Information Sciences 334–335 (2016) 122–143

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Formal reasoning in preference-based multiple-source rough set model☆

Md. Aquil Khan*

Indian Institute of Technology Indore, Indore, India

ARTICLE INFO

Article history:
Received 23 March 2015
Revised 23 October 2015
Accepted 26 November 2015
Available online 2 December 2015

Keywords:
Rough set theory
Lower and upper approximations
Modal logics
First order predicate logic
Axiomatization

ABSTRACT

We propose a generalization of Pawlak's rough set model for the multi-agent situation, where information from an agent can be preferred over that of another agent of the system while deciding membership of objects. Notions of lower/upper approximations are given, which depend on the knowledge base of the sources as well as on the position of the sources in the hierarchy giving the preference of sources. A quantified modal logic is proposed to reason about the properties of the proposed approximations. A sound and complete deductive system for the logic is also presented. Moreover, it is shown how the properties of the proposed approximations can be deduced as theorems of this deductive system.

1. Introduction

In Pawlak's rough set model [24], knowledge is understood as the ability to classify objects, which are taken from a fixed domain U. Any classification is a partition of the domain. Thus, it is based on the simple structure of *approximation space*, consisting of an equivalence relation R over a set U of objects. Objects being in the same equivalence class of R are indiscernible by means of knowledge provided by R. Based on this simple idea, any concept X given as a subset of U, is then approximated from within and outside using the knowledge base R. The union of all equivalence classes $[x]_R$ with $[x]_R \subseteq X$ gives the lower approximation of X, denoted as $\underline{X}_R$, while the union of all equivalence classes having a non-empty intersection with X gives the upper approximation of X, denoted as $\overline{X}_R$. The set $B_R(X) := \overline{X}_R \setminus \underline{X}_R$ denotes the boundary of X. The elements of the sets $\underline{X}_R$, $(\overline{X}_R)^c$, and $B_R(X)$ are respectively called the positive, negative, and boundary elements of X.

A practical source of a Pawlak approximation space is an *information system* [25], formally defined as a tuple $\mathcal{S} := (U, A, \{Val_a\}_{a\in A}, f)$, consisting of a non-empty set U of objects, a non-empty set A of attributes, a non-empty set Val_a of attribute values for each $a \in A$, and $f : U \times A \to \bigcup_{a\in A} Val_a$ such that $f(x, a) \in Val_a$. Any information system $\mathcal{S} := (U, A, \{Val_a\}_{a\in A}, f)$ and $B \subseteq A$ would induce an *indiscernibility relation* $Ind_{\mathcal{S}}(B)$ on U:

$$x\ Ind_{\mathcal{S}}(B)\ y \text{ if and only if } f(x, a) = f(y, a) \text{ for all } a \in B.$$

One can easily verify that $Ind_{\mathcal{S}}(B)$ is an equivalence relation. Thus, given an information system $\mathcal{S}$, and a set B of attributes, we obtain an approximation space $(U, Ind_{\mathcal{S}}(B))$.

Over the years it has been observed that Pawlak's simple idea of rough set theory needs extensions to make it applicable to different practical situations. Therefore, one can find several generalizations of rough set theory in literature. The variable

☆ A part of this paper appeared in preliminary form in the proceedings of the conference RSCTC 2010.

* Tel.: +91 8602365453.
E-mail address: mdaquilkhan@gmail.com

http://dx.doi.org/10.1016/j.ins.2015.11.038
0020-0255/

precision rough set model [35], the rough set model based on covering [26], neighborhood system [21] and tolerance relation [30], the Bayesian rough set model [31], the fuzzy rough set model [6] are a few instances of it. Another extension of Pawlak's rough set model is obtained by considering a natural generalization of approximation space where one is provided with a collection of relations instead of just one. The notion of *information structure* proposed by Orłowska [23] is an instance of such a structure. It is defined as a tuple $(U, \{R_B\}_{B\subseteq\mathcal{A}})$, where $\mathcal{A}$ is a non-empty set of parameters or attributes and for each $B\subseteq\mathcal{A}$, R_B is an equivalence relation on U satisfying,

