—the maximum entropy distribution [6]. By using the parameterised probabilistic semantics, the principle of maximum entropy can be applied to find a probability distribution. Once this has been found, it can be abstracted back to the ranking function representation.

The entropy of a probability distribution over a set of models, $\mathcal{M}$, is given by

$$H[P] = -\sum_{m \in \mathcal{M}} P(m) \log P(m) \tag{4}$$

The problem, then, is to select that probability distribution which maximises (4) subject to constraints imposed by the defaults.

This idea was originally proposed by Pearl [11], and developed by Goldszmidt *et al.* [5], who presented an algorithm which computed the me-ranking for a restricted class of defaults called minimal core sets. They used the same constraint for each default but pointed out that the algorithm could also be applied to variable strength defaults, provided the sets were still minimal core. The minimal core restriction ensures that all the inequality constraints are satisfied as equalities.

In the work presented here, using variable strength defaults requires all constraints to be satisfied as equalities in the maximum entropy distribution. While this means that more information is required from the knowledge engineer, who must now specify the strength he assigns to defaults explicitly, it makes the whole formalism more meaningful since it guarantees that all defaults are taken into account and stresses the necessity of specifying relative strengths for defaults, something which was not obvious in the original work. The point is that varying the strengths leads to different me-rankings so that it is not meaningful to look for *the* me-ranking without specifying a strength assignment over the defaults.

Now, specifying relative orders of magnitude for the conditional probabilities corresponding to each default (i.e., their strengths), results in a similar order of magnitude description of the probabilities of each model. This is achieved by allowing the parameter, ε, to tend to zero. This can be thought of as taking a set of assumptions (i.e., the defaults) to the extreme in order to ascertain what other information is implied. Intuitively, defaults with numerically higher strengths can be thought of as holding more strongly than, or as having priority over, those of lower strength.

Because the analysis is infinitesimal, the constraints represent asymptotic equalities and the symbol $\sim$ is used to denote this[1]. When abstracting back to ranking functions, it is only the lowest exponents in an expression which dominate and the asymptotic equalities become integer equalities. Note that since it is only the asymptotic behaviour of the probabilities that is important, there is no need to consider their coefficients explicity, nor indeed the actual value of entropy.

The strength of each default is expressed as some power of the parameter ε which has no significance other than linking all defaults together. Thus a default $a \overset{s}{\Rightarrow} b$ will be said to have relative strength s, if $P(\neg b|a) \sim \varepsilon^s$ for some integer $s > 0$. As $\varepsilon \to 0$, the term $\varepsilon^s \to 0$ and so $P(b|a) \to 1$; the default becomes arbitrarily certain. In specifying a default, it is assumed that the knowledge engineer is encoding information which he takes to be almost certain.

In a similar manner, the probability of each model $m \in \mathcal{M}$ is taken to be asymptotically equivalent to some non-negative integer power of ε, i.e., $P(m) \sim \varepsilon^{\kappa(m)}$ for $\kappa(m) \geq 0$; this exponent is denoted $\kappa(m)$ since, later, it will determine the ranking function over $\mathcal{M}$.

Given a set of variable strength defaults, $\Delta = \{r_i : a_i \overset{s_i}{\Rightarrow} b_i\}$, the constraints imposed on P for each default can be written:

$$\sum_{m \models a_i \wedge \neg b_i} P(m) \quad \sim \quad \frac{\varepsilon^{s_i}}{1 - \varepsilon^{s_i}} \sum_{m \models a_i \wedge b_i} P(m) \tag{5}$$

Using these constraints and the Lagrange multiplier technique to find the point of maximum entropy, Goldszmidt *et al.* [5] derived the following elegant and simple

[1]Two models m and m' are asymptotically equivalent iff $\lim_{\varepsilon \to 0} \frac{P(m)}{P(m')} = C$, a constant.

approximation for the probability of each model[2]:

$$P(m) \quad \sim \quad \prod_{\frac{r_i}{m \models a_i \wedge \neg b_i}} \alpha_i \tag{6}$$

where the α_i relate to the Lagrange multipliers for each default.

Making a further assumption that the α_i can also be approximated by a relative order of magnitude, thus writing $\alpha_i \sim \varepsilon^{\kappa(r_i)}$, the probability expressions (6) are substituted back into the constraints (5) yielding $|\Delta|$ simultaneous equations with $|\Delta|$ unknowns, the $\kappa(r_i)$[3].

