Some qualitative approaches to applying the Dempster-Shafer theory

Simon Parsons
Advanced Computation Laboratory
Imperial Cancer Research Fund
P.O. Box 123
Lincoln's Inn Fields
London WC2A 3PX

Abstract

This paper introduces the idea of using the Dempster-Shafer theory of evidence with qualitative values. Dempster-Shafer theory is a formalism for reasoning under uncertainty which may be viewed as a generalisation of probability theory with special advantages in its treatment of ambiguous data and the ignorance arising from it. Here we are interested in applying the theory when the numbers that it usually operates over are not universally available. To cope with this lack of numbers, we use qualitative, semiqualitative, and linguistic values, and apply a form of order of magnitude reasoning.

1 Introduction

Dempster-Shafer theory is a numerical method for evidential reasoning. The theory originated with a paper by the statistician Arthur Dempster [7] who wanted to free probability theory from the need to attach a measure of uncertainty to every hypothesis under consideration [11]. His work remained hidden in the statistics literature until Glenn Shafer, one of Dempster's students, brought the material to a wider audience in his doctoral dissertation [16]. The method has become popular, and the basic model has been extended in a number of directions in recent years [17], [19], [20]. In this paper I propose another adaptation of the model. My interest is in reasoning under uncertainty when all the numerical information required by methods such as Dempster-Shafer theory are not available, handling such a lack of information [13] [14] by using techniques from qualitative reasoning [1]. Extending the approach first suggested in [15], I consider replacing the numerical operands of more usual applications of Dempster-Shafer theory with qualitative values. These give degraded, but still useful, results which are illustrated by a number of examples.

In Section 2, the basics of Dempster-Shafer theory are explained for the benefit of those who are not familiar with the approach. Section 3 introduces the qualitative version of the theory, and Section 4 applies the theory to linguistic and other semiqualitative values, while Section 5 assesses what can be done with knowledge of the absolute order of magnitude of values. Section 6 concludes.

2 Dempster-Shafer theory

The basic idea of the Dempster-Shafer theory is that numerical measures of uncertainty, termed basic probability masses, may be assigned to sets of hypotheses as well as individual hypotheses. Consider the following example, adapted from the work of Philippe Smets [17]. Mr Jones has been murdered. We know that the murderer was one of three notorious assassins, Peter, Paul and Mary, so we have a set of hypotheses $\Theta = \{Peter, Paul, Mary\}$. The only evidence that we have initially is that of Mrs Jones who saw the killer leaving the scene of the murder and is 80% sure that it was a man. Thus all we know is that $p(Man) = 0.8$. If we were using probability theory we would have to:

(a) allocate $p(\neg Man) = p(Mary) = 1 - 0.8 = 0.2$
(b) allocate $p(Man) = 0.8 = p(Peter) + p(Paul) = 0.4 + 0.4$

The first since $p(Man)+p(\neg Man) = 1$, and the second by some principle such as the principle of maximum entropy. With evidence theory, however, we are not limited to allocating probability to the members of the set $\{\{Peter\}, \{Paul\}, \{Mary\}\}$. We have instead a mass assignment function $m(\cdot)$ where $m : 2^\Theta \mapsto [0, 1]$ assigns probabilities to any set which is a member of the power set of Θ, that is the set $2^\Theta = \{\{Peter, Paul, Mary\}, \{Peter, Paul\}, \{Peter, Mary\}, \{Paul, Mary\}, \{Peter\}, \{Paul\}, \{Mary\}, \emptyset\}$. The only restrictions on $m(\cdot)$ are:

$$\sum_{x \in 2^\Theta} m(x) = 1 \tag{1}$$

$$m(\emptyset) = 0 \tag{2}$$

so that all the assigned probabilities sum to unity, and there is no belief in the empty set. Note that any subset x of the frame Θ for which $m(x)$ is non-zero is called a focal element.

In the case of Mr Jones' murder we can assign values to equate with what we know and nothing more. We know that $p(Man) = 0.8$ so that the focal element is $\{Peter, Paul\}$ and $m(\{Peter, Paul\}) = 0.8$. We know nothing about the remaining probability so it is allocated to the whole frame of discernment— $m(\{Peter, Paul, Mary\}) = 0.2$. Now, consider that a second piece of evidence comes to light. It is reported with confidence 0.6 that Peter was leaving on a jet plane when the murder occurred, so that we have $m'(\{Paul, Mary\}) = 0.6$, and $m'(\{Peter, Paul, Mary\}) = 0.4$. We would like to combine these two pieces of evidence, and this may be done by combining the mass assignments using Dempster's rule to create a new mass assignment m'' defined by:

$$m''(C) = \sum_{\substack{i,j \\ A_i \cap B_j = C}} m(A_i)m'(B_j) \tag{3}$$

We will write $m'' = m \otimes_\cap m'$ as a shorthand for this operation. Put simply, the result of combining two assignments is that for any intersecting sets A and B, where A has mass M from assignment m and B has mass M' from assignment m', the belief accruing to their intersection is the product of M and M'. the combination for our example is given in Table 1. Having established the final mass assignments of the set of hypotheses we can

$\otimes_\cap$	$m(\{Peter, Paul\}) = 0.8$	$m(\Theta) = 0.2$
$m'(\{Paul, Mary\}) = 0.6$	$m''(\{Paul\}) = 0.48$	$m''(\{Paul, Mary\}) = 0.12$
$m'(\Theta) = 0.4$	$m''(\{Peter, Paul\}) = 0.32$	$m''(\Theta) = 0.08$

Table 1: Combining the mass functions for the murder example

assess the belief and plausibility of any set of hypotheses as follows:

$$Bel(A) = \sum_{B \subset A} m(B) \tag{4}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{5}$$

These measures are clearly related to one another:

$$Bel(A) = 1 - Pl(\neg A) \tag{6}$$

$$Pl(A) = 1 - Bel(\neg A) \tag{7}$$

The belief in any set is the sum of all the probabilities of all the subsets of that set. The plausibility is the sum of all the values not accruing to any sets that are exclusive of the one in question. The function $Bel : 2^\Theta \mapsto [0, 1]$ is known as a *belief function* [16] and its dual $Pl : 2^\Theta \mapsto [0, 1]$ is called a *plausibility function* [17]. Calculating the belief and plausibility in the case of Mr Jones' murder:

$$\begin{aligned}
Bel(\{Paul\}) &= 0.48 \\
Bel(\{Peter, Paul\}) &= m''(\{Peter\}) + m''(\{Paul\}) + m''(\{Peter, Paul\}) \\
&= 0 + 0.48 + 0.32 \\
&= 0.8 \\
Bel(\{Peter, Paul, Mary\}) &= 1 \\
Pl(\{Peter\}) &= m''(\{Peter, Paul\}) + m''(\{Peter, Paul, Mary\}) \\
&= 0.4 \\
Pl(\{Mary\}) &= 0.2
\end{aligned}$$

In a similar way, $Bel(\{Peter\}) = Bel(\{Mary\}) = Bel(\{Peter, Mary\}) = 0, Bel(\{Paul, Mary\}) = 0.6, Pl(\{Peter, Mary\}) = 0.52$, and $Pl(\{Paul\}) = Pl(\{Paul, Mary\}) = Pl(\{Peter, Paul, Mary\}) = 1$. This rule of combination is unproblematic so long as no two focal elements have an empty intersection, that is as long as

$$m(\emptyset) = \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m(A_i) m'(B_j) = 0 \tag{8}$$

Violations of this condition are problematic. What a non-zero mass for $\emptyset$ suggests is that there is belief in a hypothesis that is not in the frame of discernment. However, the frame of discernment is an exhaustive set of hypotheses, and so it is not possible to have belief in something outside it.

