A first order logic for reasoning under uncertainty using rough sets

Simon Parsons
Advanced Computation Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom.

Miroslav Kubat
Department of Medical Informatics, Institute of Biomedical Engineering, Graz University of Technology, Brockmangasse 41, A-8010, Graz, Austria.

Abstract

Reasoning with uncertain information is a problem of key importance when dealing with knowledge from real situations. Obtaining the precise numbers required by many uncertainty handling formalisms can be a problem when building real systems. The theory of rough sets allows us to handle uncertainty without the need for precise numbers, and so has some advantages in such situations. We develop a set of symbolic truth values based upon rough sets which may be used to augment predicate logic, and provide methods for combining these truth values so that they may be propagated when augmented logic formulæ are used in automated reasoning.

Keywords: uncertainty, rough sets, possible worlds, rules of inference, resolution principle, theorem proving.

1. Introduction

One fundamental requirement of all intelligent systems is that they have a means of representing knowledge, and a means of using this knowledge in order to act intelligently. This knowledge can range from the thousands of production rules that the expert system R1/XCON [23] uses to configure computers for Digital, to the knowledge of how to conduct a whimsical conversation employed by the winning entry in the First Annual Loebner Prize Competition[1] [13]. In addition to representing knowledge, intelligent systems need to be able to use that knowledge to reason from what they know, or are told, about situations in order to make decisions about what actions they should take in response to those situations. Thus an intelligent medical system must reason with its knowledge to determine the correct treatment for a patient, and intelligent fire fighting equipment [4] must reason with its knowledge to determine the best plan for putting out a fire.

Many early intelligent systems favoured the use of logic or production rules as a means of representing knowledge, since these formalisms made it easy to express the knowledge of experts in the domain in which the intelligent system was intended to operate, and systems built using these formalisms performed well when the knowledge that they were used to encode was complete and certain. However, many real situations may not be completely and certainly described, so that many systems designed to operate in the real world must, perforce, be capable of dealing with uncertain information, that is information whose certainty may not be completely established. This is a direct consequence of the complexity of the real world and the finite size of the

[1] Which provides a yearly forum in which machines attempt to fool humans in progressively more difficult versions of the Turing test.

knowledge base that any intelligent system has at its disposal. Now, uncertain information cannot be expressed using production rules or logic alone— other, non-standard, logics are required, and it is one particular non-standard logic that is the subject of this paper. Uncertain information is often represented by attaching some numerical estimate to the information in question to express the fact that it is not known to be true and most people agree that this seems an obvious and natural way to express uncertainty. The numerical information is propagated according to the axioms of some more or less well established mathematical theory which seeks to guarantee that the final degree of certainty accorded to the answer to a query is exactly that determined by the degrees of the relevant facts in the knowledge base. There are many such formalisms, all with their unique advantages and disadvantages. What we propose in this paper is something rather different, that is a logic which attaches non-numerical values to facts and uses these to express the certainty, or lack of certainty, of those facts. The advantage of such a symbolic approach is its simplicity compared with the mathematical complexities of some of the numerical formalisms, and the fact that it does not use numbers since these can at times be unintuitive. However, unlike most other symbolic approaches such as the theory of endorsements [5], these symbolic values have a solid mathematical foundation[2].

Section 2 begins our discussion with a brief survey of formalisms for reasoning under uncertainty— a necessary preliminary to proffering our approach. Then in Section 3 we introduce the basic concepts of rough set theory, including the key notions of core and envelope, along with the related idea of flou sets, and discuss the use of rough sets in knowledge representation. In Section 4, we show how the theory of rough sets may be extended to cope with computing the core and envelope of logical combinations of objects and describe a first order logic that may be quantified with the rough sets. Following from this Section 5 introduces set of symbolic quantifiers for predicate logic and provides rough versions of classical inference rules such as modus ponens. Section 6 extends these results with a rough resolution rule and proves some properties of the results of resolution. Section 7 concludes.

2. A brief survey of uncertainty handling formalisms

As mentioned above, the problem with expressing uncertain information in logic [36], or using production rules [7] is that in such formalisms facts and rules can only be true or false. Thus it is possible to express the fact that "it is raining" is true, and that the rule "if it is raining, then Simon must wear a hat" is true and to use these two pieces of information to decide that "Simon must wear a hat". However, it is not possible to express the fact that "There is a 40% chance of rain in the afternoon" or to say that the rule "if it will rain in the afternoon then Simon must take a hat to work" is correct some of the time, and so it is impossible to reason about whether Simon should take a hat to work. A natural response to this problem is to try and break up the space of possible values of facts so that there are values other than true or false. A typical response is to use the $[0, 1]$ interval, with false represented by 0 and true by 1, allowing any other value in between. This allows any value to be assigned to any fact, but raises two new problems— how should values be combined when reasoning is performed, and what do the values mean? There are any number of answers to these questions, and each pair of answers, in effect, provides a different system for reasoning under uncertainty.

The oldest formalism for reasoning under uncertainty is probability theory, which, according to Shafer [39] was founded by Pascal and Fermat in an exchange of letters in 1654. Over the subsequent 340 years the theory has been well defined and its capabilities extensively explored, so that the rules for propagating values are established without question, and may be found in any textbook on probability (for instance [22]). It is,

[2] For the purposes of this discussion we will not be considering nonmonotonic logics, since these are applicable to reasoning under conditions of incomplete information rather than with uncertainty.

however, less clear what the numbers mean. Some advance the view that the probability of an event is a measure of its frequency of occurence in the long term, while others insist that the probability is a subjective measure of one's belief in the occurence of the event. There are arguments for and against both positions, and a good summary of these may be found in [12]. An important point about the use of probability theory is that it is not truth functional. That is, it is not possible to precisely establish the probability of a combination of two or more propositions from the probabilities of the propositions alone. This is in direct contrast to logic which is truth functional. The result of this difference between logic and probability is that attaching probability values to logical propositions is not very fruitful. For instance, in the best known of these hybrids, Nilsson's [26] probabilistic logic, when two propositions are combined together it is only possible to establish the bounds on the probability of the combination, and these tend to $[0, 1]$ very quickly.

There are two possible solutions to this problem— to change the representation from logic to something that is more natural from the point of view of probability theory, or to use a numerical measure of uncertainty that is truth functional. The first approach is that taken by Pearl [34] and Lauritzen and Spiegelhalter [20] in their seminal work on probabilistic networks. Probabilistic networks explicitly record the conditional independencies between the probabilities of propositions, so that when the probability of a particular proposition is required, it is clear which other probabilities must be taken into account. This provides a means of establishing the precise probabilities of interesting propositions which is efficient in practice, despite being intractable in the general case [6]. This approach has become very propular, but, despite much recent work, does not have the flexibility of a first order formalism. The second approach, that of truth functional values, was adopted by Bundy [2] who proposed using "incidences" as a measure of uncertainty. In Bundy's approach, incidences are associated with propositions, are combined truth functionally, and may be used to establish the probabilities of the propositions within tighter bounds than are otherwise possible.

There are a vast number of other numerical methods for reasoning under uncertainty which have been proposed as alternatives to probability theory, and we will discuss a few of them. However, this is far from being an exhaustive survey, and the interested reader is directed to [21] for other methods, and more detail on the methods mentioned here.

Firstly, an honourable mention must go to certainty factors [1] which have been widely used in production rule systems, especially MYCIN and its sucessors. Certainty factors aimed to release expert system developers from the restrictive fact that in probability theory the negation of a proposition is supported by all the probability that is not assigned to that proposition. The method worked well, but was eventually discredited by Heckerman [18] who showed that certainty factors were just another application of probability theory, but one which entailed a lot of unrealistic assumptions. Another widely used formalism is that of fuzzy logic, which quantifies propositions using fuzzy sets [46], sets whose members need not be unambiguously in or out of the set (as in classical set theory), but may be in the set to some degree. Fuzzy logic, despite the philosphical attackes made on it by, for instance Haack [17], has proved immensely useful in automatic control and is now in wide use in industry in systems from rice cookers to turbines.

Fuzzy sets also provide the basis for another formalism that has provoked a lot of academic interest— possibility theory [45]. Possibility theory is based around the possibility measure, which quantifies the degree to which a proposition might have a particular property. For instance, the possibility that x is tall is a measure of the degree to which x may be tall, and in the simplest case this comes down to the degree to which x is in the fuzzy set of tall things. From this simple basis has grown a theory which largely parallels probability theory, with different methods for normalising distributions of numerical values and manipulating them. In fact, one of the strengths of possibility theory is that it has a whole family of different combinaton operations, each member of which assumes slightly different nuances of meaning of the values. However, from the point of view of providing a means of extending logic to handle uncertain information, possibility theory has gone beyond probability theory. Possibilistic logic [10], [11] quantifies propositions with possibility values, and their dual necessity values, and provides a means for combining the values in all the situations that will be encountered in logical reasoning. The result is a first order, quantified truth functional logic, which is perfect for many instances of reasoning under uncertainty. This logic is very similar in many ways to the one introduced in this paper.

It is also worth mentioning the Dempster-Shafer theory [39] since this has many similarities with the theory of rough sets on which this paper is based [43] [29]. The theory addresses the problem that in probability theory one must attach a measure of uncertainty to every proposition of interest by allowing measures to be attached to sets of propositions. It has been extended to permit logical inference by McLeish [24] and Saffiotti [37].

All of the approaches mentioned above use numerical quanitifiers. However, despite the many successes that such methods have provided, this may well not be the best approach. To begin with there is a wealth of evidence, of which [19] is a good example, to support the view that human beings, who after all compile the knowledge bases that we are interested in, are not very good at dealing with numbers. They are not good at estimating them, even when they are very familiar with the theories that underlie them, and they are very bad at reasoning with them. Since human beings are very adept at reasoning under uncertainty, this suggests that there may be a lot to gain by investigating non-numerical techniques. Another problem with strictly numerical methods is their over-precision. As Wellman [42] points out, crafting a system with a particular set of numbers that cover particular instances (which has to be done in many formalisms) means that it is then invalid to apply the system to other cases because the numbers would be incorrect, despite the fact that the underlying principles are applicable. To combat this problem robust methods based on techniques from qualitative reasoning have been developed [42], [28].

There are also many papers, including [1] and [2], which demonstrate that simple symbolic techniques which essentially count the numbers of reasons for and against a hypothesis are as good, in many cases, at dealing with uncertainty as complex numerical ones. This simple approach was refined by Cohen [5] with his theory of endorsements which attempted to give some structure to the symbolic reasons, so that it was possible to capture more subtle details of their strength. At the same time Fox was attempting to logically capture the meaning of linguistic terms expressing uncertainty, so that, for instance, a hypothesis is “possible” if there is no reason to exclude it [16]. A natural successor to these ideas is the concept of argumentation [15] [21] in which the intuitive idea of establishing the certainty of a hypothesis by weighing all the reasons for and against it is given a formal treatment. Our aim in this paper is to take the germ of these symbolic ideas and combine them with logic, developing a simple truth-functional, quantified logic that has symbolic quantifiers with intuitive meanings which are mathematically well-behaved since they are soundly based on rough set theory.

3. Rough set theory

Rough sets, originally introduced by Pawlak [30], have been further developed and applied to a number of problems by various authors [14], [27], [31], [32], [33] and [44]. Here we discuss the basic ideas behind the theory before relating them to knowledge representation using logic and the handling of uncertainty.

3.1 Basic concepts of rough sets

Consider a set of elementary concepts or attributes $\mathbf{A} = \{A_1,...,A_n\}$ such as {green, blue, egg, ball, cube}. These concepts are the language which is available to describe the set of objects $\mathbf{E} = \{E_1,...,E_m\}$ in which we are interested. For instance, $\mathbf{E}$ may be the set {green cube, blue cube, blue ball, egg}. Now, if we base our description of the E_j on the A_i alone, it could well be the case that some of the E_j are indistinguishable since the values that distinguish them are not present in $\mathbf{A}$. For instance using the set of concepts {green, blue, egg, ball, cube} we cannot tell a blue prism from a blue cone since there is nothing in our set of descriptive terms to distinguish prism from cone.

Thus the use of a finite set $\mathbf{A}$ implies the existence of an equivalence relation $\approx$ such that $E_j \approx E_k$, and E_j and E_k cannot be distinguished from one another for a given $\mathbf{A}$ if for every i, A_i is an attribute of E_j if and only if it is an attribute of E_k. Thus there is a partition on $\mathbf{E}$:

$$\mathbf{P} = \{P_1,...,P_r\} \quad \text{where} \quad \bigcup P_i = \mathbf{E} \quad \text{and} \quad P_i \cap P_j = \emptyset \quad \text{for } i \neq j = 1,...,r \qquad (1)$$

and each P_i is an equivalence class. Thus in our example where $\mathbf{A}$ = {green, blue, egg, ball, cube} and $\mathbf{E}$ = {green cube, blue cube, blue ball, egg}, P = {{green cube}, {blue cube, blue ball}, {egg}}. Let $T \in \mathbf{T} = \{T_1,...,T_p\}$ be an object, whose attributes are $T_A \subseteq \mathbf{A}$, that we wish to describe in terms of the partitioned set of attributes E_i. Let:

$$T^c(\mathbf{P}, \mathbf{E}) = \{e: e \in P_i, P_i \subset T\} \qquad (2)$$

$$T^e(\mathbf{P}, \mathbf{E}) = \{e: e \in P_i, P_i \cap T \neq \emptyset\} \qquad (3)$$

where $T^c(\mathbf{P}, \mathbf{E})$ is the core[3] of T based on $\mathbf{E}$ and $\mathbf{P}$, the set of all equivalent objects in $\mathbf{E}$ all of whose attributes are possessed by T, and $T^e(\mathbf{P}, \mathbf{E})$ is the envelope of T based on $\mathbf{E}$ and $\mathbf{P}$, the set of all equivalent objects in $\mathbf{E}$ at least one of whose attributes is possessed by T. So for our example if T = {blue egg}, then T^c = {egg}, T^e = {egg, blue cube, blue ball}. The pair $[T^c(\mathbf{P}, \mathbf{E}), T^e(\mathbf{P}, \mathbf{E})]$ is a *rough set*. The boundary of T is the set of equivalent objects in $\mathbf{E}$ in its envelope that are not in its core:

$$T^b(\mathbf{P}, \mathbf{E}) = T^e(\mathbf{P}, \mathbf{E}) - T^c(\mathbf{P}, \mathbf{E}) \qquad (4)$$

and the indifference set is the set of all equivalent objects in $\mathbf{E}$ which are not involved in the description of T.

$$T^i(\mathbf{P}, \mathbf{E}) = \mathbf{P} - T^e(\mathbf{P}, \mathbf{E}) \qquad (5)$$

Let the set of all rough sets that may be defined using $\mathbf{E}$ partitioned as $\mathbf{P}$ be denoted by $\mathbf{R}$. Consider $R = [R^c, R^e], R' = [R'^c, R'^e] \in \mathbf{R}$. It is simple to show that the following set theoretic relations hold where the symbol '~' stands for complement[4]:

[3]In Pawlak's original work on rough sets the envelope and core were named 'upper approximation' and 'lower approximation' respectively.

[4]Note that $\sim R^c$ is $\sim(R^c)$, the complement of the core of R and should be distinguished from $(\sim R)^c$, the core of the complement of R

$$
\begin{array}{llllll}
(R \cup R')^{c} & \supseteq & R^{c} \cup R'^{c} & (R \cup R')^{e} & = & R^{e} \cup R'^{e} \\
(R \cap R')^{c} & = & R^{c} \cap R'^{c} & (R \cap R')^{e} & \subseteq & R^{e} \cap R'^{e} \\
(\sim R)^{c} & = & \sim(R^{e}) & (\sim R)^{e} & = & \sim(R^{c})
\end{array} \qquad (6)
$$

It is also straightforward to show that $\mathbf{R}$ is a lattice [35] with maximal element $(\mathbf{E}, \mathbf{E})$, and minimal element $(\emptyset, \emptyset)$:

$$
\begin{array}{ll}
R \cup R' = R' \cup R, & R \cap R' = R' \cap R \\
(R \cup R') \cup R'' = R \cup (R' \cup R''), & (R \cap R') \cap R'' = R \cap (R' \cap R'') \\
R \cup R = R, & R \cap R = R \\
R \cup (R \cap R') = R \cap (R \cup R') & = R
\end{array} \qquad (7)
$$

The central idea of this paper is that rough sets may be used to handle uncertainty, providing a new method of dealing with an old problem. In Section 3.3 we discuss how rough sets may be used to capture uncertainty. First we introduce the related idea of flou sets.

3.2 Flou sets

For completeness it is worth briefly considering the notion of a flou set since it is closely related to that of a rough set. A flou set [25] is a pair of sets. For a some $\mathbf{U}$, we have:

$$
F = (\mathbf{E}, \mathbf{F}) \quad \text{where} \quad \mathbf{E} \subseteq \mathbf{F}; \quad \mathbf{E}, \mathbf{F} \subseteq \mathbf{U} \qquad (8)
$$

The set of all flou sets is denoted by $FL(\mathbf{U})$, and has a natural ordering, so that for $F, F' \in FL(\mathbf{U})$, $F = (\mathbf{E}, \mathbf{F})$ and $F' = (\mathbf{E}', \mathbf{F}')$:

$$
F \subseteq F' \quad \Leftrightarrow \quad \mathbf{E} \subseteq \mathbf{E}' \text{ and } \mathbf{F} \subseteq \mathbf{F}' \qquad (9)
$$

For $F \in FL(\mathbf{U})$, $F = (\mathbf{E}, \mathbf{F})$, $\mathbf{E}$ is the sure region, $\mathbf{F}$ the maximum region, and $\mathbf{F}/\mathbf{E}$ the flou region. Set operations on flou sets are defined as follows. For $F, F' \in FL(\mathbf{U})$, $F = (\mathbf{E}, \mathbf{F})$ and $F = (\mathbf{E}', \mathbf{F}')$:

$$
\begin{array}{lll}
F \cup F' & = & (\mathbf{E} \cup \mathbf{E}', \mathbf{F} \cup \mathbf{F}') \\
F \cap F' & = & (\mathbf{E} \cap \mathbf{E}', \mathbf{F} \cap \mathbf{F}') \\
\sim F & = & (\sim \mathbf{F}, \sim \mathbf{E})
\end{array} \qquad (10)
$$

It is simple to show that $FL(U)$ is a completly distributive lattice [35], so that, for $F, F', F'' \in FL(U)$:

$$
\begin{array}{ll}
F \cup F' = F' \cup F, & F \cap F' = F' \cap F \\
(F \cup F') \cup F'' = F \cup (F' \cup F''), & (F \cap F') \cap F'' = F \cap (F' \cap F'') \\
F \cup F = F, & F \cap F = F \\
F \cup (F \cap F') = F \cap (F \cup F') & = F \\
F \cup (\bigcap_{i \in I} F_i) = \bigcap_{i \in I} (F \cup F_i), & F \cap (\bigcup_{i \in I} F_i) = \bigcup_{i \in I} (F \cap F_i)
\end{array} \qquad (11)
$$

The maximal element of the lattice $FL(\mathbf{U})$ is clearly $(\mathbf{U}, \mathbf{U})$, and the minimal element $(\emptyset, \emptyset)$. Comparing the results in the last two sections, it is clear that the properties of rough sets and flou sets are closely related. Both are pairs of sets, one of which is a subset of the other as shown by (2), (3) and (8). Comparing (6) and (10) shows that set theoretic operations on rough and flou sets are broadly similar, but that flou sets allow the exact determination of the outcome of operations in all cases whereas rough sets at times only allow limits to be set on the outcomes.

The fact that flou sets always give exactly known combinations under set theoretic operations means that a set of flou sets is a distributive lattice (11) while a set of rough sets forms a non-distributive lattice (7). In this sense flou sets are a special case of rough sets which are well behaved under union and intersection. Thus all the results that are derived below could be re-derived using flou sets— the only difference being that in every result that is bounded by $\supseteq$ or $\subseteq$ the $\supseteq$ or $\subseteq$ would become equality.

3.3 Rough sets and knowledge representation

The essential idea of using rough sets for knowledge representation is that the basic item, the object whose core and envelope are manipulated, is a logical fact rather than the object of Pawlak's original work. We will consider the use of rough sets in knowledge representation from the stand-point of a knowledge base used to reason about a physical system. A knowledge item is the basic unit from which a knowledge base is constructed. In this case we will consider knowledge items to be logical propositions that may be associated with truth values, but in general they may be any atomic unit of any knowledge representation scheme.

We have a set of knowledge items $\mathbf{E} = \{E_1,\ldots, E_m\}$ about the system we are interested in, for instance the symptoms which require an explanation such as {headaches, fever, spots, rash}. We also have a set of hypotheses $\mathbf{T} = \{T_1,\ldots, T_p\}$ such as {measles, tuberculosis, whooping cough} which may be related to the facts $\mathbf{E}$ by observation. For instance we may decide that $E_i \wedge E_j \rightarrow T_k$, a possible medical gloss being "fever and spots implies measles". Now, rather than quantify the underlying uncertainty in this rule by attaching a number to it, we turn to the set of underlying concepts $\mathbf{A} = \{A_1,\ldots,A_n\}$. Often we don't know what these are, and they rarely map one to one onto the E_j and T_k, the facts that the domain in question forces us to deal with. All we know is that the A_i map onto the E_j so that one or more A_i relate to each E_j. Similarly, one or more concept relates to each T_k.

Because the concepts A_i are the fundamental concepts of the domain, there is an exact relationship between sets of A_i for every piece of information about that domain. For example, the rule relating to the symptoms of measles may be $A_1 \wedge A_2 \wedge A_3 \rightarrow A_4 \wedge A_5$. However, because the things that we observe do not map cleanly onto these concepts, and we are constrained to make statements in terms of the things that we observe, we can only make statements that are approximately true. Thus we say $E_i \wedge E_j \rightarrow T_k$ which for $E_i = A_1$, $E_j = A_3 \wedge A_6$ and $T_k = A_4$ amounts to $A_1 \wedge A_3 \wedge A_6 \rightarrow A_4$ which is clearly not quite right. Thus the use of observable facts to describe systems means that in general we can only approximate the underlying concepts. Rough sets enable us to keep track of these approximations since they relate the E_j and T_k to the A_i. If we have the E_j and T_k given in terms of the A_i we can write down logical expressions relating them, and manipulate them using the techniques described in Section 3. Alternatively we can write down the logical relations between the E_j and T_k, based on an unknown set $\mathbf{A}$, from expert knowledge, quantifying the relations in much the same way as other certainty values [38] are obtained. In this case the propagation of the quantifiers is based on rough set theory, and thus takes into account the mismatch between the observable facts and the underlying concepts. This is covered in Section 4. Since all ideas of accuracy are relative to the set of basic concepts $\mathbf{A}$, we can relate quantifiers based on rough sets to the kind of truth values discussed in other formalisms by careful choice of $\mathbf{A}$. This matter is discussed in [29].

4. Reasoning with rough sets

Having introduced the basic concepts of rough set theory, we proceed, in this section, to show how they may be used to define a new method for dealing with imprecisely known information.

4.1 Combining rough sets

The degree to which an object T may be defined within $(\mathbf{A}, \mathbf{E})$ depends on the cardinality of T^c and T^e. Pawlak [31] gives the following:

If	$T^c = T^e$	then A is precisely defined by $(\mathbf{A}, \mathbf{E})$	
If	$T^c \neq T^e$ and $T^e \neq \emptyset$	then A is roughly defined by $(\mathbf{A}, \mathbf{E})$	
If	$T^c = \emptyset$	then A is internally undefined by $(\mathbf{A}, \mathbf{E})$	(12)
If	$T^e = \mathbf{E}$	then A is externally undefined by $(\mathbf{A}, \mathbf{E})$	
If	$T^c = \emptyset$ and $T^e = \mathbf{E}$	then A is totally undefined by $(\mathbf{A}, \mathbf{E})$	

We can determine the degree to which logical combinations of roughly defined objects may themselves be defined. The usual logical operations of disjunction, $\vee$, conjunction, $\wedge$, and negation, $\neg$, may be defined in terms of set operations on the core and envelope of the objects concerned. We define:

$$
\begin{array}{lllll}
(T \vee S)^c = (T \cup S)^c & & (T \vee S)^e = (T \cup S)^e \\
(T \wedge S)^c = (T \cap S)^c & & (T \wedge S)^e = (T \cap S)^e \\
(\neg T)^c = (\sim T)^c & & (\neg T)^e = (\sim T)^e
\end{array}
\qquad (13)
$$

where we give $\{\neg, \vee, \wedge\}$ their usual meaning so that $\neg T$ means not T, $T \wedge S$ means both T and S, and $T \vee S$ means either T or S or both. Using (6) we obtain:

$$
\begin{array}{lllll}
(T \vee S)^c \supseteq T^c \cup S^c & & (T \vee S)^e = T^e \cup S^e \\
(T \wedge S)^c = T^c \cap S^c & & (T \wedge S)^e \subseteq T^e \cap S^e \\
(\neg T)^c = \sim(T^e) & & (\neg T)^e = \sim(T^c)
\end{array}
\qquad (14)
$$

we can use these results to define other logical operations such as material implication, $\rightarrow$, where $T \rightarrow S \equiv \neg T \vee S$:

$$
(T \rightarrow S)^c \supseteq \sim T^e \cup S^c \qquad (T \rightarrow S)^e = \sim T^c \cup S^e \qquad (15)
$$

Note that the core of the disjunction of two terms may not be specified precisely, we only ever know the lower bound. Since the core is itself a lower bound on the accurate description, this is not a problem. Similarly, given the inequality in the description of the envelope of the conjunction, the envelope of any combination involving conjunction will only be defined by an upper bound. Again, this is not problematic.

Given the results of equations (13)–(15), we can deduce the logical relationships between objects if the set of concepts $\mathbf{A}$ is known. For instance, consider a set of concepts $\mathbf{A} = \{A_1, A_2, A_3, A_4, A_5\}$, and a pair of objects E_1 and E_2 where $E_1 = [\{A_1, A_2\}, \{A_1, A_2, A_3\}]$, which is to say that the core of E_1 is $\{A_1, A_2\}$ and its envelope is $\{A_1, A_2, A_3\}$, and $E_2 = [\{A_2, A_3\}, \{A_2, A_3, A_4\}]$. Here we can say that $T_1 = [\{A_2\},$

$\{A_2, A_3\}]$, is equivalent to $E_1 \wedge E_2$, while $T_1 = [\{A_1, A_2, A_3\}, \{A_1, A_2, A_3, A_4\}]$, is equivalent to $E_1 \vee E_2$.

Similarly we can establish the validity of logical statements about objects whose rough descriptions we know. Given $E_3 = [\{A_3, A_4\}, \{A_2, A_3, A_4, A_5\}]$ and $E_4 = [\{A_2, A_3\}, \{A_1, A_2, A_3\}]$ we can say that $E_3 \rightarrow E_4$ will have the rough description $[\{A_1, A_2, A_3\}, \{A_1, A_2, A_3, A_5\}]$.

4.2 A logic for rough reasoning

The ideas introduced in the previous section can be adapted to create a quantified logic in which rough sets are used to model the upper and lower bounds on the value of the propositions. The core and envelope are still composed of partitions of E, the language with which we may describe objects. Thus the value of a proposition is determined by the accuracy with which it is determined by the set of partitions. If we consider that the full set of partitions alone suffices to accurately describe true propositions we can relate the accuracy of description to truth.

We will deal with a first order predicate logic language PRL of roughly described terms where (after Reeves and Clarke [36]) the constant terms are the names N of all the roughly described objects that we are interested in. In addition we have a set of connectives $\{\neg, \rightarrow, \vee, \wedge\}$, a set of punctuation symbols {(,),,}, and a set of quantifier symbols $\{\forall, \exists\}$. Finally there is a set P of sets P_n of n-ary predicate symbols for each $n \geq 0$, a set F_n of n-ary function symbols {*f, g, h*...} for each $n \geq 0$, and a set V of variables $\{x_1, x_2, ...\}$. The set of sentences L(<P N F>), based on this logic is defined by:

(1) any element of N is a *term* based on L(<P N F>).

(2) If p is a member of P, and $x_1, ..., x_n$ are members of V, then $\forall p(x_1, ..., x_n)$ and $\exists p(x_1, ..., x_n)$ is a *formula* based on L(<P N F>).

(3) If S and T are formulae based on L(<P N F>) then so are $\neg S$, $S \rightarrow T$, $S \wedge T$ and $S \vee T$.

(4) If $S_1, ..., S_n$ are terms or formulae based on L(<P N F>) then $f(S_1, ..., S_n)$, where the arity of *f* is n, is a term based on L(<P N F>).

(5) If $S_1, ..., S_n$ are are terms or formulae based on L(<P N F>) then $p(S_1, ..., S_n)$ where p is amember of P, and the arity of p is n, is a term based on L(<P N F>).

(6) Nothing else is term or a formula based on L(<P N F>).

If x_1 is a variable in a sentence, and x_1 is not in the scope of any quantifier, then x_1 is a *free* variable. If S is a member of L(<P N F>) and x_1 is a free variable in S then the term t is *free for* v in S iff there is no variable x_2 in t such that x_1 appears within the scope of the quantifier that binds x_2 in S.

PRL has a set of axioms generated by the following axiom schemas. If S, R and T are members of L(<P N F>), then:

(A1) $(S \rightarrow (T \rightarrow S))$

(A2) $((S\to(T\to R))\to((S\to T)\to(S\to R)))$
(A3) $(((\neg S)\to(\neg T))\to(T\to S))$
(A4) $(\forall x_i\ S\to S)$, if x_i does not occur free in S
(A5) $(\forall x_i\ S\to S[t/x_i])$ if S is a formula of the language in which x_i may appear free, and t is free for x_i in S.
(A6) $(\forall x_i\ (S\to T)\to(S\to\forall x_i\ T))$, if x_i does not occur free in S

We have a number of rules of inference for PRL :

(MP) Modus ponens: From S and $S\to T$, for any S and T, deduce T
(MT) Modus tollens: From $\neg T$ and $S\to T$, for any S and T, deduce $\neg S$
(R) Resolution: From $S\vee T$ and $\neg S\vee R$, for any S, R and T, deduce $T\vee R$
(S) Syllogism: From $S\to T$ and $T\to R$, for any S, R and T, deduce $S\to R$
(UI) Universal instantiation: From $\forall x_i\ Q(x_i)$ deduce $Q(a)$ where Q is any formula, and a is any term not containing free variables.

We are interested in the rough set descriptions of the sentences based on $L(\langle P\ N\ F\rangle)$, which we manipulate as quantifiers expressing our knowledge about these concepts. We define a rough measure R over the sentences of $L(\langle P\ N\ F\rangle)$ such that for any element p of $L(\langle P\ N\ F\rangle)$, $R(p) = [p^{\geq C}, p^{\leq E}]$ where $p^{\geq C}$ is the lower bound on the core of p and $p^{\leq E}$ is the upper bound on the envelope of p. $R(p)$ gives us the rough set that describes p. The components of the rough set are drawn from a finite set U of descriptors. Thus, rather than dealing with numerical quantifiers expressing the truth of the sentences we quantify the sentences with examples of the concept that they embody. It is possible, however, to use rough sets to model truth values [29].

5. Rough truth values

The rough measure described in Section 3 maintains the core and envelope of each distinct sentence, allowing a precise estimation of the degree to which it is defined by the system. In this section we investigate coarsenings of this measure. Instead of attaching a core and envelope to each sentence, we use the core and envelope to define a rough truth values for each sentence which are then propagated in place of the approximations themselves.

5.1 Symbolic truth values

The rough measure of a sentence $p \in L(\langle P\ N\ F\rangle)$ is determined by the degree to which its core and envelope are defined by the set of descriptors $\mathbf{A}$. We can distinguish the following boundary cases for $\emptyset\subset X\subset\mathbf{A}$, and $\emptyset\subset Y\subset\mathbf{A}$ which correspond to the definitions of (12):

If	$R(p)$	=	$[\mathbf{A}, \mathbf{A}]$	then p is true	
If	$R(p)$	=	$[X, \mathbf{A}]$	then p is roughly true	
If	$R(p)$	=	$[\emptyset, \mathbf{A}]$	then p is of unknown value	(16)
If	$R(p)$	=	$[\emptyset, Y]$	then p is roughly false	
If	$R(p)$	=	$[\emptyset, \emptyset]$	then p is false	

These rough values form a lattice, ordered by set inclusion $\subseteq$, giving the following:

$$[\emptyset, \emptyset] \subseteq [\emptyset, Y] \subseteq [\emptyset, \mathbf{A}] \subseteq [X, \mathbf{A}] \subseteq [U, \mathbf{A}] \quad (17)$$

false roughly false unknown roughly true true

This suggests the introduction of a rough truth measure RV over $L(\langle P\ N\ F\rangle)$ which identifies which of these five ordered states the rough measure of each $p \in L(\langle P\ N\ F\rangle)$ falls into. The advantages of such a measures are its extreme simplicity and robustness, a direct result of the simple conditions used to define the values, and the fact that the values are ordered. The latter allows the axioms of the rough truth measure to be simply stated:

$$\begin{aligned} RV(p \vee q) &= \max(RV(p), RV(q)) \\ RV(p \wedge q) &= \min(RV(p), RV(q)) \end{aligned} \quad (18)$$

These may be easily verified by considering the set operations on the rough measure for each proposition. Similar considerations will validate the negation operator:

RV(p)	true	roughly true	unknown	roughly false	false
RV¬(p)	false	roughly false	unknown	roughly true	true

(19)

The quantified logic may be applied without consideration for the set of basic concepts that underlie the rough values. However, the sound mathematical basis on which the logic is built ensures that it is well behaved.

It is possible to identify a further set of truth values. These are the contradictory[5] value $R(\perp) = [U, \emptyset]$, the indeterminate value $R(\mathbf{i}) = [Y, X]$, and two 'partially contradictory' values $R(s) = [U, Y]$ and $R(t) = [X, \emptyset]$. These, happily, need not concern us, since they are not generated by operations on the values introduced above.

5.2 Reasoning with rough truth values

In order to use our symbolic truth values for practical reasoning purposes, we need to provide a set of rules for propagating inference. It is trivial to establish the rough value of a material implication from that for disjunction:

$$RV(p \rightarrow q) = \max(RV(\neg p), RV(q)) \quad (20)$$

This is, however, of limited practical use since it only allows us to establish the strength of an implication from that of its antecedent and consequent. Rules giving the strength of the consequent (antecedent) based on the strength of the material implication and its antecedent (consequent) are of more use in an expert system context. To this end we can specify how the rough truth values are propagated when we use the logical inference rules of Section 4.2.

[5]From the definition of core and envelope as being lower and upper approximations, respectively, it is clear that a set of rough values for p that have $p^c \supset p^e$ are contradictory to the underlying rough set theory.

Theorem 5.1: For modus ponens (MP) we have:

$$\begin{array}{ccccc} & & RV(p \rightarrow q) & = & \alpha \\ & & RV(p) & = & \beta \\ \hline \alpha & \geq & RV(q) & \geq & \min(\alpha, \beta) \end{array} \qquad (21)$$

Where p, q are any sentences of L(<P N F>).

Proof: $RV(q) \geq RV(p \cap q) = RV((\neg p \vee q) \cap p) = \min(\alpha, \beta)$, so the lower bound on RV(q) is $\min(\alpha, \beta)$. In addition, $\alpha = RV(\neg p \vee q) \geq RV(q)$, so the upper bound on RV(q) is α.■

Theorem 5.2: For modus tollens (MT):

$$\begin{array}{ccccc} & & RV(p \rightarrow q) & = & \alpha \\ & & RV(\neg q) & = & \beta \\ \hline \alpha & \geq & RV(\neg p) & \geq & \min(\alpha, \beta) \end{array} \qquad (22)$$

Where p, q are any sentences of L(<P N F>).

Proof: as for Theorem 5.1.■

In order to reason about specific roughly described objects if we are given general statements about classes of such objects we need to know how our rough measure responds to instantiation:

Theorem 5.3: For universal instantiation (UI):

$$\begin{array}{ccccc} & & RV(\forall x_i P(x_i)) = & & \alpha \\ \hline t & \geq & RV(P(a)) & \geq & \alpha \end{array} \qquad (23)$$

where a is any rough object.

Proof: $RV(\forall x_i P(x_i)) = RV(P(a)) \wedge RV(P(b)) \wedge \ldots \wedge RV(P(n)) = \min[RV(P(a)), RV(P(b)), \ldots, RV(P(n))]$. Thus $RV(P(a)) \geq RV(\forall x_i P(x_i))$.■

These theorems enable us to propagate rough measures in our logic of rough objects PRL.