$$R_\emptyset = U\times U \tag{1}$$

$$R_{B\cup C} = R_B\cap R_C. \tag{2}$$

The condition (1) signifies that we can distinguish objects only using the information about the objects regarding the attributes. With the aim to obtain new properties, some authors (e.g. [8,10]) have also considered operations on the collection of indiscernibility relations. For instance, in [8], the intersection and transitive closure of union of indiscernibility relations are considered.

The collection of equivalence relations over the same domain is also studied in [28]. The multi-agent scenario is investigated where each agent has its own knowledge base represented by an equivalence relation. Thus in [28], a structure of the form $(U, \{R_t\}_{t\in\mathcal{T}})$ is considered, where $\mathcal{T}$ is a set of terms built using a set T of individual agent constants and two binary operations $\wedge$ and $\vee$ such that

- for each $a\in\mathcal{T}$, R_a is an equivalence relation on U,
- for $a, b\in\mathcal{T}$,

$$U|R_{a\vee b} := \{[x]_{R_a}\cap[y]_{R_b} : [x]_{R_a}\cap[y]_{R_b}\neq\emptyset\} \text{ and}$$
$$U|R_{a\wedge b} := \{[x]_{R_a}\cup[y]_{R_b} : [x]_{R_a}\cap[y]_{R_b}\neq\emptyset\}.$$

If R_a and R_b represent the knowledge base of the agents a and b respectively, then $R_{a\vee b}$ and $R_{a\wedge b}$ are respectively called the *strong distributed knowledge base* and *weak distributed knowledge base* of the group $\{a, b\}$ of agents. It is instructive to note that $R_{a\vee b} = R_a\cap R_b$, and hence the notion of strong distributed knowledge base can be identified with the notion of distributed knowledge in epistemic logic [7]. Therefore, from now onwards, we will write distributed knowledge to mean strong distributed knowledge.

In [13–15], rough set theory is again explored in multi-agent scenario, although the more general term 'source' is used there instead of 'agent'. The authors in [13] raised the important issues of counterparts of the standard rough set concepts such as approximations of sets, definability of sets, membership functions in the multi-agent case, and proposed the notions of strong/weak lower and upper approximations. At this juncture, we would like to add that the notions of pessimistic and optimistic approximations proposed by Qian et al. [27] in multi-granulation rough set model are mathematically the same as the notions of strong/weak approximations.

Since the proposals of Khan and Banerjee [13] and Qian et al. [27], there has been a lot of research on the multi-granulation rough set model (e.g. [11,17,19,20,29,32,33]). In this article, we also focus on such a model, but with motivation from the multi-agent situation. At this point, it is pertinent to note that most of the studies of rough set theory under multi-agent system, including the one done in [13], are based on the assumption that each source is equally preferred. The interest of the present work lies in the situation where one source may be preferred over another source of the system in deciding membership of an object. For instance, we could make the assumption that a source will always prefer himself/herself (i.e. his/her knowledge base) over that of the other sources of the system. Thus with this assumption, if we find that $x\in\underline{X}_{R_1}\cap B_{R_2}(X)$ and $y\in\underline{X}_{R_2}\cap B_{R_1}(X)$, R_1, R_2 being the knowledge bases of sources 1 and 2 respectively, then source 1 will put more possibility on x to be an element of X than y. Observe that in the above conclusion, not only the knowledge base of the sources but also the preference of source 1 is playing a role. The current article presents a rough set model where a preference order on the set of sources will also be considered. Moreover, we shall define notions of approximations that depend on the knowledge base of the sources as well as on the position of the sources in the hierarchy giving the preference of sources.

There have been extensive studies on the logics with semantics based on the structures inherited from rough set theory. The modal nature of the lower and upper approximations was evident from the beginning. