Rough Sets for Uncertainty Reasoning

S.K.M. Wong[1] and C.J. Butz[2]

[1] Department of Computer Science, University of Regina, Regina, Canada, S4S 0A2, wong@cs.uregina.ca
[2] School of Information Technology & Engineering, University of Ottawa, Ottawa, Canada, K1N 6N5, butz@site.uottawa.ca

Abstract. Rough sets have traditionally been applied to decision (classification) problems. We suggest that rough sets are even better suited for reasoning. It has already been shown that rough sets can be applied for reasoning about knowledge. In this preliminary paper, we show how rough sets provide a convenient framework for uncertainty reasoning. This discussion not only presents a new topic for future research, but further demonstrates the flexibility of rough sets.

1 Introduction

The theory of rough sets [4] generalizes traditional set theory by allowing a concept to be described approximately by a lower and upper bound. Although rough sets have been extensively studied, most of these investigations demonstrated the usefulness of rough sets in decision (classification) problems. Wong [7] first demonstrated that rough sets can also be applied for reasoning about knowledge. This observation was also made later by Salonen and Nurmi [5].

In this preliminary paper, we extend the work in [5, 7] by demonstrating that rough sets can also be applied for *uncertainty* management. In [5, 7], rough sets are used as a framework to represent formulas such as "player 1 *knows* ϕ". By incorporating probability, we can now represent sentences such as "the probability of ϕ, according to player 1, is at least α", where ϕ is a formula and α is a real number in $[0, 1]$. Thereby, not only does this discussion present a new topic for future research, but it further demonstrates the flexibility of rough sets.

The remainder of this paper is organized as follows. Kripke semantics for modal logic are given in Section 2. The key relationships between rough sets and the Kripke semantics for modal logic are stated in Section 3. In Section 4, probability is incorporated into the logical framework. In Section 5, we demonstrate that rough sets are also a useful framework for uncertainty reasoning. The conclusion is given in Section 6.

2 Kripke semantics for Modal Logic

Consider an ordered pair $< W, R >$ consisting of a nonempty set W of possible worlds and a binary relation R on W. Let Q denote the set of sentence letters (primitive propositions). An evaluator function f:

$$f : W \times Q \to \{\top, \bot\},$$

assigns a truth-value, $\top$ or $\bot$, to each ordered pair (w, q), where $w \in W$ is a *possible world*, and $q \in Q$ is a sentence letter. We call the triple $M =< W, R, f >$ a *model* (structure), and R a *possibility* (accessibility) relation.

The function of an evaluator is to determine which primitive proposition q is to be *true* at which world w in a model. We write $(M, w) \models q$, if $f(w, q) = \top$. We can now define what it means for a proposition (formula) to be true at a given world in a model by assuming that $\models$ has been defined for all its subformulas of φ. That is, for all propositions φ and ψ,

$$(M, w) \models \varphi \wedge \psi \text{ iff } (M, w) \models \varphi \text{ and } (M, w) \models \psi,$$
$$(M, w) \models -\varphi \text{ iff } (M, w) \not\models \varphi,$$

and

$$(M, w) \models \Box\varphi \text{ iff } (M, x) \models \varphi, \text{ for all } x \text{ such that } (w, x) \in R.$$

The above definition enables us to infer inductively the truth-value, i.e., $(M, s) \models \varphi$, of all other propositions from those of the primitive propositions. We say "φ is true at (M, s)" or "φ holds at (M, s)" or "(M, s) satisfies φ", if $(M, s) \models \varphi$.

In order to establish a connection with rough set theory, we review the notion of an *incidence* mapping [1], denoted by I. To every proposition φ, we can assign a set of worlds $I(\varphi)$ defined by:

$$I(\varphi) = \{w \in W | (M, w) \models \varphi\}.$$

This function is used in establishing the relationship between a Kripke structure and an Auman structure in the recent work of Fagin et al. [2]. The important point of this discussion is that the incidence mapping I provides a set-theoretic interpretation of Kripke semantics.

