

UNCERTAIN REASONING AND DECISION MAKING

Qing Zhou

The Software Institute of Zhongshan University, Guangzhou, Guangdong, P.R.China, 510275

lnsq@zsulink.zsu.edu.cn

Wei Peng

Science and Technology Institute of GRC, Guangzhou, Guangdong, P.R.China, 510060

pengwei99@tom.com

Abstract In this paper we discuss uncertain reasoning and decision making. Our proposal is based on the knowledge we have and entirely formalized within the so called classical two valued logic, so it has a solid foundation. Basic notions of various items are defined formally; formulas of supporting degree for uncertain reasoning and supporting degree with safety for decision making are introduced. Evaluation of "weighted facts", which represents the different importance of facts, is clearly presented within our proposal without anything else. The relation between uncertain reasoning and decision making is discussed in detail. Examples in the paper are comprehensively exhibited, which shows that our proposal is reasonable and computer-operative.

Keywords: uncertainty, decision making, knowledge-based systems

1. Introduction

In this paper we discuss the uncertain reasoning and decision making on the base of knowledge. As our discussion is entirely built on the base of the so called classical two valued logic, it has a strong foundation.

Our purpose is to find a method similar to the way we often use when we deal with uncertain problems. So first let us take a look at what we usually do in such situations:

Suppose P is a problem, h_1, \dots, h_k are propositions about the solution of P , called the hypothesis. Our problem is to decide which of $h_i, 1 \leq i \leq k$, is true. If such a h_i can be deduced from a set K (corpus of knowledge) of propositions, which has been accepted as true, the question is solved definitely and hence P is certain; otherwise, P is uncertain. Then our problem turns to whether there is a further way to solve the question when P is uncertain.

A common way used by experts can be described as follows: 1. Gather an evidence set E as completely as possible. 2. Try to infer h_i from $K \cup E$. If (2) succeeds, the question is solved, and the uncertainty of h_i is removed absolutely. Otherwise, 3. try to estimate the degree of the possibility for h_i to be true under $K \cup E$.

From the above discussion, we can see that several concepts are basic. To determine the truth value of a proposition called a hypothesis, we need some knowledge, which is a set of formulas syntactically. The set is called a corpus of knowledge, which are regarded as true such as logical axioms, basic laws in a field, etc. When the hypothesis is uncertain, we employ some evidences, which are usually some facts (closed formulas, syntactically) on the specific problem P , to help us determine the truth value of the hypothesis. It is only required that all of the evidences are closed formulas, although they are supposed to be true for our purpose.

From the view point of uncertain reasoning, the most important part is, of course, to find a method to estimate the degree of the possibility for h_i to be true. Here we bring in an easy fact from classical logic: As indicated above, due to the limitation of our knowledge, an uncertain problem is an undecidable formula in the logical system which contains our knowledge. This means that our knowledge on the problem is incomplete, hence it has many models by the Completeness Theorem; In some of them h_i is true, while in others h_i is false, so our purpose is to determine which model is the one we want. Now since all the evidences are true in the model we want, h_i should have some common consequences with evidences if h_i is also in the model since all the consequences of evidences and h_i have to be true in the model from the definition of models. Then it is logical to believe that if h_i has more common consequences with evidences than h_j , it is more possible to be true than h_j . The following example illustrates the point:

EXAMPLE 1 Suppose we are in a murder case and we have a corpus of knowledge: If x is the murder, 1. x was in scene of the crime during certain time period; 2. x has a reason of murdering; 3. x has the weapon in the crime. Now after investigation, we found the evidences: Michael was in the scene of the crime during the time period and has a reason of murdering, John has a weapon like the one in crime. Then we have reasons to believe that Michael did the crime more likely than John as he has more suspicions than John.

The above example shows that the more same consequences of $K \cup \{h_i\}$ and $K \cup E$ we have, the more likely we believe h_i to be true. So it makes sense to consider these common consequences as important facts which support the solution h_i . Then number of common consequences of each possible solution h_i represents the strength of the supports of it from the evidences we have. This is the idea of this paper.