5.3 An example

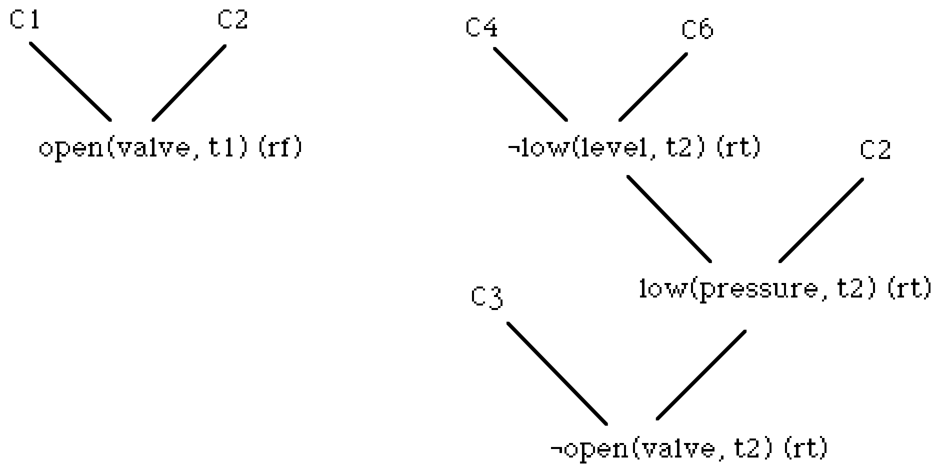
To illustrate the kind of reasoning possible with our rough valued logic, consider the following simple example. An intelligent system has a set of rules which it uses to determine when it is appropriate to open and close a pressure regulating valve. These rules are expressed as logical sentences, each of which is quantified with a rough truth value:

C1	high(pressure, x)	→	open(valve, x)	(t)
C2	¬low(pressure, y)	→	low(level, y)	(t)
C3	low(pressure, z)	→	¬open(valve, z)	(t)
C4	high(temperature, w)	→	¬low(level, w)	(rt)

The system is also supplied with two predictions about the state of the system at times t_1 and t_2:

C5	high(pressure, t_1)	(rf)
C6	high(temperature, t_2)	(rt)

From this data the system can use the rules of modus ponens (21) and modus tollens (22) along with universal instantiation to find the lower biunds on the rough truth value of new facts:



to establish that the valve may be open at time t_1, and is very likely to be closed at time t_2.

5.4. Robust reasoning in rule based systems

For most practical purposes, intelligent knowledge-based systems are rule-based , with knowledge encoded in the form of 'if...then...' rules. In many domains detailed numerical estimates of the certainty of rules and facts may be impossible to obtain, and the reasoning mechanism adopted must be capable of dealing with vague estimates. Especially important is the robustness of the mechanism— its ability to deal with rules and facts whose certainty is unknown. In this section we analyse the robustness of rough valued logic in the context of rule based reasoning.

The knowledge base of a typical rule-based system consists of a series of rules of the form 'if p then q' with a certainty value attached to each. In forward chaining inference starts with one or more facts, also with an associated certainty, which match the antecedents of particular rules. These rules are fired to obtain their consequents, with the certainty of the consequent being determined by a combination of the certainties of rule and antecedent, and the consequents used to fire more rules. This process continues until there are no facts that match the antecedents of unfired rules, or the goal fact has been deduced. If we assume that rules of the form 'if p then q' are translated by use of material implication into logical statements of the form $p \rightarrow q$, then the mechanism of forward chaining is the rule of modus ponens (21). This allows us to establish when consequent values of unknown certainty will be generated, taking the lower bound on the value of q:

p	p → q	q
t	u	u
rt	u	u
u	u	u
rf	u	rf
f	u	f

p	p → q	q
u	t	u
u	rt	u
u	u	u
u	rf	rf
u	f	f

No combinations of values of antecedent and rule other than those shown, have either antecedent, consequent, or rule valued as unknown. The truth tables show that the value of the consequent can be determined if the value of the rule and the antecedent is given, or, if one has an unknown value the other has the value false (f) or roughly false (rf).

A similar analysis may be performed for backward chaining. Here we are interested in determining the antecedent of a rule from the rule and its consequent. Reasoning proceeds from the goal, and continues until a known fact is identified as the antecedent of a rule that must be fired in order to generate the goal. Using the pattern for rough modus tollens (22) we obtain a similar truth tables to those above for the lowest bound on the value of p:

¬q	p → q	¬p
t	u	u
rt	u	u
u	u	u
rf	u	rf
f	u	f

¬q	p → q	¬p
u	t	u
u	rt	u
u	u	u
u	rf	rf
u	f	f

Once again, no combinations of values of antecedent and rule other than those shown, have either antecedent, consequent, or rule valued as unknown.The tables show that a fact of known value may be deduced from a fact and a rule of known value, or from a fact of unknown value and a rule that is false (f) or roughly false (rf), or from a rule of unknown value and a fact that is false or roughly false. In both these latter cases, the deduced fact has value false or roughly false. Thus, in both forward and backward chaining, uncertainty can be absorbed by the logic at the cost of reducing the certainty value of the facts deduced. In situations in which the certainty of conclusions is secondary to the need to continue to operate in the face of degraded information such behaviour will be an advantage.

6. Other patterns of rough reasoning

There are two other patterns of reasoning that are important, the resolution rule and, to a lesser extent, syllogism. These are discussed at some length in this section. Firstly we indicate how rough truth values may be propagated across these patterns, then we prove two results about the value of conclusions established using the resolution rules. Finally we show the resolution rule in action.

6.1 Resolution and syllogism

Rough truth values may be propagated across these inference patterns as follows:

Theorem 6.1: For resolution (R):

$$\frac{RV(p \vee q) = \alpha \qquad RV(\neg p \vee r) = \beta}{\max(\alpha, \beta) \geq RV(q \vee r) \geq \min(\alpha, \beta)} \tag{24}$$

Where p, q , r are any sentences of $L(\langle P\ N\ F\rangle)$.

Proof: $RV(q \vee r) = RV((p \vee q \vee r) \wedge (\neg p \vee q \vee r)) = \min(RV(p \vee q \vee r), (\neg p \vee q \vee r))$. Now, $RV(p \vee q \vee r) \geq RV(p \vee q) = \alpha$, and $RV(\neg p \vee q \vee r) \geq RV(\neg p \vee q) = \beta$, and we get the lower limit. For the upper limit, consider the fact that $RV(p \vee q) = \max(RV(p), RV(q))$, so that the maximum value of $RV(q)$ is α, and the maximum value of $RV(r)$ is β. Since $RV(q \vee r) = \max(RV(q), RV(r))$, $RV(q \vee r) \leq \max(\alpha, \beta)$.■

Theorem 6.2: Syllogism (S):

$$\frac{RV(p \rightarrow q) = \alpha \qquad RV(q \rightarrow r) = \beta}{\max(\alpha, \beta) \geq RV(p \rightarrow r) \geq \min(\alpha, \beta)} \tag{25}$$

Where p, q , r are any sentences of $L(\langle P\ N\ F\rangle)$.

Proof: This follows from the resolution principle. Rewriting the pattern remembering that $(a \rightarrow b) \equiv (\neg a \vee b)$, we get: $RV(\neg p \vee q) = \alpha$, $RV(\neg q \vee r) = \beta$ which resolve to give $\max(\alpha, \beta) \geq RV(\neg p \vee r) \geq \min(\alpha, \beta)$.■

And we have sufficient results to handle most reasoning tasks using PRL.

6.2 The results of resolution

The rules discussed above allow us to us the quantified versions of PRL to infer the truth of new facts from the truth of existing facts and relations. Thus a system using quantified PRL can infer new information from its observations. In this section we investigate how the truth values of the conclusions are related to the truth values of the facts on which the conclusions are based. We will centre our investigation around the commonly used resolution rule, following the approach of Dubois *et al* [11]. Given a set of clauses quantified with truth values derived from rough set theory, we are interested in determining the bounds on the truth value of the clauses derived by applying the resolution rule. We have:

Theorem 6.3: the truth value of the result of resolution between the members of a rough valued set of clauses must lie between the maximum and minimum values attached to members of that set of clauses.

Proof: Let C be a set of clauses of quantified RL, and $R(C)$ be the union of C with all the results of resolving together every pair of clauses in C that may be resolved together. We write $R^n(C)$ to denote the result of iterating this procedure n times. We can easily see that for $C = \{C_1, \ldots, C_m\}$, where $\forall i = 1, \ldots, m$, $RV(C_i) \geq \alpha_i$, and C^n denotes any clause in $R^n(C)$, we have $\forall n \geq 0$, $RV(C^n) \geq \min_{i=1,\dots,m} \alpha_i$, and $\max_{i=1,\dots,m} \alpha_i \geq RV(C^n)$.■

While resolution may be used to reason 'forwards' from a set of known clauses in order to establish new facts, it is often used to reason 'backwards' in proofs by refutation. In proof by refutation, in order to assess whether a clause C_q follows from a set of clauses C we repeatedly apply the refutation rule to the set of clauses $C \cup \neg C_q$. If we deduce the empty clause {} (ie. by resolving a and $\neg a$) then the clause C_q follows from the set C. In both of our logics we will have truth values attached to all clauses, including the empty clause at the end of the proof. What we would like to establish is the correspondence between the truth value we derive for the empty clause and the truth value of the clause we set out to prove.

Theorem 6.4: the rough truth value of the empty clause that results from resolution is the lower bound on the value of the clause we are trying to prove.

Proof: Let $C = \{C_1, \dots, C_m\}$ be the set of clauses from which we are trying to prove C_q.After the proof, we obtain $\beta \geq RV(\{\}) \geq \alpha$ from a subset of C, $C_\alpha \cup \{\neg C_q\}$, where $C_\alpha = \{C_i \mid \beta \geq RV(C_i) \geq \alpha\}$ and $RV(\neg C_q) = t$. Since {} would be obtained from C_α by applying the classical resolution rule (ignoring degrees of truth), C_q is a logical consequence of C_α so that $RV(\neg(\cap_i \{C_i \in C_\alpha\}) \vee C_q) = t$. Now, $\beta \geq RV(\cap_i \{C_i \in C_\alpha\}) \geq \alpha$, so applying the resolution rule to the two clauses gives us $t \geq RV(C_q) \geq \alpha$. ■

6.5 Theorem proving using rough truth values

An example of reasoning using the rough valued resolution rule is the following problem. Another control system has a number of pieces of information that relate the states of a number of valves to the pressure of some vessel:

C1. If $Valve_1$ is open, then $Valve_2$ is not open.
C2. $Valve_1$ will be open at time t_1.
C3. If $Valve_3$ is open then pressure is likely to be low.
C4. $Valve_3$ may be open at time t_2
C5. If $Valve_4$ is open, and $Valve_2$ is not open, then it is almost certain that the pressure will be low.
C6. It is likely that $Valve_2$ or $Valve_5$ will be open at t_1
C7. If $Valve_5$ is open then $Valve_4$ will be open

The system wishes to ascertain if the pressure will be low at any time so we add the clause:

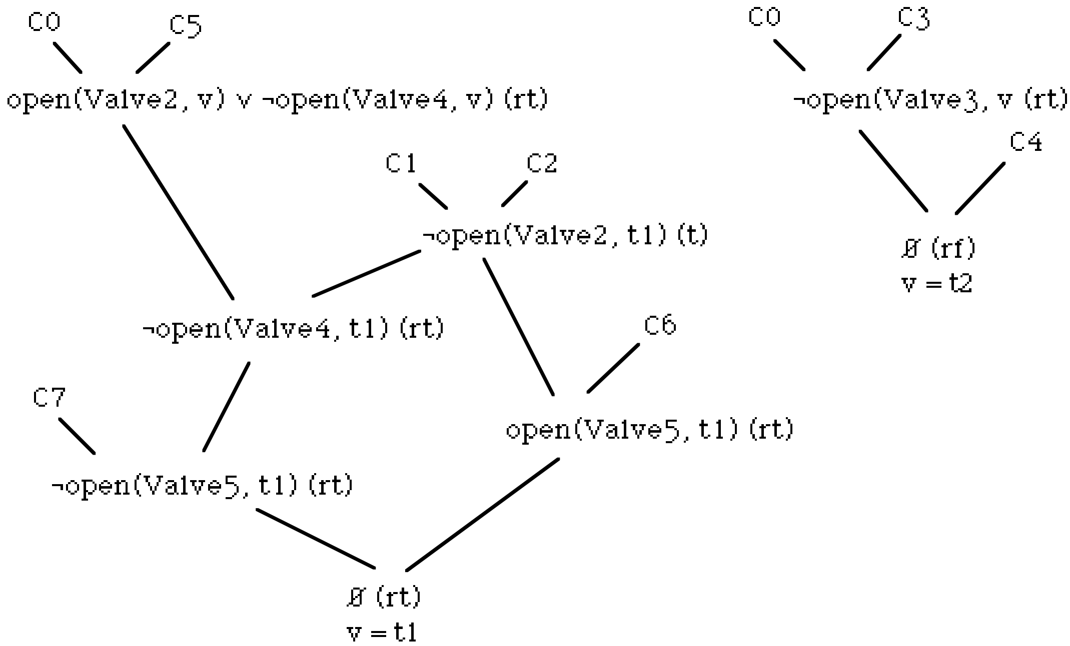
C0. The pressure will not be low.

The clauses may be translated into first order logic quantified with rough truth values:

C0. $\neg low(pressure, v)$ (t)
C1. $\neg open(Valve_1, x) \vee \neg open(Valve_2, x)$ (t)
C2. $open(Valve_1, t_1)$ (t)
C3. $\neg open(Valve_3, y) \vee low(pressure, y)$ (rt)
C4. $open(Valve_3, t_2)$ (rf)
C5. $open(Valve_2, z) \vee \neg open(Valve_4, z) \vee low(pressure, z)$ (rt)
C6. $open(Valve_2, t_1) \vee open(Valve_5, t_1)$ (rt)

C7. ¬open(Valve$_5$, w) v open(Valve$_4$, w) (t)

There are two possible refutations of C0:



Thus we can conclude that the pressure may be low at time t_2, and is likely to be low at time t_1.

7. Discussion

There are a three points arising from the previous sections that are worthy of brief discussion. The first concerns the alternative view of uncertainty that this paper espouses, as the mismatch between what relates cause to effect and what is observed, and which the rough set approach models. This seems to explain the presence of uncertainty in many domains quite neatly, but it remains to be seen if it is really helpful. Further work will help to resolve this concern, especially work to use rough logic and symbolic truth values in systems that must contend with uncertainty. The second point concerns the symbolic truth values and their provision of a simple method of handling uncertainty where only a few different truth values are required. These values have the interesting property of absorbing uncertainty in certain circumstances, in that the combination of some known values with the value “unknown” produces the known values. This is in contrast to other approaches to reasoning under uncertainty where combination with “unknown” values causes values to become less certain. Again, the usefulness of this property remains to be proven, and the development of exacting applications would seem to be the best means of doing so. The third point is the similarity between our symbolic truth values and possibility measures [10], [11]. From the results presented, it does seem as though the symbolic measures behave in much the same way as possibility measures, a fact that might be expected given the similarity between the formalisms on which the two measures are based, namely rough and fuzzy sets. However, it is not yet clear whether this connection is meaningful or superficial, and to what extent it is related to the close relationship that exists between rough sets, and measures defined over them, and belief functions [29], [43]. We hope to be able to investigate this relationship more closely in the near future.

8. Summary

In summary, we have presented several different ways in which rough sets may be used in reasoning under uncertainty. It is simple to define quantities in terms of rough sets, and formalisms based upon them are robust in the face of scarce information; a great advantage in dealing will ill-known domains such as biotechnology and ecology. Extending Pawlak's original work on rough sets we have defined symbolic truth values based on rough measures. These values are simple and robust, being exceptionally good at assimilating ignorance. The values have been used to quantify a predicate logic, and rules for combining values when the inference rules of the logic are applied have been introduced. Some results have been given for theorem proving using the quantified logic, and several examples have demonstrated the kind of automated reasoning that may be achieved using the logic.

References

[1] Buchanan, B. G. and Shortliffe, E. H. *Rule-based expert systems: the MYCIN experiments of the Stanford heuristic programming project*, Addison Wesley, Reading, Mass, 1984.

[2] Bundy, A. Incidence calculus: a mechanism for probabilistic reasoning, *Journal of Automated Reasoning*, 1, 263–284, 1985.

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

[4] Cohen, P. R., Greenberg, M. L. Hart, D. M. and Howe, A. E. Trial by fire: understanding the design requirements for agents in complex environments, *AI Magazine*, 10, 34–48, 1989.

[5] Cohen, P. R. *A heuristic approach to reasoning with uncertainty*, Pitman, London, 1985.

[6] Cooper, G. F. The computational complexity of probabilistic inference using Bayesian belief networks, *Artificial Intelligence*, 42, 393–405, 1990.

[7] Davis, R., Buchanan, B. and Shortliffe, E., Production rules as a representation for a knowledge-based consultation program, *Artificial Intelligence* 8, 15–45, 1977.

[8] Dawes, R. M. The robust beauty of improper linear models in decision making, in *Judgement under uncertainty: Heuristics and biases* (D. Kahnemann, P. Slovic, and A. Tversky, Eds.), Cambridge University Press, Cambridge, 391–407, 1982.

[10] Dubois, D. and Prade, H. Resolution principles in possibilistic logic, *International Journal of Approximate Reasoning*, 4, 1–21, 1990.

[11] Dubois D., Lang J., Prade H. Automated reasoning using possibilistic logic: semantics, belief revision and variable certainty weights. *Proceedings of the 5th Workshop on Uncertainty in Artificial Intelligence*, Windsor, Ontario 81–87, 1989.

[12] Earman, J. *Bayes or bust? A critical examination of bayesian confirmation theory*, MIT Press, Cambridge, MA.

[13] Epstein, R. Can machines think, *AI Magazine*, 13, 2 80–95, 1992.

[14] Fariñas del Cerro, L., Orłowska, E., DAL - a logic for data analysis, *Theoretical Comp. Sci.* 36, 251–264. 1985.

[15] Fox, J. Decision theory and autonomous systems, in *Decision Support Systems and Qualitative Reasoning*, (M. G. Singh and L. Travé-Massuyès, eds), Elsevier, North-Holland, 1991.

[16] Fox, J. Three arguments for extending the framework of probability, in *Uncertainty in Artificial Intelligence* (L.N. Kanal and J. F. Lemmer, eds), Elsevier, North-Holland, 1986.

[17] Haack, S. Do we need "fuzzy logic"? *International Journal of Man-Machine Studies*, 11, 437–445, 1979.

[18] Heckerman, D. E. Probabilistic interpretation for MYCIN's certainty factors, in *Uncertainty in Artificial Intelligence* (L.N. Kanal and J. F. Lemmer, eds), Elsevier, North-Holland, 1986.
[19] Kahneman, D., Slovic, P. and Tversky, A. *Judgement under uncertainty: Heuristics and biases*. Cambridge University Press, Cambridge 1982.
[20] Lauritzen, S. L. and Spiegelhalter D. J. Local computations with probabilities on graphical structures and their applications to expert systems, *The Journal of the Royal Statistical Society*, B 50, 157–224, 1988.
[21] Krause, P. J. and Clark, D. A. *Representing uncertain knowledge: and artificial intelligence approach*, Intellect, Oxford, 1993.
[22] Lindley, D. V. *Introduction to probability theory and statistics from a bayesian point of view*, Cambridge University Press, Cambridge, 1965.
[23] McDermott, J. R1: a rule based configurer of computer systems, *Artificial Intelligence*, 19, 39–88, 1982.
[24] McLeish, M, Probabilistic logic: some comments and possible use for nonmonotonic reasoning, in *Uncertainty in Artificial Intelligence 2* (L.N. Kanal and J. F. Lemmer, eds), Elsevier, North-Holland, 1988.
[25] Negoita, C. V. and Ralescu, D. A. *Applications of Fuzzy Sets to Systems Analysis*, Bikhäuser Verlag, Basel, 1975.
[26] Nilsson, N. J. Probabilistic logic, *Artificial Intelligence*, 28, 71–87.
[27] Orłowska, E. and Pawlak, Z. Expressive power of knowledge representation systems, *International Journal of Man-Machine Studies*, 20, 485–500, 1984.
[28] Parsons, S. Qualitative methods for reasoning under uncertainty, Ph.D. Thesis, Department of Electronic Engineering, Queen Mary and Westfield College, London, 1993.
[29] Parsons, S., Kubat, M. and Dohnal, M. A rough set approach to reasoning under uncertainty, Technical Report, *Journal of Experimental and Theoretical Artificial Intelligence*, (to appear).
[30] Pawlak, Z. Rough classification, *International Journal of Man-Machine Studies*, 20, 469–483, 1984.
[31] Pawlak, Z. Rough Sets, *International Journal of Information and Computer Sciences*, 11, 341–356, 1982.
[32] Pawlak, Z., Wong, S. K. M., and Ziarko, W. Rough sets: probabilistic versus deterministic approach, *International Journal of Man-Machine Studies*, 29, 81–95, 1988.
[33] Pawlak, Z., Słowinski, K., and Słowinski, R. Rough classification of patients after highly selective vagotomy for duodenal ulcer, *International Journal of Man-Machine Studies,* 24, 81–95, 1986.
[34] Pearl, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufman, San Mateo, CA, 1988.
[35] Rutherford, D. E. *Introduction to Lattice Theory*, Oliver and Boyd Ltd, Edinburgh, 1965.
[36] Reeves, S. and Clarke, M. *Logic for computer science*, Addison Wesley Publishing Company, Wokingham, England, 1990.
[37] Saffiotti, A. A belief function logic, *Proceedings of the 10th National Conference on Artificial Intelligence*, San Jose, 1992.
[38] Saffiotti, A. An AI view of the treatment of uncertainty. *The Knowledge Engineering Review*, 2, 75–97 1987.
[39] Shafer, G. The early development of mathematical probability, Working Paper 228, School of Business, The University of Kansas.
[40] Shafer, G. *A mathematical theory of evidence*, Princeton University Press, Princeton, 1976.
[41] Smets, Ph, Mamdani, E. H., Dubois, D. and Prade, H.*Non-standard logics for automated reasoning*, Academic Press, London, 1988.
[42] Wellman, M. P. Fundamental concepts of qualitative probabilistic networks, *Artificial intelligence*, 44, 257–303, 1990.

[43] Wong, S. K. M., Wang, L. S. and Yao, Y. Y. Interval structure: a framework for representing uncertain information, *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, Stanford, 1992.

[44] Wong, S. K. M., Ziarko W. and Li Ye, R. Comparison of rough-set and statistical methods in inductive learning, *International Journal of Man-Machine Studies*, 24, 53–72, 1986.

[45] Zadeh, L. A. Fuzzy sets as a basis for a theory of possibility, *Fuzzy sets and systems*, 1, 1–28, 1978.

[46] Zadeh, L. A. Fuzzy sets, *Info. Control*, 8, 338–353, 1965.