3 Rough Sets versus Kripke semantics

The original motive of rough sets [4] was to characterize a particular *concept* (represented by a subset of a finite universe W of interest) based on the information (knowledge) on hand. This *knowledge* is represented by a binary relation R on W. Rough sets can be viewed as an extension of ordinary sets, in which a set $A \subseteq W$ is described by a pair $(\underline{A}, \overline{A})$ of subsets of W. Note that $\underline{A}$ and $\overline{A}$ are not necessarily distinct. For our exposition here, we may assume that R is an equivalence relation. In this case, rough sets are defined by the following *knowledge operator* K: for all $A \subseteq W$

$$\underline{A} = K(A) = \{w \in W \mid [w]_R \subseteq A\},$$

and

$$\overline{A} = -K(-A) = \{w \in W \mid [w]_R \cap A \neq \emptyset\},$$

where $[w]_R$ denotes the equivalence class of R containing the elements $w \in W$. In the theory of rough sets, we call $\underline{A}$ the lower approximation and $\overline{A}$ the upper approximation of A.

It was shown [6] that the Kripke semantic model is equivalent to the characterization of modal propositions by a rough-set model. That is, each proposition $\varphi \in L$ can be represented by a subset of possible worlds and the modal operator $\Box$ by the knowledge operator K defined above. The key relationships between the Kripke semantic model and the rough-set model are summarized as follows:

$$\begin{array}{ll} (i) & (M, w) \models \varphi \text{ iff } w \in I(\varphi), \\ (ii) & (M, w) \models \Box\varphi \text{ iff } w \in K(I(\varphi)). \end{array}$$

The above results enable us to adopt rough sets for reasoning about knowledge instead of using the framework based on modal logic as suggested by Fagin et al. [2].

We conclude this section with an example [2] to illustrate how the rough-set model is used in reasoning. Consider a deck of cards consisting of three cards labeled X, Y and Z. Assume there are two players (agents), i.e., $G = \{1, 2\}$. Players 1 and 2 each gets one of these cards. The third card is left face down. We describe a possible world by the cards held by each player. Clearly, there are six possible worlds, i.e., $W = \{(X,Y), (X,Z), (Y,X), (Y,Z), (Z,X), (Z,Y)\}$ $= \{w_1, w_2, w_3, w_4, w_5, w_6\}$. For example, $w_2 = (X, Z)$ says that player 1 holds card X and player 2 holds card Z. The third card Y is face down. We can easily construct the two partitions π_1 and π_2 of W, which respectively represent the knowledge of the two players. For example, $w_1 = (X, Y)$ and $w_2 = (X, Z)$ belong to the same block of π_1 because in a world such as $w_1 = (X, Y)$, player 1 considers two worlds possible, namely $w_1 = (X, Y)$ itself and $w_2 = (X, Z)$. That is, when player 1 holds card X, he considers it possible that player 2 holds card Y or card Z. Similarly, in a world $w_1 = (X, Y)$, player 2 considers the two worlds $w_1 = (X, Y)$ and $w_6 = (Z, Y)$ possible, i.e., w_1 and w_6 belong to the same block of π_2. Based on this analysis, one can easily verify that:

$$\begin{aligned} \pi_1 &= \{[w_1, w_2]_{1X},\ \ [w_3, w_4]_{1Y},\ \ [w_5, w_6]_{1Z}\}, \\ \pi_2 &= \{[w_3, w_5]_{2X},\ \ [w_1, w_6]_{2Y},\ \ [w_2, w_4]_{2Z}\}. \end{aligned}$$

It is understood that in both worlds w_1 and w_2 of the block $[w_1, w_2]_{1X}$ in π_1, player 1 holds card X; in both worlds w_1 and w_6 of the block $[w_1, w_6]_{2Y}$, player 2 holds card Y, and so on. The corresponding equivalence relations R_1 and R_2 can be directly inferred from π_1 and π_2. In this example, we have six primitive propositions: $1X$ denotes the statement "player 1 holds card X", $1Y$ denotes the statement "player 1 holds card Y", ..., and $2Z$ denotes the statement "player

2 holds card Z". Each of these propositions is represented by a set of possible worlds. By the definition of the mapping I, we obtain:

$$I(1X) = \{w_1, w_2\}, \quad I(1Y) = \{w_3, w_4\}, \quad I(1Z) = \{w_5, w_6\},$$

$$I(2X) = \{w_3, w_5\}, \quad I(2Y) = \{w_1, w_6\}, \quad I(2Z) = \{w_2, w_4\}.$$