We will describe this process formally and on the base of the two-valued classical logic we will develop a computer-operative method of finding out the strength of believing h_i when h_i is uncertain.

An important area of applications of uncertain reasoning is decision making. Actually, all decision making problems are uncertain since all the knowledge we have now is about the past and the results of our decisions will happen in the future. This makes decision problems difficult. With the help of uncertain reasoning in this paper, we will propose a knowledge based method of decision making. The method is an analogue of "what...if..." analysis method, which is often used by experts. The following example is a simple decision making problem:

EXAMPLE 2 John wants to make an investment and he has 3 choices: in a bar, a restaurant or a super market. Investing in a bar needs small investment and most of the investment is in the dinner equipments but produces small profits. Investing in a restaurant needs large investment and most of the investment is in the dinner equipments and produces large profits. Investing in a super market needs large investment and most of the investment is in the merchandiser and produces large profits. John thinks that large profits is very important for his investment and he has to consider the possibility of changing his mind when he has to. Now if he choose to invest in a restaurant, he can have large profits and he can easily shift to a bar when he wants. So John gains more advantages from investing in restaurant than the other choice, and hence he has reasons to choose investing in a restaurant.

So in a decision making problem, we always analyze the result under the assumption that we take a specific choice (what ... if ... analysis) to try to have the most advantages and consider the easiness of changing our mind to reduce risks. This consideration will be embodied into the formalizations in the paper, too.

The rest of this paper breaks into three sections. The basic notions about uncertain reasoning are given in the first section. The supporting degree of uncertain reasoning is also introduced here. All of these are completely done within the classical two valued logic. In the second section we deal with "weighted facts". This is important since it often happens that some facts are believed more important than others during the uncertain reasoning or decision making in our real life. We show in this section that one needs nothing else to deal with weighted facts with our method, weighted facts can be represented in our method without any extra efforts. The last section takes care of the decision making problem. As indicated above, our decision making method is built on the method of uncertain reasoning discussed in the first section. A method of evaluation of "overlapped" facts is also discussed in the section, which help us to make smart decision with less risks.

2. Uncertain Reasoning and Its Supporting Degree

Now we formalize our propose within the classical logic. Let T be a fixed first order reasoning system, and L its language. In the following discussion, we will use K to be the corpus of knowledge, which is a finite set of formulas of L . Further it is assumed that K is consistent; E to be the evidence set, which is a finite set of closed formulas of L . It is also assumed that $K \cup E$ is consistent; P to be a formula of L , which is called problem. Semantically P is the problem we want to solve and K consists of those basic rules on P and E consists of some facts related to P .

Let X be a set of formulas of L . If a formula A of L can be derived from X , we use $X \vdash A$ to denote it.

Now we define that a closed formula s is uncertain if K can not deduce both s and $\neg s$. Then let $H = \{h_1, \dots, h_n\}$ be the set of hypotheses for the problem P . From the discussion in the Introduction, for all $h_i \in H$, h_i is undecidable in K , i.e. we have no way to determine whether a proposition is certainly true or false from the knowledge and the facts we have got. So we define:

DEFINITION 3 We call a set of closed formulas H a hypotheses set and for $h_i \in H$, we call h_i a hypothesis, if every sentence in H is undecidable in K .

As indicated in the Introduction, common consequences of both the hypothesis h_i and the evidence set E play important roles in the uncertain reasoning. So we define the common consequence set D_{h_i} as:

DEFINITION 4 Let h_i be a hypothesis. Then

$$D_{h_i} = \{A : A \text{ is a closed formula of } L \text{ and } K \cup E \vdash A \text{ and } K \cup \{h_i\} \vdash A\}$$

Then we say that $A, B \in D_{h_i}$ are equivalent in D_{h_i} , denoted as $A \sim B$, if $K \cup E' \vdash A$ if and only if $K \cup E' \vdash B$ for all $E' \subseteq E$. It is trivial to show that \sim is an equivalent relation defined on D_{h_i} . Now we define

DEFINITION 5 B_{h_i} be the equivalent class on D_{h_i} , i.e. $B_{h_i} = \{U : U \subseteq D_{h_i} \text{ and for any } A, B \in U, A \sim B\}$.

Then it is easy to see that B_{h_i} is finite.

From above definitions, it is obvious that for any $A \in D_{h_i}$, A is a common consequence derived by both E and h_i , so it supports h_i . Then it is logical to believe that the more elements B_{h_i} has, the more facts support h_i . From this observation, we use the word "the supporting degree " to formalize it which are given in the following definition:

DEFINITION 6 The supporting degree of a hypotheses h_i in H , $s(h_i, H)$, is a real number:

$$s(h_i, H) = \frac{|B_{h_i}|}{|\bigcup_{h_k \in H} B_{h_k}|}$$

EXAMPLE 7 Let us take a look on how to use the above system to solve the Example 1. in the Introduction.

The reasoning system we construct is:

Our language L has the following formulas::

1. $murderer(x)$, which indicates that x was the murderer semantically;
2. $inscene(x)$, which indicates that x was in scene of the crime during certain time period semantically;

3. $reason(x)$, which indicates that x had a reason of murdering;

4. $weapon(x)$, which indicates that x had the weapon in the crime;

Here x can be assigned a constant in $\{Michael, John\}$. which indicates the man with such a name, semantically.

The corpus of knowledge K consists of the following axioms:

1. $murder(x) \rightarrow inscene(x)$, which semantically indicates that if x was the murder, x was in scene of the crime during certain time period;

2. $murder(x) \rightarrow reason(x)$, which semantically indicates that if x was the murder, x had a reason of murdering;

3. $murder(x) \rightarrow weapon(x)$, which semantically indicates that if x was the murder, x had the weapon in the crime;

4. $murder(x) \rightarrow \neg murderer(x_1)$, which semantically indicates that if x was the murder, x_1 was not;

5. $murder(x) \vee murder(x_1)$, which semantically indicates that x or x_1 was the murder.

Here $x, x_1 \in \{Michael, John\}$ and $x \neq x_1$. Formula 4 and 5 show that only one was the murder among $John$ and $Michael$.

We also assume that the system has two deduction rules: 1. if $A \rightarrow B$ and A , then B ; 2. if $A \vee B$ and $\neg A$, then B .

The evidence set E consists of:

$inscene(Michael), reason(Michael), weapon(John)$

The hypotheses set:

$H = \{murder(Michael), murder(John)\}$

Now we have done our construction of the reasoning system. Let us look at what we can have with the system.

Obviously, from $K \cup \{murder(Michael)\}$ we can have from the rules and the corpus of knowledge that: $inscene(Michael), reason(Michael)$ and $weapon(Michael)$.

And from $K \cup E$ we can have: $inscene(Michael), reason(Michael)$

So $|B_{murder(Michael)}| = |\{inscene(Michael), reason(Michael)\}| = 2$

From $K \cup \{murderer(John)\}$ we can deduce the following,

$inscene(John), reason(John)$ and $weapon(John)$.

And from $K \cup E$ we can have: $weapon(John)$

So $|B_{murder(John)}| = |\{weapon(John)\}| = 1$

Then we can get $s(\text{murder}(\text{Michael}), H) = 2/3 > s(\text{murder}(\text{John}), H) = 1/3$.

Therefore from our system we have that Michael did the murder more likely than John did. This is exact what we should believe from the evidences we have, since Michael has more suspicions than John has.

3. Evaluations for "Weighted Facts"

It seems that we can get the best solution by comparing $|B_{h_i}|$ for all $h_i \in H$. But in reality it is not so simple. It often happens that some facts are supposed to be more important than others. For example, when John wants to invest in a project he thinks that the profit is more important than other factors. So, our uncertain reasoning system should reflect the difference in numbers. But in the supporting degree introduced in the last section every common consequence of E and the hypothesis is used only once and hence every common consequence plays the same role as others. So the importance of some evidences seem to be ignored.

Now we show that "weighted facts" can be easily represented in our system without any extra efforts.