In the limit as $\varepsilon \to 0$ those models with the lowest powers of ε will dominate, and the constraints reduce to:

$$\min_{m \models a_i \wedge b_i} [\kappa(m)] + s_i \quad = \min_{m \models a_i \wedge \neg b_i} [\kappa(m)] \tag{7}$$

Given a set of me-ranks for defaults, $\kappa(r_i)$, the me-ranking over models, $\kappa(m)$, can be found using the abstraction of equation (6). The me-rank of each model is given by the sum of the me-ranks of those defaults it falsifies:

$$\kappa(m) = \sum_{\frac{r_i}{m \models a_i \wedge \neg b_i}} \kappa(r_i) \tag{8}$$

This completes the derivation of the maximum entropy ranking with $\kappa(m)$ defining the me-consequence relation. Note that in deriving these constraints (7) and (8), the assumption is that an me-ranking, i.e., an integral solution to these equations, exists. Given the extra requirement that the constraints be satisfied as equalities, this may not always be the case. Reasons for this will be explored in section 4, but in the next section an algorithm is given which computes the me-ranking, if one does exist.

3 The me-algorithm

The me-ranking can be found if the equations (7) and (8) can be solved. However, since this is a set of non-linear simultaneous equations, there is no guarantee either that there is a solution or that a given solution is unique. Expanding equations (7) and (8) illustrates how the algorithm for computing the me-ranking was be devised.

Let v_r (respectively, f_r) represent a minimal verifying (respectively, falsifying) model of r in some ranking κ.

Definition 3.1 *An integer ranking, κ, over a set of defaults, $\{r_i : a_i \stackrel{s_i}{\Rightarrow} b_i\}$, is said to be* me-valid *with respect to that set if, for all r,*

$$\kappa(v_r) + s_r = \kappa(f_r) \tag{9}$$

and the ranking over models is determined by (8).

Since each falsifying model of a default has a contribution from its own me-rank, equation (9) can be re-written as

$$\kappa(r) + (\kappa(f_r) - \kappa(r)) = \kappa(v_r) + s_r \tag{10}$$

This equation can be expanded to give:

[2]Note that this approximation is, on the face of it, independent of the strengths s_i.

[3]Note that the function κ is used to represent both the ranking function over models and the me-ranks of the defaults themselves.

me-algorithm

```
Input:  a set of variable strength defaults, {r_i : a_i =s_i=> b_i}.
Output:  an me-valid ranking, κ, if one exists.

 [1] Initialise all κ(r_i) = INF.
 [2] While any κ(r_i) = INF do:
     (a) For all r_i with κ(r_i) = INF, compute
         MINV(r_i) + s_i.
     (b) For all such r_i with minimal MINV(r_i) + s_i,
         compute MINF(r_i).
     (c) Select r_j with minimal MINF(r_i).
     (d) If MINF(r_j) = INF let κ(r_j) := 0
         else let κ(r_j) := s_j + MINV(r_j) − MINF(r_j).
 [3] Assign ranks to models using equation (8).
 [4] Check constraints (7) to verify this is an me-valid ranking[4].
```

Figure 1: The me-algorithm

$$\kappa(r_i) + \min_{m \models a_i \wedge \neg b_i} \left[\sum_{\substack{r_j, j \neq i \\ \overline{m \models a_j \wedge \neg b_j}}} \kappa(r_j) \right] = s_i + \min_{m \models a_i \wedge b_i} \left[\sum_{\substack{r_j, j \neq i \\ \overline{m \models a_j \wedge \neg b_j}}} \kappa(r_j) \right] \tag{11}$$

Let the function MINV(r) (respectively, MINF(r)) be defined so that it returns the current minimal rank of all verifying models of r (respectively, the current minimal rank of all falsifying models of r *excluding its own contribution*) using equation (8). In the me-algorithm of Figure 1, equation (10), along with the functions MINV(r) and MINF(r), is used to compute the me-rank of each default iteratively via the assignment:

$$\kappa(r) := s_r + \mathrm{MINV}(r) - \mathrm{MINF}(r) \tag{12}$$

The remainder of this section sets out to demonstrate the claim that this algorithm computes an me-ranking, if one exists.

The first lemma shows that the me-algorithm always computes a finite set of ranks for the defaults provided the input set is p-consistent—meaning that it is possible to build a probability distribution over the defaults which respects the constraints they embody.

Lemma 3.2 *Given a p-consistent set of variable strength defaults, the me-algorithm assigns a finite rank to each default.*

Proof. Provided the minimal computed value for the function MINV(r) is finite at each pass of the loop, then the rank assigned to the chosen default will also be finite since if the computed value of MINF(r) is infinite it will be assigned rank 0, otherwise it will be assigned MINV(r) + s_r − MINF(r) which is also finite. Suppose therefore that at some pass of the loop the minimal computed value for MINV(r) is infinite for all unranked r. This means that all verifying models of each unranked default also falsify an unranked default, i.e., the set of defaults remaining to be ranked is unconfirmable. This contradicts the p-consistency of the original set and hence each default will be assigned a finite rank. □

[4]As will be seen later, this step may be replaced by a check for 1-fold cycles at step [2](c).

Given a p-consistent set of defaults, therefore, some set of finite ranks will be produced, which in turn implies a finite set of ranks over models. The next lemma shows that this represents a ranking function over models, i.e., that all ranks for models are non-negative and that at least one has zero rank.

Lemma 3.3 *Given a p-consistent set of variable strength defaults, the me-algorithm assigns a non-negative rank to each model.*
Proof. This is shown by induction. The rank of each model at any given stage equals the sum of the current ranks of those defaults it falsifies. At the start, as all defaults have infinite rank, the current rank of a model is either zero, if it falsifies no defaults, or infinite. Assume that at some intermediate stage all models have non-negative rank before the chosen default, r, is assigned a rank. Now, if the computed value of $\mathrm{MINF}(r)$ is infinite, the default is assigned a rank of 0 but this will not change the current rank of any model since all its falsifying models also falsify other unranked defaults. If, on the other hand, $\mathrm{MINF}(r)$ is finite then the default is assigned a rank of $\mathrm{MINV}(r) + s_r - \mathrm{MINF}(r)$. Now any falsifying models of r which only falsify other previously ranked defaults will all have a rank of greater than or equal to $\mathrm{MINV}(r) + s_r$ because $\mathrm{MINF}(r)$ was minimal among them. Any other falsifying models of r will still have infinite rank. The lemma follows by induction. $\square$

This lemma does not preclude a default from having a negative rank. Note that, at this stage, there is no guarantee that the computed ranking over models is me-valid, or even admissible, only that it represents a ranking. The following lemma shows that the defaults are ranked in an order corresponding to the ascending order of their $\kappa(v_r) + s_r$ in the final ranking.

Lemma 3.4 *Given a p-consistent set of variable strength defaults, the me-algorithm assigns ranks to defaults in ascending order of the final ranks of their minimal verifying models plus their strengths.*
Proof. The proof of lemma (3.3) shows that, at any stage, if a model's rank becomes finite it will be greater than or equal to that of the current default's computed $\mathrm{MINV}(r) + s_r$. Since, at each pass of the loop, r is chosen so that this is minimal, it also follows that no model which has infinite rank can subsequently obtain a lower final rank than the current $\mathrm{MINV}(r) + s_r$. This implies both that $\kappa(v_r) = \mathrm{MINV}(r)$ for the current r, and that the defaults are ranked in ascending order of their final $\kappa(v_r) + s_r$. $\square$

Corollary 3.5 *κ is admissible, that is, for all r*

$$\kappa(v_r) + s_r \leq \kappa(f_r)$$

Proof. Note that all falsifying models of a default have infinite rank when it is being ranked and so cannot have a final rank of less than $\kappa(v_r) + s_r$. $\square$

In order to show that the algorithm is sound, it is necessary to notice under what circumstances no me-valid ranking exists. In fact, this can be determined from the behaviour of the me-algorithm itself. The following definition will be helpful.

Definition 3.6 (Cycles) *An* n-fold cycle *is said to be encountered if, at stage [2](c) of the me-algorithm, n defaults are available to be ranked (i.e., have minimal $MINV(r) + s_r$) and all have infinite $MINF(r)$.*

Now the soundness of the me-algorithm can be established under the circumstance that no 1-fold cycle is encountered.