Using these primitive representations, the representations of more complex propositions can be easily derived from properties $(i1) - (i5)$. For example,

$$\begin{aligned} I(1X \wedge 2Y) &= I(1X) \cap I(2Y) \\ &= \{w_1, w_2\} \cap \{w_1, w_6\} = \{w_1\}, \end{aligned}$$

$$\begin{aligned} I(2Y \vee 2Z) &= I(2Y) \cup I(2Z) \\ &= \{w_1, w_6\} \cup \{w_2, w_4\} = \{w_1, w_2, w_4, w_6\}. \end{aligned}$$

More interesting is the following expression which indicates that if player 1 holds card X, then he *knows* that player 2 holds card Y or card Z:

$$\begin{aligned} I(\Box_1(2Y \vee 2Z)) &= K_1(I(2Y \vee 2Z)) \\ &= \{w \mid [w]_{\pi_1 \sqcap \pi_2} \subseteq I(2Y \vee 2Z)\} \\ &= \{w \mid [w]_{\pi_1 \sqcap \pi_2} \subseteq \{w_1, w_2, w_4, w_6\}\} \\ &= \{w_1, w_2\}. \end{aligned}$$

4 Incorporating Probability

The discussion here draws from that given by Halpern [3]. The language is extended to allow formulas of the form $P_i(\phi) \geq \alpha$, $P_i(\phi) \leq \alpha$, and $P_i(\phi) = \alpha$, where ϕ is a formula and α is a real number in the interval [0,1]. A formula such as $P_i(\phi) \geq \alpha$ can be read "the probability of ϕ, according to player i, is at least α".

To give semantics to such formulas, we augment the Kripke structure with a probability distribution. Assuming there is only one agent, a *simple probability structure* M is a tuple (W, p, π), where p is a discrete probability distribution on W. The distribution p maps worlds in W to real numbers in [0,1] such that $\sum_{w \in W} p(w) = 1.0$. We extend p to subsets A of W by $p(A) = \sum_{w \in A} p(w)$. We can now define satisfiability in simple probability structures: the only interesting case comes in dealing with formulas such as $P_i(\phi) \geq \alpha$. Such a formula is true, if:

$$(M, w) \models P(\phi) \geq \alpha \text{ if } p(\{w | (M, w) \models \phi\}) \geq \alpha.$$

That is, if the set of worlds where ϕ is true has probability at least α. The treatment of $P_i(\phi) \leq \alpha$, and $P_i(\phi) = \alpha$ is analogous.

Simple probability structures implicitly assume that an agent's (player's) probability distribution is independent of the state (world). We can generalize simple probability structures with *probabilistic Kripke structures* by having p depend on the world and allowing different agents to have different probability distributions.

A *probabilistic Kripke structure* M is a tuple $(W, p_1, \ldots, p_n, \pi)$, where for each agent i and world w, we take $p_i(w)$ to be a discrete probability distribution, denoted $p_{i,w}$, over W. To evaluate the truth of a statement such as $P_i(\phi) \geq \alpha$ at world w we use the distribution $p_{i,w}$:

$$(M, w) \models P_i(\phi) \geq \alpha \text{ if } p_{i,w}(\{w|(M, w) \models \phi\}) \geq \alpha.$$

We now combine reasoning about knowledge with reasoning about probability. A *Kripke structure for knowledge and probability* is a tuple $(W, \mathcal{K}_1, \ldots, \mathcal{K}_n, p_1, \ldots, p_n, \pi)$. This structure can give semantics to a language with both knowledge and probability operators. A natural assumption in this case is that, in world w, agent i only assigns probability to those worlds $\mathcal{K}_i(w)$ that he considers possible. (However, in some cases this may not be appropriate [3].)

We use the following example from [3] to illustrate a logical approach to reasoning about uncertainty. Alice has two coins, one of which is fair while the other is biased. The fair coin has equal likely hood of landing heads and tails, while the biased coin is twice as likely to land heads as to land tails. Alice chooses one of the coins (assume she can tell them apart by their weight and feel) and is about to toss it. Bob is not given any indication as to which coin Alice chose.