Let S be a first order reasoning system. Suppose F is a fact which is assumed to be more important than others. We do not know whether F can be derived from our system or not. But what we want is as soon as F is derived we count it twice. Can this be represented in our system? The answer is Yes.

To this end, we add two new predicates, *dummy* and *dummy1*, to the system S . Then we add rules $F \wedge \text{dummy} \rightarrow \text{dummy1}$ and $h_i \rightarrow \text{dummy1}$ for all $h_i \in H$ to the corpus of knowledge K , and add *dummy* to the evidence set E . We use S_1 to denote the resulting system. Then the following results show that with the supporting degree introduced in the last section, B_{h_i} will be the same in both S and S_1 if F is not in D_{h_i} and B_{h_i} in S_1 will be 1 more than in S if F is in D_{h_i} .

THEOREM 8 *In S_1 , $F \in D_{h_i}$ if and only if $\text{dummy1} \in D_{h_i}$.*

Now we use $B_{h_i}^S$ and $B_{h_i}^{S_1}$ to B_{h_i} in S and S_1 , respectively. Then from the theorem we immediately get:

COROLLARY 9 *If $F \in D_{h_i}$ then $|B_{h_i}^{S_1}| = |B_{h_i}^S| + 1$.*

This theorem shows that "weighted facts", i.e. the importance of facts, can be reflected by the increments in number.

4. Uncertain Reasoning in Decision Making

As well known, decision making is an important application area of uncertain reasoning. In our life, we always face decision problems which can be

described as follows: We have a problem P , which could be our goal, or a difficulty in our life. To reach our goal or overcome the difficulty we have a few choices c_1, \dots, c_n which are supposed to be the possible solutions. The decision problem is to find out the best solution from the choice set $\{c_1, \dots, c_n\}$.

As indicated in the Introduction, the main difficulty of making a choice is that there always is uncertainty on what will exactly happen since we can only have knowledge on the past and the result of our decision is about future.

Now suppose we have a choices set $\{c_1, \dots, c_n\}$ and we have to make a decision. As there are always some uncertainties on what will exactly happen, decision making problems are uncertain and we can use uncertain reasoning for such problems. So for the discussion of decision making we should first define the notion "choice" formally.

In the real life, before we choose one from the choice set we always consider what results it will bring to us, i.e. what...if... analysis. Different choices will make different resulting sets. Among these different resulting sets, if one contains more advantages than others, we have reasons to take the one which produces this set. These results are usually able to be derived from the knowledge we have when a choice c_i is assumed and they can be supported by evidences we gathered, that is, they can be deduced from both $K \cup E$ and $K \cup \{c_i\}$.

In terms of Model Theory the decision making problem is to choose the best model, then if some evidences are in the same model as c_i they should have something in common with c_i , i.e. they and c_i should have common consequences. From this observation, we say that a subset of E favorites c_i if it has a common consequence with c_i . (Due to the length limitation, we only consider those evidences which give advantages in this paper, for the case of disadvantages, one can easily have the formulation from the idea of this paper.)

Now let $C = \{c_1, \dots, c_n\}$ be the set of choices for the problem P . From the discussion above, for all $c_i \in C$, c_i is undecidable in K , i.e. we have no way to determine whether a choice is certainly true or false from the knowledge and the facts we have got. So we define:

DEFINITION 10 *We call a set of closed formulas C a choices set and for $c_i \in C$, we call c_i a choice, if every sentence in C is uncertain.*

Comparing this definition with the one for hypothesis in the first section, we can easily see similarity of them. So it is reasonable to use the formulas in the first section and choose the choice from the choice set which has the largest supporting degree. Actually, from the discussions above, this is what we always do when we are in the cases that we have to make decisions in some simple situations, but in many cases, we will miss some important factors if we only use that supporting degree.

In the real world, some common results might be derived by more than one choices. In the example in the Introduction, the dinner equipments are needed

for both bars and restaurants, so it will be easier for John to change investment in a restaurant to a bar when he has to. This is certainly an important factor he has to consider when he is making decision on his investment. Of course, this is also an important factor we have to take account of when we consider decision making problems. We call such a fact the "overlapped problem" in the field of decision making. Described formally, for two choice $c_i, c_j \in C$, we say that they are overlapped if they have same consequences, i.e. $B_{c_i} \cap B_{c_j} \neq \emptyset$.

Obviously different decisions between two overlapped choices c_i and c_j will make different results. Our problem is how to evaluate the common results $B_{c_i} \cap B_{c_j}$ derived by both of them and use it to help us make decision. We think that such results play different roles for c_i and c_j . Suppose our choice is c_i . Then we will have all the advantages in B_{c_i} , but also have partial advantages in B_{c_j} , i.e. the $B_{c_i} \cap B_{c_j}$ part in B_{c_j} . Therefore, to give a measurement to this partial advantage for c_j when c_i is chosen it is logical to consider the number $\frac{|B_{c_i} \cap B_{c_j}|}{|B_{c_j}|}$, which we call "the relevancy coefficient":

DEFINITION 11 For two choices c_i and c_j , the relevancy coefficient of c_i to c_j in C , $rc(c_i, c_j, C)$, is a real number:

$$rc(c_i, c_j, C) = \frac{|B_{c_i} \cap B_{c_j}|}{|B_{c_j}|}$$

It is also obvious that not only the supporting degree of a choice but also the common effects of all choices in the choices set should be taken account when we want to make a smart decision and consider the situations that we may have to change our mind later. The number rc could measure the effects of such common consequences in each of the choices in our decision making procedure. We will use it in the following discussion.

Because $B_{c_i} \cap B_{c_j}$ contains those common consequences which support both c_i and c_j and they can still work for c_j even when c_i is chosen and c_j is not, $rc(c_i, c_j, C)$ measures the advantages for c_j , so it represents the possibility of changing c_i to c_j . The relevancy coefficient represents the degree of the easiness of changing from one choice to another. The larger $rc(c_i, c_j, C)$ is, the easier we change c_i to c_j . It is very important and necessary when we are in the decision-making cases in which we have often to change our mind due to situations. We define $rc(c_i, c_j, C) = 0$ if $|B_{c_j}| = 0$. Then it is easy to have the formula for taking care of such consideration, which we call "the supporting degree with safety":

DEFINITION 12 For any choice $c_i \in C$, the supporting degree with safety $ssd(c_i, C)$ of c_i is defined to be:

$$ssd(c_i, C) = s(c_i, C) + \sum_{c_j \in C, c_j \neq c_i} rc(c_i, c_j, C) \times s(c_j, C)$$

LEMMA 13 $ssd(c_i, C) \leq |C|$

This definition gives a balanced consideration which is often needed in reality while making decisions. Once we have the safe supporting degree of c_i for every c_i in the choice set C , we can easily make our decision: It is the one which has the largest safe supporting degree as it is the most possible solution for the problem and has the least risk from the knowledge we have now. Of course in many cases we do not have chances to change our plan then we can omit the safe degrees and take one with the highest supporting degree instead.

With our purpose, some emotional factors, such as one's interests or hobbies, are not taken in our proposal because they are very hard to be represented completely within the classical logic.

EXAMPLE 14 *Now we consider the Example 2. in the Introduction.*

To solve the problem we have to build up a reasoning system for it first.

We define a first order language L as follows:

We will use the symbols BAR , RST and SM to denote a bar, a restaurant or a super market respectively and the letter x to denote the variables representing a project which John will invest to. Obviously $x \in \{BAR, RST, SM\}$.

Also L contains the following predicate:

$Invest(x)$: a predicate which represents that John invests to the project x ;

$SI(x)$: a predicate which represents that the project x needs small investment;

$LP(x)$: a predicate which represents that the project x produces large profits;

FA : a predicate which represents most of the investment is in the fixed assets(dinner equipments);

$dummy, dummy1$: two predicate symbols without any sense.

Then our problem turns out to determine which can satisfy John's needs more in the choice set $C = \{c_1 = Invest(BAR), c_2 = Invest(RST), c_3 = Invest(SM)\}$.

Now we can construct our corpus of knowledge K as following:

$Invest(BAR) \rightarrow SI(BAR)$: which means that the investment in a bar needs small investment;

$Invest(BAR) \rightarrow FA$: which means that most of the investments in a bar is for fixed assets(dinner equipment);

$Invest(RST) \rightarrow FA$: which means that most of the investments in a restaurant is for fixed assets(dinner equipment);

$Invest(RST) \rightarrow LP(RST)$: which means that the investment in a restaurant produces large profits;

$Invest(SM) \rightarrow LP(SM)$: which means that of the investment in a restaurant produces large profits ;

$LP(x) \wedge dummy \rightarrow dummy1$: which mean that large profits of an investment is important;

$Invest(Bar) \rightarrow dummy1, Invest(Bar) \rightarrow dummy1, Invest(SM) \rightarrow dummy1$: meaningless, only for the completeness as in the Section 2.

Then our evidence set E consists of the facts: $SI(BAR), LP(RST), LP(SM), FA, dummy$.

So from $K \cup \{c_1\}$ and $K \cup E$ we can get $B_{c_1} = \{[SI(BAR)], [FA]\}$; From $K \cup \{c_2\}$ and $K \cup E$ we can get $B_{c_2} = \{[LP(RST)], [FA], [dummy1]\}$; From $K \cup \{c_3\}$ and $K \cup E$ we can get $B_{c_3} = \{[LP(SM)], [dummy1]\}$. So

$$|\bigcup_{c_k \in C} B_{c_k}| = |\{[SI(BAR)], [FA], [LP(RST)], [LP(SM)], [dummy1]\}| =$$

5. Then we can get the supporting degrees as:

$$s(B_{c_1}, C) = \frac{2}{5}, s(B_{c_2}, C) = \frac{3}{5}, s(B_{c_3}, C) = \frac{2}{5}$$

The relevancy coefficients are: $rc(B_{c_1}, B_{c_2}, C) = \frac{1}{3}, rc(B_{c_2}, B_{c_1}, C) = \frac{1}{2}$ and $rc(B_{c_3}, B_{c_1}, C) = rc(B_{c_1}, B_{c_3}, C) = rc(B_{c_2}, B_{c_3}, C) = rc(B_{c_3}, B_{c_2}, C) = 0$

So we can get the supporting degrees with safety for such choices: $ssd(B_{c_1}, C) = \frac{2}{5} + \frac{2}{5} \times \frac{1}{2} = \frac{3}{5}, ssd(B_{c_2}, C) = \frac{3}{5} + \frac{3}{5} \times \frac{1}{3} = \frac{4}{5}$, and $ssd(B_{c_3}, C) = \frac{2}{5}$

Thus John has more reasons to invest in a restaurant. If we do not consider the importance of large profits of investment, investing in a restaurant gains the same advantages as that in a bar. But when the importance is considered, we get different results. It shows that our purpose can work for such problems.

References

- Doyle, J., Maintenance and Belief Revision, in Belief Revision, Ed. by P. Gardenfors, Cambridge University Press, 1992.
- Kanal, L., Lemmer, J.F., Uncertainty in Artificial Intelligence, North-Holland, Amsterdam, New York, Oxford, Tokyo, 1986.
- Russell, S., Norvig, P. Artificial Intelligence - A Modern Approach, Chapter 16 "Making Simple Decisions", Prentice Hall, 1995.
- George J. Klir, Richard M. Smith, On Measuring Uncertainty and Uncertainty-Based Information: Recent Developments, Annals of Mathematics and Artificial Intelligence, Volume 32, Issue 1-4, 2001
- Guerlain, S., Brown, D. and Mastrangelo, C. (2000). Intelligent decision support systems. Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, Nashville, TN, pp. 1934-1938.
- Zhou, Q., Peng, W., KURS: a Knowledge-based Uncertain Reasoning System, Artificial Intelligence & Its Applications, Proceedings of the ICAAI'03, 2003
- Chansarkar, S.R., Expert System Shell: A Case Study, Artificial Intelligence & Its Applications, Proceedings of the ICAAI'03, 2003