Theorem 3.7 (Soundness) *Given a p-consistent set of variable strength defaults, the me-algorithm produces an me-valid ranking iff no 1-fold cycle is encountered.*

Proof. For the if part. Assume that a 1-fold cycle occurs when some r is being ranked. Clearly r is assigned rank zero but all its falsifying models remain infinite. Now, since there were no other defaults available to be ranked along with r, all unranked defaults have a strictly higher value for $\mathrm{MINV}(r) + s_r$. When any of r's falsifying models become finite they will have a final rank strictly greater than $\kappa(v_r) + s_r$ and so κ is me-invalid.

For the only if part. From the corollary (3.5), κ is admissible. For κ to be me-invalid, it must be the case that for some r, $\kappa(v_r) + s_r < k(f_r)$. This can only occur if all its falsifying models also falsify another default, r', which has a higher final rank for $\kappa(v_{r'}) + s_{r'}$.[5] The me-algorithm will therefore encounter a 1-fold cycle when it ranks r because r' will not be available for ranking as it does not have minimal $\mathrm{MINV}(r) + s_r$. □

To establish the final result for the me-algorithm, that it is complete in the sense that it computes an me-valid ranking if one exists, it is necessary to show that the assignments for the ranks for defaults are compatible with an me-valid ranking, if one exists. Note that it is not important which me-valid ranking is computed just that one will be computed if any exist.

Theorem 3.8 (Completeness) *The ranks assigned to defaults in the me-algorithm are compatible with an me-valid ranking, if one exists.*

Proof. Assume that at least one me-valid ranking, κ', exists. The proof is inductive. Note that at the start, when no defaults have been ranked, this is compatible with κ'. Assume that at some point, when r_n is selected to be ranked, all previously ranked rules, r_i, $i < n$, have $\kappa(r_i) = \kappa'(r_i)$. Now, if v'_{r_n} is a minimal verifying model for r_n in κ', it falsifies only previously ranked defaults so $\kappa(v'_{r_n}) = \kappa'(v'_{r_n})$. As shown in theorem (3.7), the minimal verifying model of r_n in κ has rank $\mathrm{MINV}(r_n)$ which must equal $\kappa'(v'_{r_n})$ and v'_{r_n} is minimal in both κ' and κ.

Consider f'_{r_n}, a minimal falsifying model for r_n in κ'. Suppose it, too, falsifies only previously ranked defaults, apart from r_n, then $\mathrm{MINF}(r_n)$ will be finite and equal to $\kappa'(v'_{r_n}) + s_n - \kappa'(r_n)$. The rank assigned to $\kappa(r_n)$ is $\kappa(v'_{r_n}) + s_n - \mathrm{MINF}(r_n)$ which must equal $\kappa'(r_n)$. The assignment of $\kappa(r_n)$ is therefore compatible with κ'.

Suppose, on the other hand, that f'_{r_n} falsifies some as yet unranked defaults and let one of these be r_m with $m > n$. Now f'_{r_n} will also be a minimal falsifying model for r_m in κ' and the constraint of me-validity (9) is given by:

$$\kappa'(v'_{r_n}) + s_n = \kappa'(f'_{r_n}) = \kappa'(r_n) + \kappa'(r_m) + C$$

where C is a constant being the sum of the κ' ranks of all other defaults falsified by f'_{r_n}. Now there are infinitely many me-valid rankings which satisfy this equation with $\kappa'(r_n)$ and $\kappa'(r_m)$ varying, at least one of which has $\kappa'(r_n) = 0$ and so the assignment $\kappa(r_n) = 0$ will be compatible with at least one of these κ'. The theorem follows by induction. □

Theorems (3.7) and (3.8) lead to the following corollary.

Corollary 3.9 *An me-valid ranking exists iff the me-algorithm does not encounter a 1-fold cycle.*

It follows that step [4] in the me-algorithm may be replaced by a simple check for the existence of a 1-fold cycle at step 2[c]. If a 1-fold cycle is encountered then an exception can be raised to indicate that the ranking produced is not me-valid (though it will be admissible).

[5] Note that there may be several defaults r'.

m	a	b	c	r_1	r_2	r_3	$P(m)$
m_1	0	0	0	-	-	-	1
m_2	0	0	1	-	-	-	1
m_3	0	1	0	-	-	-	1
m_4	0	1	1	-	-	-	1
m_5	1	0	0	f	f	-	$\alpha_1\alpha_2$
m_6	1	0	1	f	v	-	α_1
m_7	1	1	0	v	f	f	$\alpha_2\alpha_3$
m_8	1	1	1	v	v	v	1

Table 1: Unnormalised probabilities for example 4.1.

Complexity

The main disadvantage of this new algorithm, which it shares with that of [5], is its intractability, requiring enumeration of models. Thus if $\mathcal{L}$ has n propositions, the me-algorithm will be polynomial in 2^n. The issue of complexity is not addressed in this paper and this is a severe problem with the me-approach in general [3]. However, the intention of this paper is to expound the theoretical benefits of the maximum entropy ranking, not its practicability. A current implementation[6] requires $\mathcal{L}$ to contain 16 or fewer propositions.

4 Uniqueness and existence criteria

In the derivation of the me-constraint equations, it was assumed that considering the asymptotic behaviour of the probabilities, and fixing the relative strength of defaults, determines the me-ranking. In both cases this amounts to ignoring their coefficients. It turns out that this assumption does not hold in general although it is possible to identify situations in which there are no solutions, and in which the solution is unique.

First, an example which illustrates why the assumption may not always be valid. Consider the following example where the probabilities (6) are used to show what may happen when all defaults have the same strength but the coefficients of these strengths are allowed to vary.

Example 4.1

$$\Delta = \{r_1 : a \overset{s_1}{\Rightarrow} b, r_2 : a \overset{s_2}{\Rightarrow} c, r_3 : a \wedge b \overset{s_3}{\Rightarrow} c\}$$

Table 1 shows whether a model falsifies or verifies each default and gives its (unnormalised) probability using equation (6):

Using the substitution $u = \frac{\varepsilon}{1-\varepsilon}$, with all defaults having equal strength of 1, and letting the coefficients of these be c_1, c_2, c_3, respectively, the constraint equations (5) give rise to three simultaneous equations:

$$\begin{aligned} \alpha_1\alpha_2 + \alpha_1 &= c_1 u(1 + \alpha_2\alpha_3) \\ \alpha_1\alpha_2 + \alpha_2\alpha_3 &= c_2 u(1 + \alpha_1) \\ \alpha_2\alpha_3 &= c_3 u \end{aligned}$$

Solving these for the α_i in terms of u gives[7]:

$$\alpha_1 = \frac{u(c_1 + c_1 c_3 u - c_2 + c_3)}{1 + c_2 u}$$

[6] Available at: `http://www2.elec.qmw.ac.uk/~rach/drs.html`.

[7] Note that $\sim$ has been replaced by $=$ since the coefficients are now relevant, however, this analysis is somewhat lax since by using (6) some approximation has already occurred. Nevertheless, the point being made is a valid one.

m	a	b	c	$(1,0,1)$	$(1,1,0)$	$(2,-1,2)$
m_1	0	0	0	0	0	0
m_2	0	0	1	0	0	0
m_3	0	1	0	0	0	0
m_4	0	1	1	0	0	0
m_5	1	0	0	1	2	1
m_6	1	0	1	1	1	2
m_7	1	1	0	1	1	1
m_8	1	1	1	0	0	0

Table 2: Multiple me-rankings for example 4.1.

$$\alpha_2 = \frac{c_1c_2u + c_1c_2c_3u^2 + c_2 - c_3}{c_1 + c_1c_2u - c_2 + c_3}$$
$$\alpha_3 = \frac{c_3u(c_1 + c_1c_2u - c_2 + c_3)}{c_1c_2u + c_1c_2c_3u^2 + c_2 - c_3}$$

Now consider what happens asymptotically for various values of the coefficients.

Case 1: Let $c_1 = 2(c_2 - c_3)$ (for $c_2 > c_3$). This gives a solution of $\alpha_1 \sim u$, $\alpha_2 \sim 1$, $\alpha_3 \sim u$ and leads to an me-ranking over defaults of (1,0,1).