There are four possible worlds:

$$W = \{w_1 = (F, H),\ w_2 = (F, T),\ w_3 = (B, H),\ w_4 = (B, T)\}.$$

The world $w_1 = (F, H)$ says that the fair coin is chosen and it lands heads. We can easily construct two partitions π_{Alice} and π_{Bob} of W, which represent the respective knowledge of Alice and Bob:

$$\pi_{Alice} = \{[w_1, w_2],\ [w_3, w_4]\},$$
$$\pi_{Bob} = \{[w_1, w_2, w_3, w_4]\}.$$

The corresponding equivalence relations R_{Alice} and R_{Bob} can be directly inferred from π_{Alice} and π_{Bob}. In this example, we consider the following four propositions: f - Alice chooses the fair coin; b - Alice chooses the biased coin; h - The coin will land heads; t - The coin will land tails.

We first define a probability distribution $p_{Alice,w}$, according to Alice, for each of the worlds $w \in W$. In world $w_1 = (H, T)$, $p_{Alice,w_1}(w_1) = 1/2$, $p_{Alice,w_1}(w_2) = 1/2$, $p_{Alice,w_1}(w_3) = 0.0$, $p_{Alice,w_1}(w_4) = 0.0$. For world $w_3 = (B, T)$, $p_{Alice,w_3}(w_1) = 0.0$, $p_{Alice,w_3}(w_2) = 0.0$, $p_{Alice,w_3}(w_3) = 2/3$, $p_{Alice,w_3}(w_4) = 1/3$. These definitions are illustrated in Figure 1.

It can be verified that $p_{Alice,w_2} = p_{Alice,w_1}$ and $p_{Alice,w_4} = p_{Alice,w_3}$. Moreover, Bob's probability distributions are the same as Alice's, namely,

$$p_{Bob,w_i} \;=\; p_{Alice,w_1}, \quad i = 1, 2, 3, 4.$$

	Coin	*Lands*	p_{Alice,w_1}	p_{Alice,w_3}
w_1	*fair*	*heads*	1/2	0
w_2	*fair*	*tails*	1/2	0
w_3	*biased*	*heads*	0	2/3
w_4	*biased*	*tails*	0	1/3

Fig. 1. A knowledge system for Alice.

The truth evaluation function π maps $\pi(h, w_1) = true$, $\pi(h, w_2) = true$, $\pi(h, w_3) = false$, $\pi(h, w_4) = false$. Thus, $I(h) = \{w_1, w_2\}$.

It can now be shown that

$$(M, w_1) \models P_{Alice}(h) = 1/2,$$

since $p_{Alice,w_1}(\{w_1, w_3\}) = 1/2$. Similarly,

$$(M, w_2) \models P_{Alice}(h) = 1/2.$$

This means that Alice *knows* the probability of heads is 1/2 in world w_1:

$$(M, w_1) \models \Box_{Alice}(P_{Alice}(h) = 1/2),$$

since $(M, w_1) \models P_{Alice}(h) = 1/2$, $(M, w_2) \models P_{Alice}(h) = 1/2$, and $[w_1, w_2]$ is an equivalence class in R_{Alice}.

The same is not true for Bob. Note that

$$(M, w_1) \models P_{Bob}(h) = 1/2,$$

since $p_{Bob,w_1}(\{w_1, w_3\}) = 1/2$. However,

$$(M, w_3) \not\models P_{Bob}(h) = 1/2,$$

since $p_{Bob,w_3}(\{w_1, w_3\}) = 2/3$. Therefore,

$$(M, w_1) \not\models \Box_{Bob}(P_{Bob}(h) = 1/2),$$

since for instance $(w_1, w_3) \in R_{Bob}$. This says that Bob does *not* know that the probability of heads is 1/2 in world w_1.

5 Rough Sets for Uncertainty Reasoning

Recall that each proposition is represented by a set of possible worlds. The proposition $P_{Alice}(h) = 1/2$, for instance, is represented by

$$I(P_{Alice}(h) = 1/2) = \{w_1, w_2\}.$$

Similarly, proposition $P_{Bob}(h) = 1/2$ is represented by

$$I(P_{Bob}(h) = 1/2) = \{w_1, w_2\}.$$

Recall the following results obtained in a previous section using a logical framework:

$$(M, w_1) \models \Box_{Alice}(P_{Alice}(h) = 1/2),$$
$$(M, w_2) \models \Box_{Alice}(P_{Alice}(h) = 1/2).$$