One possible way around the problem is to normalise the result of applying Dempster's rule, dividing the mass assigned to every focal element of m'' by:

$$1 - \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m(A_i) m'(B_j) \tag{9}$$

This is the approach advocated by Shafer [16]. Thus the normalised version of Dempster's rule is:

$$m''(C) = \frac{\sum_{\substack{i,j \\ A_i \cap B_j = C}} m(A_i)m'(B_j)}{1 - \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m(A_i)m'(B_j)} \tag{10}$$

Now, normalisation is only required when the evidence that is described by the mass assignments disagree. Thus, in the case of Mr Jones' murder, normalisation would be required if the second piece of evidence were that both Peter and Paul were on the 'plane. Then one piece of evidence would indicate that the culprit was either Peter or Paul and the other would indicate that it was Mary. Recognising this has led some authors, including Zadeh [21], to criticise the use of normalisation since it can lead to counter-intuitive results if most of the mass after a combination, but before normalisation, is assigned to the empty set.

The problem with assigning mass to the empty set arises from the assumption that Θ is exhaustive. Smets [17] gets around the problem by making an "open world assumption" that the real solution may lie outside the frame of discernment. Under such an assumption any mass that is assigned to $\emptyset$ after a combination is taken to be the belief that the solution is a hypothesis that is not included in Θ. This position is not without its critics, see for example [3].

One point that should be noted is the interpretation of the belief and plausibility measures. There are many points of view. Smets [18] takes Dempster-Shafer belief to be a quantification of subjective *credal* belief that is distinct from the probabilistic *pignistic* belief that is necessary for decision making. Others take belief and plausibility as lower and upper bounds on the probability that may be assigned to an event. Under such an interpretation, the results of applying the Dempster-Shafer theory are consistent with any probabilistic analysis of the same problem, they just make less assumptions.

Finally in this section, it is worth noting that, despite the fact that it is often criticised as being computationally intractable [1] the Dempster-Shafer theory has been used to build complex applications, including the combination of visual evidence in a robot navigation system [10].

3 A qualitative approach to evidence theory

As it stands, evidence theory works very well as long as all the necessary numerical information is available. Provided that we can put a basic probability mass on any piece of evidence that comes to light then the theory gives us intuitive results. However, a problem arises when we do not have easily quantifiable evidence. For instance we may be taking readings from faulty sensors, or we may be dealing with data which relates to occurrences that happen so rarely that no accurate numbers are available. In such cases all we can say about a particular piece of evidence is that it indicates that certain hypotheses are true to 'a certain degree'. To what degree 'a certain degree' is we have no idea. What we would like is to use the intuitive evidence theory style of reasoning to combine such pieces of evidence to give us some idea of what the evidence implies. The following sections describe a method, which is an extension of that discussed in [15], in which qualitative values are used to provide representations of 'a certain degree' which

[1] This problem should have been comprehensively laid to rest by Wilson [20] who has provided efficient algorithms for both exact and approximate computation of the result of combining mass functions

may be combined with other pieces of evidence and summed to give degrees of belief and plausibility

3.1 Basic qualitative values

Consider what would have happened in the enquiry into Mr Jones' murder if Mrs Jones could not say how sure she was that the person she saw running from the house was a man. Say that she is sure that it could be a man so that $m(Man) > 0$, and she is unable to say for sure that it was a man so that $m(Man) < 1$, but is unwilling to commit herself further. In such a case we have $m(\{Peter, Paul\}) = something$, $m(\{Peter, Paul, Mary\}) = 1 - something$ where $0 < something < 1$. The value *something* may not be manipulated using normal arithmetic. Instead we convert it to a qualitative value using the mapping $[\![\cdot]\!] : \Re \mapsto \{+, 0\}$:

$$\text{For any } x \in \Theta,\ m(x) = M, \text{ becomes } [\![m(x)]\!] = \begin{cases} +, & \text{if } 0 < M < 1; \\ 0, & \text{if } M = 0. \end{cases} \tag{11}$$

which says that any mass that is known to be non-zero is represented by the qualitative value + while zero values are represented by the value 0. It also makes the assumption that a given basic probability assignment assigns mass to at least two members of 2^Θ ensuring that $0 < m(x) < 1$. Applying this mapping from unknown numerical values to qualitative values makes no assumptions about the value of *something*, but enables us to manipulate the unknown value using the well-established methods of qualitative arithmetic [1]. Thus when Mrs Jones is unwilling to give a numerical estimate, we have $p(Man) = +$ so that $m(\{Peter, Paul\}) = +$ and $m(\{Peter, Paul, Mary\}) = +$. Mass assignments are combined using Dempster's rule as before, with arithmetic being performed using restricted versions of the standard combinator tables for qualitative addition $\oplus$ and qualitative multiplication $\otimes$ [1].(see Table 2). So when the witness in the aeroplane also

$\oplus$	+	0
+	+	+
0	+	0

$\otimes$	+	0
+	+	0
0	0	0

Table 2: Qualitative combinator tables

refuses to give a numerical estimate of how sure they are that Peter was on the plane, we have $m'(\{Paul, Mary\}) = +$, $m'(\{Peter, Paul, Mary\}) = +$, and these may be combined as in Table 3 to give the following qualitative beliefs and plausibilities:

$$\begin{aligned} Bel(\{Paul\}) &= + \\ Bel(\{Peter, Paul\}) &= m''(\{Peter\}) + m''(\{Paul\}) + m''(\{Peter, Paul\}) \\ &= + \\ Bel(\{Peter, Paul, Mary\}) &= + \end{aligned}$$

$\otimes_\cap$	$m(\{Peter, Paul\}) = +$	$m(\Theta) = +$
$m'(\{Paul, Mary\}) = +$	$m''(\{Paul\}) = +$	$m''(\{Paul, Mary\}) = +$
$m'(\Theta) = +$	$m''(\{Peter, Paul\}) = +$	$m''(\Theta) = +$

Table 3: Combining the qualitative mass functions for the murder example

$$
\begin{aligned}
Pl(\{Peter\}) &= m''(\{Peter, Paul\}) + m''(\{Peter, Paul, Mary\}) \\
&= + \\
Pl(\{Mary\}) &= +
\end{aligned}
$$

At first sight these results don't seem to be very helpful, since all the sets of hypotheses have the same degree of support from the evidence. However, this first impression is not really correct. What the stripping away of the numbers makes extremely clear is that the beautiful and intuitive mechanism of evidence theory works just as well without numbers as it does with them, and it continues to lay bare the implication of the evidence.

What we can see from this, just as well as we can see from the numerical example, is that there is only one singleton hypothesis that is indicated by the evidence, $\{Paul\}$, and that if we want to consider hypotheses of the form 'A or B', then there is evidence for $\{Paul, Mary\}$ and $\{Peter, Paul\}$[2]. The method will even detect evidence for solutions other than those in the frame of discernment $\{Peter, Paul, Mary\}$, if the open world assumption is accepted, by the accruing of a + to the empty set $\emptyset$ when the focal elements of the mass functions do not intersect. This will be discussed further in Section 3.3.

Of course it is possible to invent pathological cases where the intuitive result is the wrong one. Consider what would happen if $m'(\{Paul, Mary\})$ were 0.1. The final result of the weighing of the evidence would be $Bel(\{Paul\}) = 0.08$, $Bel(\{Peter, Paul\}) = 0.8$ which suggests that there is little evidence against Paul alone, while the qualitative solution would be the same as before. However, this does not mean that there is no virtue in using the qualitative approach to establish which way the evidence points in particular situations where no numerical probability masses may be established, it merely makes the point that this qualitative approach is heuristic. Thus it will often produce correct answers where the numerical method would be unable to do so, but will, on occasion produce incorrect answers.

3.2 A simple example

By way of illustrating the usefulness of qualitative evidence theory we will consider an example from decision making in gastroenterology which was also studied in [15]. We consider a clinic specialising in the diagnosis of gastroenterological complaints. These complaints are gastric cancer, peptic ulcers (both gastric and duodenal ulcers), gallstones, and functional disorders. The latter are conditions with no identifiable organic cause, and are often stress related. Over many years, a number of symptoms and signs which provide useful information for discriminating between complaints have been recorded from many patients. These are signs of jaundice, pain after meals, weight loss and the age of the patient.

The clinic's research into gastric disorders has progressed since it was reported in [15]. Workers at the clinic have now established that particular symptoms are evidence for particular sets of diseases. Thus jaundice indicates gallstones or functional disorder $\{gs, fd\}$, pain after meals indicates gastric cancer, peptic ulcer or functional disorder $\{gc, pu, fd\}$, weight loss indicates gastric cancer $\{gc\}$ and if the patient is elderly then she is likely to be suffering from gastric cancer, peptic ulcer or gallstones $\{gc, pu, gs\}$. We

[2]Since the set $\{Peter, Paul\}$ not only represents the set of hypotheses { 'Peter killed Mr Jones','Paul killed Mr Jones' } but is also the hypothesis 'Peter or Paul killed Mr Jones', I will use the term 'hypothesis' to denote both the individual members of $\{Peter, Paul\}$ and the set itself. Which is being referred to will be clear from the context.

$\otimes_\cap$	$m(\{gc, pu, fd\}) = +$	$m(\{gc, pu, gs, fd\}) = +$
$m'(\{gc, pu, gs\}) = +$	$m''(\{gc, pu\}) = +$	$m''(\{gc, pu, gs\}) = +$
$m'(\{gc, pu, gs, fd\}) = +$	$m''(\{gc, pu, fd\}) = +$	$m''(\{gc, pu, gs, fd\}) = +$

Table 4: The result of considering two pieces of evidence about Jack

$\otimes_\cap$	$m'''(\{gc\}) = +$	$m'''(\{gc, pu, gs, fd\}) = +$
$m''(\{gc, pu\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu\}) = +$
$m''(\{gc, pu, gs\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu, gs\}) = +$
$m''(\{gc, pu, fd\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu, fd\}) = +$
$m''(\{gc, pu, gs, fd\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu, gs, fd\}) = +$

Table 5: The result of considering a third piece of evidence about Jack

are interested in the case of Jack, an elderly patient who shows no signs of jaundice but has recently lost weight and often has pain after eating.

Considering the evidence of Jack's age and the fact that he suffers pain after eating gives us the results of Table 4, which suggests that the most specific diagnosis is that Jack is suffering from either gastric cancer or peptic ulcer. Now we consider the evidence that Jack has recently lost weight. This gives the results of Table 5.which strongly indicates that Jack has gastric cancer since not only is gastric cancer the only singleton hypothesis indicated by the evidence, but it is also a member of every single other set of hypotheses that the evidence points to.

3.3 Extending the basic model

The basic approach discussed in Sections 3.1 and 3.2 is unproblematic when dealing with simple sets of evidence. When there is only one set of hypotheses that is smaller than any other, then it is clear which is the most specific set of hypotheses favoured by the evidence, and it seems natural to make the heuristic assumption that this set of hypotheses is to be preferred as an explanation of the evidence. Thus in Jack's case the diagnosis of $\{gc\}$ is the one that best fits the evidence of his three symptoms. However, it will not always be the case that the set of preferred hypotheses are so obvious. Take the case of Irwin, another patient who comes to the clinic of Section 3.2 as an example of the problems that might arise.

When Irwin's symptoms are related to the diseases that the clinic specialises in, it is clear that there is evidence for two possible sets of diseases, gallstones or functional disorder $\{gs, fd\}$ and gastric cancer or peptic ulcer $\{gc, pu\}$. The result of combining this evidence is given in Table 6. The combined mass distribution given in Table 6 is somewhat more complicated to understand than those considered previous, largely because it introduces the vexed issue of normalisation. Part of the basic probability assignment m'' is given to the empty set $\emptyset$. Under the closed world assumption this is not permitted, and the mass must be re-distributed by dividing the masses of all the

$\otimes_\cap$	$m(\{gs, fd\}) = +$	$m(\{gc, pu, gs, fd\}) = +$
$m'(\{gc, pu\}) = +$	$m''(\emptyset) = +$	$m''(\{gc, pu\}) = +$
$m'(\{gc, pu, gs, fd\}) = +$	$m''(\{gs, fd\}) = +$	$m''(\{gc, pu, gs, fd\}) = +$

Table 6: The result of considering two pieces of evidence about Irwin

remaining members of 2^Θ by $1 - +$. This is easy to accomplish. Dividing the masses by $1 - +$, which is positive since the $+$ discussed here is known to represent a number strictly less than one (11), does not change anything, giving a new distribution:

$$\begin{aligned} m''_n(\{gs, fd\}) &= + \\ m''_n(\{gc, pu\}) &= + \\ m''_n(\{gc, pu, gs, fd\}) &= + \end{aligned}$$

This clearly indicates that there are two preferred sets of hypotheses; $\{gs, fd\}$ and $\{gc, pu\}$. However, what normalising misses is the main conclusion of the evidence. The reason that mass is assigned to the empty set is that the two pieces of evidence that generate m and m' are completely contradictory. One suggests gastric cancer or peptic ulcer, the other suggests gallstones or functional disease, so that there is absolutely no common ground between them. If we want to acknowledge this contradiction, we must accept the open world assumption.

Under the open world assumption, the restriction concerning the assignment of mass to the empty set (2) is dropped, and belief and plausibility in a set of hypotheses is computed, as according to Smets [17], by:

$$Bel(A) = \sum_{\substack{B \subset A \\ B \neq \emptyset}} m(B) \tag{12}$$

As a result of the assumption, the mass assigned to the empty set has a new and useful meaning. It allows us to compute the amount of belief in a hypothesis outside the frame of discernment, that is a hypothesis that was not originally thought to be worth considering (or even a hypothesis that was not even thought of) but which becomes a likely option in the light of the evidence. It is worth noting that here the expression 'a hypothesis' (rather than 'a set of hypotheses') is used advisedly, not least because Smets takes $\emptyset$ to be a single hypothesis. It makes sense to assume it is a single hypothesis on similar grounds to those on which the single fault hypothesis [5] is adopted— that if the evidence is pointing to something that previously escaped notice, this is more likely to be one thing than many things.

In the case of Irwin's visit to the clinic, the meaning of the mass assigned to the empty set is that he is suffering from none of the usual diseases. This interpretation seems to be a natural and reasonable conclusion to draw from the conflict in the evidence, and it certainly seems to be a better conclusion that that drawn using the closed world assumption. Thus, we can say that when the redistribution of mass by normalisation fails to suggest a single preferred hypothesis, then the open world assumption may provide a better interpretation of the evidence [3].

This adoption of the open world assumption solves one possible problem with the purely qualitative quantifiers, but there is another. Consider how the preferred hypothesis changes when a third piece of evidence, which indicates that Irwin is suffering from gastric cancer, is considered. The result of doing this is given in Table 7. Since the empty set is taken to represent a single hypothesis outside Θ, the result of this third piece of evidence is to suggest that there are two singleton hypotheses as to which disease Irwin is sufering from. These are that Irwin is suffering from gastric cancer, and that Irwin is suffering from something other than gastric cancer, peptic ulcer, gallstones or functional disorder.

[3] And it is interesting to note that in most criticisms of normalisation this is indeed the case— equal masses are assigned to the two contradictory hypotheses.

$\otimes_\cap$	$m''(\{gc\}) = +$	$m''(\{gc, pu, gs, fd\}) = +$
$m'''(\{gs, fd\}) = +$	$m''''(\emptyset) = +$	$m''''(\{gs, fd\}) = +$
$m'''(\{gc, pu\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu\}) = +$
$m'''(\emptyset) = +$	$m''''(\emptyset) = +$	$m''''(\emptyset) = +$
$m'''(\{gc, pu, gs, fd\}) = +$	$m''''(\{gc\}) = +$	$m''''(\{gc, pu, gs, fd\}) = +$

Table 7: The result of considering a third piece of evidence about Irwin

This is as much as this qualitative application of the Dempster-Shafer theory will tell us— that is there are two possibilities, each with weight +. The question is, what is the diagnosis?

There are a couple of options. One is not to choose one disease over another. We could argue that since both $\emptyset$ and $\{gc\}$ are singleton hypotheses, and so are both 'most specific' sets of hypotheses, both satisfy the heuristic condition discussed above for being preferred hypotheses and so both should be taken as the most likely hypothesis. Thus we have a preferred set of hypotheses $\{gc, \emptyset\}$. This is a rather cautious approach, which does not discard any hypothesis which is strongly suggested by the evidence, and this caution no doubt makes it attractive for some applications [4]. However, this interpretation of results goes somewhat against the intuitions of the original Dempster-Shafer theory since it suggests that the preferred set of hypotheses is one which is not a focal element of any mass assignment (since it is a composite of two such focal elements). What this means is that it should be possible to calculate seperate beliefs for both $\emptyset$ and $\{gc\}$ and use these to make a choice between them, and doing so is perhaps more in keeping with the original theory.

To calculate the beliefs for $\emptyset$ and $\{gc\}$ with a view to selecting one as more likely than the other involves the use of another heuristic, albeit one for which there is a good deal of experimental support [2] [6]— an improper linear model. Using an improper linear model in this case is rather simple. We add up the number of pieces of evidence for each of the two preferred hypotheses (since there are no pieces of evidence directly against them) and select as most likely the singleton hypothesis with the most pieces of evidence in its favour. In Irwin's case, since there are more indications that he is suffering from a disease outside the frame of discernment (three) than there are that he is suffering from gastric cancer (two), we say that belief in him suffering from gastric cancer is less than belief in him suffering from a disease that is not under consideration.

4 Semiqualitative approaches

It is possible to further refine the qualitative approach if we have some idea of the numerical values of the mass assignments to the focal elements. This section discusses two ways in which this may be done. Both combine elements of qualitative and quantitative information, and so are known as semiqualitative approaches.

4.1 Using linguistic values

For instance, consider that we have information that allows us to quantify the mass assignments in terms of a set of linguistic labels, which correspond to a subintervals of

[4] Especially medical ones. If I were a patient at the clinic with the same symptoms as Irwin I would like the suggestion of gastric cancer to be taken very seriously indeed.

the unit interval $[0, 1]$, in a similar manner to that considered by [8]. We take the following symmetric set of subintervals: $\mathcal{P} = \{0, (\epsilon, a], [a, 1-a], [1-a, 1-\epsilon), 1\}$ whose names are, respectively; $None$, $Little$, $(About)Half$, $Much$ and All. These subintervals are ordered:

$$None \leq Little \leq Half \leq Much \leq All \tag{13}$$

As in [8] we take ϵ to be some positive infinitesimal quantity while a is some number in $(0, 0.5)$ and assume that we have enough knowledge to place the value of a mass assignment into one of the intervals. In practice we take $a = 0.3$, so that it is appropriate to note (as in [8]) that 'about half' is short for 'neither little, nor much but somewhere in between'. In order to manipulate these linguistic values we need some means of performing arithmetic on numbers and intervals. This ability is provided by interval arithmetic [12]. For any pair of intervals $[a, b]$ and $[c, d]$, where $a \leq b$ and $c \leq d$, interval addition $\oplus'$, multiplication $\otimes'$, subtraction $\ominus'$ and division $\oslash'$ are defined by:

$$[a, b] \oplus' [c, d] = [a + c, b + d] \tag{14}$$
$$[a, b] \otimes' [c, d] = [a \times c, b \times d] \tag{15}$$
$$[a, b] \ominus' [c, d] = [a - d, b - c] \tag{16}$$
$$[a, b] \oslash' [c, d] = [a/d, b/c] \tag{17}$$

Note that 'degenerate'[12] intervals $[a, b]$ where $a = b$ are allowed so that interval arithmetic may be applied to numbers and intervals together, and that interval division is not defined for $c = 0$ or $d = 0$. Given that the mass assignments are quantified with the set of linguistic labels, we can use interval arithmetic and Dempster's rule to compute the result of combining evidence in terms of the linguistic labels and combinations of labels, such as $[Little, Much] = \{X \in \mathcal{P} | Little \leq X \leq Much\}$. Such combinations of intervals are necessary given the tendency of interval arithmetic to expand the bounds on intervals when they are combined. Thus when adding $Little$ and $Little$, which is performing the interval calculation $[0, 0.3] + [0, 0.3]$, we get $[0, 0.6]$ by (14). This is an interval which does not have a linguistic label in $\mathcal{P}$, but it is clear that the value which is known to lie in $[0, 0.6]$ must lie in the interval $[0, 0.7]$ which is represented by $[Little, Much]$, and is the smallest compound interval that will contain $[0, 0.6]$. This compound interval may be given the gloss 'between a little and most'.

Now, to hark back to Jack's trip to the clinic, consider what would have happened if we had the information that if a patient has pain after meals we should have much belief in his having gastric cancer, peptic ulcer or functional disorder $\{gc, pu, fd\}$, whilst if a patient is elderly then we should have a middling belief that he is suffering from gastric cancer, peptic ulcer or gallstones $\{gc, pu, gs\}$. For Jack, we now have the combined mass assignment given in Table 8 which is rather more precise than before. We could, of course,

$\otimes_\cap$	$m(\{gc, pu, fd)\} = Much$	$m(\{gc, pu, gs, fd\}) = Little$
$m'(\{gc, pu, gs\}) = Half$	$m''(\{gc, pu\}) = [Little, Half]$	$m''(\{gc, pu, gs\}) = Little$
$m'(\{gc, pu, gs, fd\}) = Half$	$m''(\{gc, pu, fd\}) = [Little, Half]$	$m''(\{gc, pu, gs, fd\}) = Little$

Table 8: The result of considering two pieces of semiqualitative information about Jack

add in the third mass assignment to make use of the fact that if the patient has recently suffered a weight loss then we should have much belief in his having gastric cancer. This gives the result that $Bel(\{gc\}) = [Little, Much]$ while all other sets of hypotheses have belief $Little$ (Table 9).

$\otimes_\cap$	$m'''(\{gc\}) = Much$	$m'''(\{gc, pu, gs, fd\})$
$m''(\{gc, pu\}) = [Little, Half]$	$m''''(\{gc\}) = [Little, Half]$	$m''''(\{gc, pu\}) = Little$
$m''(\{gc, pu, gs\}) = Little$	$m''''(\{gc\}) = Little$	$m''''(\{gc, pu, gs\}) = Little$
$m''(\{gc, pu, fd\}) = [Little, Half]$	$m''''(\{gc\}) = [Little, Half]$	$m''''(\{gc, pu, fd\}) = Little$
$m''(\{gc, pu, gs, fd\}) = Little$	$m''''(\{gc\}) = Little$	$m''''(\{gc, pu, gs, fd\}) = Little$

Table 9: The result of considering the third piece of semiqualitative information about Jack

Clearly the method of semiqualitative linguistic labels would work equally well for larger sets of labels, and, as Table 9 illustrates, mass assignments may be made in terms of compound intervals. The use of linguistic labels works equally well with either the open or closed world assumption since (16) and (17) provide us with the machinery to perform normalisation when mass is assigned to the empty set.

4.2 Pre-compiling combinations of linguistic labels

Another way of reasoning with linguistic quantifiers is to 'pre-compile' the results of all possible assignments of linguistically quantified mass assignments. This is possible since there is a finite number of possible combinations (which is clearly not the case for the quantitative theory), and means that the process of combining mass functions is replaced by looking up the answer in a table. To demonstrate the idea, we deal with the case of mass assignments where $m(\cdot)$ assigns belief to a single subset of θ, and we have just two such assignments. The concept can of course be extended to arbitrarily large numbers of focal elements and mass assignments.

For two mass assignments m_1 and m_2 each with a single focal element F_1 and F_2, such that $m_1(F_1) = M_1$, $m_2(F_2) = M_2$, there are four sets to which the combined mass is assigned; $F_1 \cap F_2$, F_1, F_2 and Θ. These have belief masses $M_1.M_2$, $M_1.(1-M_2)$, $(1-M_1).M_2$ and $(1-M_1).(1-M_2)$ respectively, assigned to them. The set of hypotheses (which can under the open world assumption include the hypothesis $\emptyset$ indicating something outside Θ) with the largest belief is the one preferred by the evidence. With a set of linguistic labels, we can compute the most believable set of hypotheses for every possible mass assignment, and the belief in that hypothesis. The result of doing this is summarised in Table 10. Note that the Table assumes the adoption of the open world assumption since it considers $F_1 \cap F_2$ to be a valid hypothesis whether or not F_1 and F_2 have a non-empty intersection. However, if the closed world assumption is desired, it is of course possible to calculate the hypothesis which has the second largest mass in those situtations in which $F_1 \cap F_2$ is preferred since this will clearly be the preferred hypothesis when $F_1 \cap F_2 = \emptyset$. To illustrate the use of Table 10 consider what happens when Old Bull Hubbard attends the clinic. Bull's symptoms indicate that $m_1(\{pu, gs, fd\}) = Much$ and $m_2(\{gc, fd\}) = Little$ so we can say that the most likely diagnosis is $\{pu, gs, fd\}$ and that we should have between half and much belief in this.

4.3 Interval values

It is possible to generalise the approach suggested in Section 4.1 by allowing mass assignments to be made using any interval value, and carrying out the usual addition and multiplication of the application of Dempster's rule using interval arithmetic. This gives us an interval version of Dempster's rule in both its normalised (10) and unnormalised

Mass of focal elements	Preferred hypothesis
$m_1(F_1) = All$ $m_2(F_2) = All$	$Bel(F_1 \cap F_2) = All$
$m_1(F_1) = All$ $m_2(F_2) = Much$	$Bel(F_1 \cap F_2) = Much$
$m_1(F_1) = All$ $m_2(F_2) = Half$	$Bel(F_1 \cap F_2) = Half$ $Bel(F_1) = Half$
$m_1(F_1) = All$ $m_2(F_2) = Little$	$Bel(F_1) = Much$
$m_1(F_1) = All$ $m_2(F_2) = None$	$Bel(F_1) = All$
$m_1(F_1) = Much$ $m_2(F_2) = Much$	$Bel(F_1 \cap F_2) = [Half, All]$
$m_1(F_1) = Much$ $m_2(F_2) = Half$	$Bel(F_1 \cap F_2) = [Little, Half]$ $Bel(F_1) = [Little, Half]$
$m_1(F_1) = Much$ $m_2(F_2) = Little$	$Bel(F_1) = [Half, Much]$
$m_1(F_1) = Much$ $m_2(F_2) = None$	$Bel(F_1) = Much$
$m_1(F_1) = Half$ $m_2(F_2) = Half$	$Bel(F_1 \cap F_2) = [Little, Half]$ $Bel(F_1) = [Little, Half]$ $Bel(F_2) = [Little, Half]$ $Bel(\Theta) = [Little, Half]$
$m_1(F_1) = Half$ $m_2(F_2) = Little$	$Bel(F_1) = [Little, Half]$ $Bel(\Theta) = [Little, Half]$
$m_1(F_1) = Half$ $m_2(F_2) = None$	$Bel(F_1) = Half$ $Bel(\Theta) = Half$
$m_1(F_1) = Little$ $m_2(F_2) = Little$	$Bel(\Theta) = [Half, Most]$
$m_1(F_1) = Little$ $m_2(F_2) = None$	$Bel(\Theta) = Most$
$m_1(F_1) = None$ $m_2(F_2) = None$	$Bel(\Theta) = All$