Case 2: Let $c_2 = c_3$. This gives a solution of $\alpha_1 \sim u$, $\alpha_2 \sim u$, $\alpha_3 \sim 1$, and an me-ranking over defaults of (1,1,0).

Case 3: Let $c_1 + c_3 = c_2$. This gives a solution of $\alpha_1 \sim u^2$, $\alpha_2 \sim \frac{1}{u}$, $\alpha_3 \sim u^2$ and an me-ranking over defaults of (2,-1,2).

The me-rankings corresponding to each of these cases are given in Table 2. It is clear that different choices for the coefficients may lead to different me-rankings over the defaults and, more importantly, over the models. This corresponds to there being multiple solutions to the constraint equations given by (7) and (8).

For maximum entropy entailment to be well-defined, it is desirable to be able to determine when a unique me-ranking can be found. It has already been established that the me-algorithm can be used to identify when no solution exists (encountering a 1-fold cycle). The following results identify a unique me-ranking.

Definition 4.2 *An integer ranking, κ, over models is said to be* robust[8] *with respect to a set of defaults, $\{r_i : a_i \overset{s_i}{\Rightarrow} b_i\}$, if no two defaults share a common minimal falsifying model in κ.*

Theorem 4.3 *Given a finite set of defaults, $\{r_i : a_i \overset{s_i}{\Rightarrow} b_i\}$, if an me-valid ranking, κ, is robust then it is unique.*

The proof can be found in [4]. This definition of robustness is only a sufficient condition for uniqueness, however. It may be possible to find a necessary condition and this is the subject of ongoing research.

5 Examples

In the first example, the solution is tabulated explictly to illustrate the method of finding the me-ranking but later this is omitted to save space.

Example 5.1 (Exceptional inheritance)

$$\Delta = \{r_1 : b \overset{s_1}{\Rightarrow} f, r_2 : p \overset{s_2}{\Rightarrow} b, r_3 : p \overset{s_3}{\Rightarrow} \neg f, r_4 : b \overset{s_4}{\Rightarrow} w\}$$

[8]Adopting the use of "robustness" to indicate existence of a unique solution from [2].

m	b	f	p	w	r_1	r_2	r_3	r_4	$\kappa(m)$
m_1	0	0	0	0	-	-	-	-	0
m_2	0	0	0	1	-	-	-	-	0
m_3	0	0	1	0	-	f	v	-	$\kappa(r_2)$
m_4	0	0	1	1	-	f	v	-	$\kappa(r_2)$
m_5	0	1	0	0	-	-	-	-	0
m_6	0	1	0	1	-	-	-	-	0
m_7	0	1	1	0	-	f	f	-	$\kappa(r_2)+\kappa(r_3)$
m_8	0	1	1	1	-	f	f	-	$\kappa(r_2)+\kappa(r_3)$
m_9	1	0	0	0	f	-	-	f	$\kappa(r_1)+\kappa(r_4)$
m_{10}	1	0	0	1	f	-	-	v	$\kappa(r_1)$
m_{11}	1	0	1	0	f	v	v	f	$\kappa(r_1)+\kappa(r_4)$
m_{12}	1	0	1	1	f	v	v	v	$\kappa(r_1)$
m_{13}	1	1	0	0	v	-	-	f	$\kappa(r_4)$
m_{14}	1	1	0	1	v	-	-	v	0
m_{15}	1	1	1	0	v	v	f	f	$\kappa(r_3)+\kappa(r_4)$
m_{16}	1	1	1	1	v	v	f	v	$\kappa(r_3)$

Table 3: The me-ranking for the penguin example.

The intended interpretation of this knowledge base is that birds fly, penguins are birds, penguins do not fly and birds have wings. Table 3 shows whether a model falsifies or verifies each default. The column headed $\kappa(m)$ gives the me-rank of each model in terms of the $\kappa(r_i)$ using equation (8).
Substituting the $\kappa(m)$ into the reduced constraint equations (7) gives rise to:

$$\begin{aligned}
\kappa(r_1) &= s_1 \\
\kappa(r_2) &= s_2 + \min(\kappa(r_1), \kappa(r_3)) \\
\kappa(r_3) &= s_3 + \min(\kappa(r_1), \kappa(r_2)) \\
\kappa(r_4) &= s_4
\end{aligned}$$

Clearly, the only solution to these equations is $\kappa(r_1) = s_1$, $\kappa(r_2) = s_1 + s_2$, $\kappa(r_3) = s_1 + s_3$, and $\kappa(r_4) = s_4$.