This knowledge can be expressed using the following proposition in rough sets:

$$K_{Alice}(P_{Alice}(h) = 1/2).$$

By definition, this proposition is represented by the following worlds:

$$\begin{aligned} K_{Alice}(I(P_{Alice}(h) = 1/2)) &= K_{Alice}(\{w_1, w_2\}) \\ &= \{w \in W \mid [w]_{Alice} \subset \{w_1, w_2\}\} \\ &= \{w_1, w_2\}. \end{aligned} \tag{1}$$

This result is consistent with our earlier result that:

$$I(\Box_{Alice}(P_{Alice}(h) = 1/2)) = \{w_1, w_2\}.$$

Even though Bob using the same probability distributions, he is still uncertain as to when the fair coin is used:

$$(M, w_1) \not\models \Box_{Bob}(P_{Bob}(h) = 1/2).$$

The same knowledge (or lack there of) can be expressed using rough sets as:

$$A = I(P_{Bob}(h) = 1/2) \quad = \quad \{w_1, w_2\}.$$

However,

$$\underline{A} = K(A) \quad = \quad \{w \mid [w]_{Bob} \subset A\} \quad = \quad \emptyset,$$

since

$$[w_1]_{Bob} \;=\; \{w_1, w_2, w_3, w_4\} \;=\; [w_2]_{Bob} \;=\; [w_3]_{Bob} \;=\; [w_4]_{Bob}.$$

Finally, let us determine when Alice *knows* that the coin is fair and also *knows* that the probability of heads is 1/2. This sentence is represented in rough sets as:

$$K_{Alice}(f) \wedge K_{Alice}(P_{Alice}(h) = 1/2).$$

Now

$$I(K_{Alice}(f)) = \{w_1, w_2\}.$$

By Equation (1),

$$K_{Alice}(I(_{Alice}(h) = 1/2)) = \{w_1, w_2\}.$$

By the definition of the incidence mapping:

$$\begin{aligned} &I(K_{Alice}(f) \wedge K_{Alice}(P_{Alice}(h) = 1/2)) \\ &= I(K_{Alice}(f)) \cap K_{Alice}(I(_{Alice}(h) = 1/2)) \\ &= \{w_1, w_2\} \cap \{w_1, w_2\} \\ &= \{w_1, w_2\}. \end{aligned}$$

6 Conclusion

Rough sets have primarily been applied to classification problems. Recently, it has been shown that rough sets can also be applied to reasoning about knowledge [5, 7]. In this preliminary paper, we have added probability. This allows us to represent formulas such as "the probability of ϕ, according to player 1, is at least α", where ϕ is a formula and α is a real number in $[0, 1]$. Thus, the only extension to the work in [7] is to allow *formulas* involving probability.

On the other hand, our original objective was to introduce a probability *operator* P in the same spirit as the knowledge *operator* K in [7]. Unfortunately, while P behaves nicely with K, P does not always interact nicely with itself. We are currently working to resolve these problems.

References

1. Bundy, A., Incidence calculus: a mechanism for probability reasoning. *International Journal of Automated Reasoning*, **1**, 263-283, 1985.
2. Fagin, R., Halpern, J.Y., Moses, Y. and Vardi, M.Y., *Reasoning about knowledge.* MIT Press, Cambridge, Mass., 1996.
3. Halpern, J.: A logical approach to reasoning about uncertainty: a tutorial. in: Arrazola, X., Korta, K., and Pelletier, F.J., (Eds.), Discourse, Interaction, and Communication. Kluwer, 1997.
4. Pawlak, Z., *Rough sets - Theoretical Aspects of Reasoning about Data.* Kluwer Academic Publishers, 1991.
5. Salonen, H. and Nurmi, H.: A note on rough sets and common knowledge events. European Journal of Operational Research, 112, 692–695, 1999.
6. Wong, S.K.M., Wang, L.S. and Bollmann-Sdorra, P., On qualitative measures of ignorance. *International Journal of Intelligent Systems*, **11**, 27-47, 1996.
7. Wong, S.K.M.: A rough-set model for reasoning about knowledge. In *Rough Sets in Knowledge Discovery*, L. Polkowski and A. Skowron (Eds.), Physica-Verlag, 276-285, 1998.