Table 10: All possible combinations of linguistically weighted mass assignments with single focal elements

(3) forms. These can be written as:

$$m''(C) = \sum_{\substack{i,j \\ A_i \cap B_j = C}}^{\oplus} m(A_i) \otimes' m'(B_j) \tag{18}$$

$$m''(C) = \sum_{\substack{i,j \\ A_i \cap B_j = C}}^{\oplus} m(A_i) \otimes' m'(B_j) \oslash' \left\{ 1 \ominus' \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}}^{\oplus} m(A_i) \otimes' m'(B_j) \right\} \tag{19}$$

where $\sum^{\oplus}$ is a summation performed using interval addition. This approach makes it possible to combine interval and exact numerical information since interval arithmetic works equally well on interval and exact values, and gives more precise results than the linguistic approach because intervals don't have to be rounded to the nearest linguistic interval.

As an illusration of this, consider the diagnosis of Jack's disorders again. When the linguistic labels are discarded, then we have $m_1(\{gc, pu, fd\}) = [0.7, 1]$, $m_1(\{gc, pu, gs,$

$fd\}) = [0, 0.3]$ and $m_2(\{gc, fd\}) = 0.5$ $m_2(\{gc, pu, gs, fd\}) = 0.5$. Combining these gives Table 11, which tells us that belief in Jack having one of the disorders in $\{gc, , pu\}$ is $[0.35, 0.5]$ which is more precise information than the value $[Little, Half]$ that was computed using the linguistic labels. However, the less precise information afforded by

$\otimes_\cap$	$m(\{gc, pu, fd)\} = [0.7, 1]$	$m(\Theta) = [0, 0.3]$
$m'(\{gc, pu, gs\}) = 0.5$	$m''(\{gc, pu\}) = [0.35.0.5]$	$m''(\{gc, pu, gs\}) = [0, 0.15]$
$m'(\Theta) = 0.5$	$m''(\{gc, pu, fd\}) = [0.35, 0.5]$	$m''(\Theta) = [0, 0.15]$

Table 11: The result of considering interval information about Jack

the use of liguistic labels may be useful especially when the initial mass assignments are imprecisely known.

Having interval results introduces the problem of ranking intervals against one another to determine which set of hypotheses are more likely. One means of doing this is to use the ordering $\geq_{Q3}$ [15] which is based on fuzzy arithmetic.

$$[a, b] \geq_{Q3} [c, d] \text{ if and only if } a \geq c \text{ and } b \geq d \tag{20}$$

Other orderings are, of course, possible.

5 Absolute orders of magnitude

In Section 4.2 we performed an exhaustive analysis of the outcomes of different mass assignments which enable us to predict the most likely set of hypotheses given the interval value of the mass assignment. What we found was that in certain cases $F_1 \cap F_2$ was the most likely hypothesis, while under other conditions F_1, F_2, or even Θ were most likely. These results suggest a similar analysis in which we consider the limits on the mass assignments to determine under which conditions particular combinations of focal elements are preferred. That is, which combinations of focal elements have the, possibly equal, largest mass assigned to them. As was the case in Section 4.2 the open world assumption will be made.

5.1 Two mass assignments

We will begin the analysis by considering the combination of two mass assignments m_1 and m_2 which, as in Section 4.2 each have a single focal element F_1 and F_2 to which they assign mass M_1 and M_2 It is clear that $F_1 \cap F_2$ is one of the most credible hypotheses if belief in it is greater than or equal to the belief in F_1, F_2 or Θ. This will be the case if $M_1.M_2 \geq M_1.(1 - M_2)$, $(1 - M_1).M_2$ and $(1 - M_1).(1 - M_2)$.Thus we can say that $F_1 \cap F_2$ is one of the most credible hypotheses if:

$$M_1 \geq 0.5 \text{ and } M_2 \geq 0.5$$

while F_1 is one of the most credible hypotheses if:

$$M_1 \geq 0.5 \text{ and } M_2 \leq 0.5$$

From these and similar calculations, we can determine a set of rules that specify the result of combining a pair of mass functions based only on the relative sizes of the masses distributed.

$$\begin{array}{llll} IF\ M_1 \geq 0.5 & \text{and} & M_2 \geq 0.5 & (21)\\ \quad THEN & F_1 \cap F_2 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \geq 0.5 & \text{and} & M_2 \leq 0.5 & (22)\\ \quad THEN & F_1 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \leq 0.5 & \text{and} & M_2 \geq 0.5 & (23)\\ \quad THEN & F_2 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \leq 0.5 & \text{and} & M_2 \leq 0.5 & (24)\\ \quad THEN & \Theta & \text{is one of the preferred hypotheses} & \end{array}$$

Clearly, these rules are not mutually exclusive, and we can have a set of hypotheses which are all equally likely, and more likely than all hypotheses not in the set. Note that the use of the size of the masses relative to the landmark value 0.5 is reminiscent of absolute order of magnitude reasoning as discussed by [9], and the duality of the rules (21)–(24) and Table 10 makes this approach similar to the use of specified rules and tabular combining functions for combining evidence [4].

To illustrate the kind of reasoning that it is possible to perform with these rules, consider what happens when Cody visits the clinic. Cody's symptoms fall into two groups, one of which suggests that he is suffering from $\{fd, gs\}$ and the other of which suggests that he has $\{gs, pu\}$. While it is not possible to put precise numbers on the degree to which the symptoms suggest the sets of diseases, the physician who examines Cody is confident that the belief mass that she assigns to the set $\{fd, gs\}$ is at least 0.5, and she is even more sure that the second set of symptoms indicate $\{gs, pu\}$. Thus the first rule may be applied to obtain the fact that the most credible diagnosis of Cody's problem is that he is suffering from $\{fd, gs\} \cap \{gs, pu\} = \{gs\}$, so that the disease that it is most believable that Cody is suffering from is gallstones.