To determine default consequences it is necessary to compare the ranks of a default's minimum verifying and falsifying models. Since this solution holds for any strength assignment (s_1, s_2, s_3, s_4), it follows that some default conclusions may hold in general. In particular, it can be seen that the default $p \wedge b \Rightarrow \neg f$ is me-entailed since

$$\begin{aligned}
\kappa(p \wedge b \wedge \neg f) &< \kappa(p \wedge b \wedge f) \\
s_1 &< s_1 + s_3
\end{aligned}$$

This result is unsurprising since $p \wedge b \Rightarrow \neg f$ is a preferential consequence of Δ. A more interesting general conclusion is $p \Rightarrow w$, which follows since

$$\kappa(p \wedge w) = s_1 < \kappa(p \wedge \neg w) = s_1 + \min(s_2, s_4)$$

Again this result holds regardless of the strength assignments and illustrates that, for this example, the inheritance of w to p via b is uncontroversial. □

Example 5.2 (Nixon diamond)

$$\Delta = \{r_1 : q \overset{s_1}{\Rightarrow} p, r_2 : r \overset{s_2}{\Rightarrow} \neg p\}$$

The intended interpretation is that quakers are pacificists whereas republicans are not pacifists. Given a strength assignment of (s_1, s_2) it is easily shown that $\kappa(r_1) =$

s_1 and $\kappa(r_2) = s_2$. The classical problem associated with this knowledge base is to ask whether Nixon, being a republican and a quaker, is pacifist or not. This is represented by the default $r \wedge q \Rightarrow p$. The two relevant models to compare are $r \wedge q \wedge p$ and $r \wedge q \wedge \neg p$ whose me-ranks in the general me-solution are

$$\kappa(r \wedge q \wedge p) = s_2 \quad \text{and} \quad \kappa(r \wedge q \wedge \neg p) = s_1 \tag{13}$$

Clearly either $r \wedge q \Rightarrow p$ or $r \wedge q \Rightarrow \neg p$, or neither, may be me-entailed depending on the comparative strengths s_1 and s_2. This result is in accordance with the "intuitive" solution that no conclusion should be drawn regarding Nixon's pacifist status unless there is reason to suppose that one default holds more strongly than the other. In the case of one default being stronger, the conclusion favoured by the stronger would prevail. □

Example 5.3 (Royal elephants/marine chaplains)

$$\Delta = \{r_1 : a \overset{s_1}{\Rightarrow} b, r_2 : c \overset{s_2}{\Rightarrow} b, r_3 : b \overset{s_3}{\Rightarrow} d, r_4 : a \overset{s_4}{\Rightarrow} \neg d\}$$

There are two interpretations of this knowledge base. In the first, the propositions a, b, c, and d, stand for royal, elephant, african and grey, respectively; in the second, the propositions stand for chaplain, man, marine and beer drinker, respectively. The constraint equations (7) give rise to:

$$\begin{aligned}
\kappa(r_1) &= s_1 + \min(\kappa(r_3), \kappa(r_4)) \\
\kappa(r_2) &= s_2 \\
\kappa(r_3) &= s_3 \\
\kappa(r_4) &= s_4 + \min(\kappa(r_1), \kappa(r_3))
\end{aligned}$$

which have the unique solution $\kappa(r_1) = s_1 + s_3$, $\kappa(r_2) = s_2$, $\kappa(r_3) = s_3$, and $\kappa(r_4) = s_3 + s_4$.

The key question relating to this knowledge base is "Are elephants which are both royal and african, not grey?", or alternatively, "Don't marine chaplains drink beer?" This translates into the default $a \wedge c \Rightarrow \neg d$ which is me-entailed in general as can be seen from:

$$\begin{aligned}
\kappa(a \wedge c \wedge \neg d) &< \kappa(a \wedge c \wedge d) \\
s_3 &< s_3 + \min(s_4, s_1 + s_2 + s_3)
\end{aligned}$$

The result in this example is unambiguous, that is, it holds for all strength assignments. However, Touretzky *et al* [14] were not entirely happy about the conclusion that marine-chaplains do not drink beer. They argued that if the rate of beer drinking amongst marines was significantly higher than normal, then this might alter the behaviour associated with marine-chaplains.

Now, the default $r_5 : c \Rightarrow d$ (marines drink beer) is in fact me-entailed by Δ, but adding it to the database with all defaults having equal strength would violate the robustness condition. If, however, r_5 were added with a higher strength, so that it represents an extra constraint in the entropy maximization, a robust solution results and the status of the default $a \wedge c \Rightarrow \neg d$ depends on the relative strengths s_4 and s_5.

So, Touretzky *et al.* were correct in supposing that if marines were heavier drinkers than men in general then it may not be clear whether marine chaplains are beer drinkers or not. However, as this information is not explicitly represented, it is unsurprising that conclusions based on it do not occur. This example illustrates an important point, sometimes seemingly overlooked [13], that any reasoning system can only reason with the information that is available to it. The beauty of the me-approach is that it highlights exactly what is implied by the data, *and only what is implied by that data.* □

It is interesting to note that many of the more complex examples from the literature (for example, see [9]), which have been devised deliberately to overcome any intuitive biases, fail to satisfy the robustness condition when all defaults are assigned equal strengths. If a set is probabilistically consistent it is possible to restore robustness by altering the strengths. This suggests that some sets may be too complex for human intuition to disentangle because they are ambiguous or underspecified. Because the me-approach requires more information from the knowledge engineer, in terms of a strength assignment over defaults, some of these ambiguities can be cleared up and the hitherto implicit biases made explicit.

6 Conclusions

Using the me-approach for default reasoning provides the same benefits as its use in statistical problems. As Jaynes [6] suggests, by encoding all known relevant information and finding the maximum entropy distribution, any observations which differ significantly from the result imply that other constraints, in this case defaults, exist. A closer approximation to the desired model is obtained by adding more defaults or by adjusting the strengths. Rather than questioning the conclusions of a default reasoning system, one needs to ensure that all relevant information has been encoded—the maximum entropy formalism enables the precise and explicit representation of this as default knowledge and moreover has an objective justification based on the principle of indifference.

This paper has refined and extended the work of [5] on applying the principle of maximum entropy to the probabilistic semantics for default rules to enable it to be applied to arbitrary sets of variable strength defaults. A new algorithm was presented which finds a maximum entropy ranking and establishes existence, and a sufficient condition to determine uniqueness was given. This extension to arbitrary sets has shed some light onto the causes of controversy among classical examples from the literature and pointed to ways of resolving them.

References

[1] E. Adams. *The Logic of Conditionals*. Reidel, Dordrecht, Netherlands, 1975.

[2] F. Bacchus, A. J. Grove, J. Y. Halpern, and D. Koller. From statistical knowledge bases to degrees of belief. *Artificial Intelligence*, 87:75–143, 1996.

[3] R. Ben-Eliyahu. *NP-complete problems in optimal horn clause satisfiability.* Technical report r-158, Cognitive Systems Laboratory, UCLA, Los Angeles, 1990.

[4] R. A. Bourne and S. Parsons. Maximum entropy and variable strength defaults. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 1999.

[5] M. Goldszmidt, P. Morris, and J. Pearl. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:220–232, 1993.

[6] E. Jaynes. Where do we stand on maximum entropy? In R. Levine and M. Tribus, editors, *The Maximum Entropy Formalism*, pages 15–118, Cambridge, MA, 1979. MIT Press.

[7] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[8] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[9] D. Makinson and K. Schlechta. Floating conclusions and zombie paths: two deep difficulties in the "directly skeptical" approach to defeasible inheritance nets. *Artificial Intelligence*, 48:199–209, 1991.

[10] J. Paris and A. Vencovská. In defense of the maximum entropy inference process. *International Journal of Approximate Reasoning*, 17:77–103, 1997.

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

[12] J. Pearl. System Z: a natural ordering of defaults with tractable applications to default reasoning. In *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 121–135, 1990.

[13] R. Reiter and G. Criscuolo. Some representational issues in default reasoning. *Computers & Mathematics with Applications*, 9:15–27, 1983.

[14] D. S. Touretzky, J. F. Horty, and R. H. Thomason. A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 476–482, 1987.