5.2 Three mass assignments

An obvious extension of the simple case analysed above is the case in which three simple mass assignments are combined. Thus to m_1 and m_2 we add m_3 which allocates M_3 to its focal element F_3 and $1 - M_3$ to Θ. The result of this additional mass assignment is to double the number of hypotheses which might be preferred, making the process of determining the conditions under which they are preferred somewhat tedious but no more difficult than for the case of two mass assignments. We have:

$$\begin{array}{llll} IF\ M_1 \geq 0.5 & M_2 \geq 0.5 & M_3 \geq 0.5 & (25)\\ \quad THEN & F_1 \cap F_2 \cap F_3 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \geq 0.5 & M_2 \geq 0.5 & M_3 \leq 0.5 & (26)\\ \quad THEN & F_1 \cap F_2 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \geq 0.5 & M_2 \leq 0.5 & M_3 \geq 0.5 & (27)\\ \quad THEN & F_1 \cap F_3 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \leq 0.5 & M_2 \geq 0.5 & M_3 \geq 0.5 & (28)\\ \quad THEN & F_2 \cap F_3 & \text{is one of the preferred hypotheses} & \\ IF\ M_1 \geq 0.5 & M_2 \leq 0.5 & M_3 \leq 0.5 & (29)\\ \quad THEN & F_1 & \text{is one of the preferred hypotheses} & \end{array}$$

$$IF\ M_1 \leq 0.5 \quad M_2 \geq 0.5 \quad M_3 \leq 0.5 \tag{30}$$
$$THEN \quad F_2 \quad \text{is one of the preferred hypotheses}$$
$$IF\ M_1 \leq 0.5 \quad M_2 \leq 0.5 \quad M_3 \geq 0.5 \tag{31}$$
$$THEN \quad F_3 \quad \text{is one of the preferred hypotheses}$$
$$IF\ M_1 \leq 0.5 \quad M_2 \leq 0.5 \quad M_3 \leq 0.5 \tag{32}$$
$$THEN \quad \Theta \quad \text{is one of the preferred hypotheses}$$

These results suggest that it is possible to predict the conditions for a certain hypothesis being preferred for any number of simple support functions [16]— that is belief functions generated by mass assignments with a single focal element. If we have n assignments, $m_1, \ldots, m_n$ each with a single focal element $F_1, \ldots, F_n$ to which it assigns, respectively, $M_1, \ldots, M_n$, then for all $1 \leq i \leq j \leq n$,

$$IF \quad M_1 \leq 0.5, \ldots, M_{i-1} \leq 0.5 \tag{33}$$
$$M_i \geq 0.5, \ldots, M_j \geq 0.5$$
$$M_{j+1} \leq 0.5, \ldots, M_n \leq 0.5$$
$$THEN \quad F_i \cap \ldots \cap F_j \quad \text{is one of the preferred hypotheses}$$
$$IF \quad M_1 \leq 0.5, \ldots, M_n \leq 0.5 \tag{34}$$
$$THEN \quad \Theta \quad \text{is one of the preferred hypotheses}$$

The discussion so far has concentrated on simple support functions since these are an important class of belief functions with a large number of applications. However, it is also interesting to analyse more complex functions, and this is the subject of the next section.

5.3 Several focal elements

The use of mass assignments that do not generate simple support functions is a little more complex largely because of the number of possible hypotheses. Consider the combination of two mass assignments m_1 and m_2 where m_1 allocates M_{11}, M_{12}, and $1 - M_{11} - M_{12}$ to F_{11}, F_{12} and Θ respectively while m_2 allocates M_{21}, M_{22}, and $1 - M_{21} - M_{22}$ to F_{21}, F_{22} and Θ. Once again determining the preferred hypotheses is tedious rather than difficult, and the result is the following set of rules:

$$IF \quad M_{11}, M_{21} \geq M_{12}, M_{22} \geq \tfrac{1}{3} \tag{35}$$
$$THEN \quad F_{11} \cap F_{21} \quad \text{is one of the preferred hypotheses.}$$
$$IF \quad M_{11}, M_{22} \geq M_{12}, M_{21} \geq \tfrac{1}{3} \tag{36}$$
$$THEN \quad F_{11} \cap F_{22} \quad \text{is one of the preferred hypotheses.}$$
$$IF \quad M_{11} \geq M_{12} \geq \tfrac{1}{3} \tag{37}$$
$$\tfrac{1}{3} \geq M_{21}, M_{22}$$
$$THEN \quad F_{11} \quad \text{is one of the preferred hypotheses.}$$
$$IF \quad M_{12}, M_{21} \geq M_{11}, M_{22} \geq \tfrac{1}{3} \tag{38}$$
$$THEN \quad F_{12} \cap F_{21} \quad \text{is one of the preferred hypotheses.}$$
$$IF \quad M_{12}, M_{22} \geq M_{11}, M_{21} \geq \tfrac{1}{3} \tag{39}$$
$$THEN \quad F_{12} \cap F_{22} \quad \text{is one of the preferred hypotheses.}$$

$$\begin{array}{llll}
IF & M_{12} \geq M_{11} \geq \frac{1}{3} & & (40)\\
 & \frac{1}{3} \geq M_{21}, M_{22} & & \\
THEN & F_{12} & \text{is one of the preferred hypotheses.} & \\
IF & M_{21} \geq M_{22} \geq \frac{1}{3} & & (41)\\
 & \frac{1}{3} \geq M_{11}, M_{12} & & \\
THEN & F_{21} & \text{is one of the preferred hypotheses.} & \\
IF & M_{22} \geq M_{21} \geq \frac{1}{3} & & (42)\\
 & \frac{1}{3} \geq M_{11}, M_{12} & & \\
THEN & F_{22} & \text{is one of the preferred hypotheses.} & \\
IF & M_{11}, M_{12}, M_{21}, M_{22} \leq \frac{1}{3} & & (43)\\
THEN & \Theta & \text{is one of the preferred hypotheses.} &
\end{array}$$

These rules may be generalised to the case of two mass functions with arbitrary numbers of focal elements. Consider we have m_1 which allocates $M_{11}, \ldots, M_{1m}$, to $F_{11}, \ldots F_{1m}$ respectively and $1-M_{11}-, \ldots -M_{1m}$ to Θ while m_2 allocates $M_{21}, \ldots, M_{2n}$, to $F_{21}, \ldots, F_{2n}$ and $1 - M_{21}-, \ldots, -M_{2n}$ to Θ. Given $1 \leq i \leq m$, and $1 \leq x \leq n$, and considering all $1 \leq j \leq m$ such that $j \neq i$, and all $1 \leq y \leq n$ such that $y \neq x$, we have:

$$\begin{array}{llll}
IF & M_{1i} \geq M_{1j} \geq \frac{1}{m} & & (44)\\
 & M_{2x} \geq M_{2y} \geq \frac{1}{n} & & \\
THEN & F_{1i} \cap F_{2x} & \text{is one of the preferred hypotheses} & \\
IF & M_{1i} \geq M_{1j} \geq \frac{1}{m} & & (45)\\
 & \frac{1}{n} \geq M_{2x} \geq M_{2y} & & \\
THEN & F_{1i} & \text{is one of the preferred hypotheses} & \\
IF & \frac{1}{m} \geq M_{1i} \geq M_{1j} & & (46)\\
 & M_{2x} \geq M_{2y} \geq \frac{1}{n} & & \\
THEN & F_{2x} & \text{is one of the preferred hypotheses} & \\
IF & \frac{1}{m} \geq M_{1i} \geq M_{1j} & & (47)\\
 & \frac{1}{n} \geq M_{2x} \geq M_{2y} & & \\
THEN & \Theta & \text{is one of the preferred hypotheses} &
\end{array}$$

And combining the results of Sections 5.2 and 5.3 will enable us to predict the outcome of many applications of Dempster's rule. It should be noted, however, that these conditions are sufficient, but not necessary. That is, there are other conditions which will result in the same preferred hypotheses. For instance, consider $M_{11} = M_{12} = 0.3$, $M_{21} = 0.8$ and $M_{22} = 0.1$. In this case the preferred hypothesis is F_{21} even though the conditions for (41) are not satisfied. The necessary conditions are not given since they may not simply be extended to the general case.

6 Summary

This paper has introduced a number of different ways in which the Dempster-Shafer theory of evidence may be applied when precise numerical weights are not given for the various pieces of evidence. Firstly, the idea of a completely qualitative theory of evidence was introduced. In this approach, all numbers are abstracted away to be replaced by the qualitative values + and 0. This strips the theory down to its bare bones, which may still prove useful in identifying which hypotheses are indicated by the evidence in situations where numerical weights may not easily be identified. By assuming that the smallest set of hypotheses is the most likely set it is possible to combine evidence to identify the most likely hypothesis. Adopting the heuristic approach of the improper linear model allows us to choose between several smallest sets. The paper then introduced a number of ways of using the Dempster-Shafer theory with limited numerical information. Firstly, the idea of using "linguistic quantifiers" in the sense of [8] was introduced, and results given for all possible combinations of a pair of simple functions whose mass assignments were taken from the set of linguistic quantifiers. Then this approach was generalised to deal with mass assignments whose values are arbitrary intervals, and it was shown how this approach may be used to combine exact and interval values. The third approach that was introduced concerned means of predicting the outcome of a combination using Dempster's rule in its unnormalised form (in the sense of predicting which hypothesis is most likely) based on how the mass of hypotheses compare to landmark values. Results are given for the combination of any number of simple support functions, and the combination of a pair of mass assignments with arbitrary numbers of focal elements.

Acknowledgement

This work was partially supported by a grant from BT. Thanks to Robin Smith and Chris Whitney at BT for help in explaining the domain for which this approach was developed, advice to the effect that it was not possible to publish anything to do with this domain, and encouragement to write the paper anyway. Thanks are also due to the Imperial Cancer Research Fund who paid for me to travel to Barcelona where a preliminary version of this paper was presented.

References

[1] Bobrow, D. *Qualitative reasoning about physical systems.* Elsevier Publishers Ltd, North-Holland, 1984.

[2] Chard, T. Qualitative probability versus quantitative probability in clinical diagnosis: a study using computer simulation, *Medical Decision Making*, **11**, 38–41, (1991).

[3] Clarke, M. R. B. Discussion of 'Belief functions' by Ph. Smets [17], in *Non-standard logics for automated reasoning*, Ed. Smets, Ph., Mamdani, E. H., Dubois, D. and Prade, H., Academic Press, London, 1988.

[4] Cohen, P. R., Shafer, G., Shenoy, P. P., Modifiable combining functions, in *Uncertainty in Artificial Intelligence 3*, L. N. Kanal, Levitt, T. S. and Lemmer, J. F. ed., Elsevier, North-Holland, 1989.

[5] Davis, R. Diagnostic reasoning based on structure and behaviour, *Artificial Intelligence*, **24** (1984) 347–410.

[6] Dawes, R. M. The robust beauty of improper linear models in decision making, in *Judgement under uncertainty: Heuristics and biases*, Ed. Kahnemann, D., Slovic, P., and Tversky, A., Cambridge University Press, Cambridge, 1982, 391–408.

[7] Dempster, A. P. - Upper and lower probabilities induced by a multi-valued mapping, *Ann. Math. Statistics* **38**, 325–339 (1967).

[8] Dubois, D., Prade, H., Godo, L., and Lopez de Mantaras, R., A symbolic approach to reasoning with linguistic quantifiers, *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, Stanford, 1992, p.74.

[9] Dubois, D., Prade, H. Fuzzy arithmetic in qualitative reasoning, in *Modelling and Control of Systems in Engineering, Quantum Mechanics, Economics and Biosciences, (Proceedings of the Bellman Continuum Workshop 1988 Sophia Antipolis)*, Ed. Blaquière, A., Springer Verlag, Berlin, 1989, p.457.

[10] Kak, A. C., Andress, K. M., Lopez-Abadia, C., Carol, M. S., and Lewis, J. R. Hierarchical evidence accumulation in the PSEIKI system, in *Uncertainty in Artificial Intelligence 5*, M. Henrion, R. D. Schachter, Kanal, L. N. and Lemmer, J. F., ed. Elsevier, North-Holland, 1990.

[11] Lindley, D. V. *Making Decisions*. Wiley and Son, London, 1985.

[12] Moore, R. E. *Interval analysis*, Prentice Hall, Englewood Cliffs, N.J. 1966.

[13] Parsons, S. Qualitative belief networks, *Proceedings of the 10th European Conference on Artificial Intelligence*, Vienna, Ed. Neumann, B., 1992.

[14] Parsons, S. Qualitative methods for reasoning under uncertainty, PhD Thesis, Department of Electronic Engineering, Queen Mary and Westfield College, 1993.

[15] Parsons, S. and Fox, J. Qualitative and interval algebras for robust reasoning with incomplete and imprecise information, in *Decision Support Systems and Qualitative Reasoning*, Ed. Travé-Massuès, L. and M. Singh, M. Elsevier Publishers Ltd. North-Holland, 1991, p.163.

[16] Shafer, G. *A mathematical theory of evidence*. Princeton University Press, 1976.

[17] Smets, Ph. Belief functions, in *Non-standard logics for automated reasoning*, Ed. Smets, Ph., Mamdani, E. H., Dubois, D. and Prade, H., Academic Press, London, 1988.

[18] Smets, Ph., and Kruse, R. The transferable belief model for belief representation, *Proceedings of the IJACI Workshop on The Mangement of Uncertainty in AI*, Chambéry, France, 1993.

[19] Strat, T. Making decisions with belief functions, *Proceedings of the 5th Workshop on Uncertainty in AI*, Windsor, (1989).

[20] Wilson, N. Some theoretical aspects of the Dempster-Shafer theory. PhD Thesis, Department of Computing and Mathematical Sciences, and Department of Hotel and Catering Management, Oxford Polytechnic, 1992.

[21] Zadeh, L. A. A mathematical theory of evidence (book review), *AI Magazine*, **5**, 3, 1984.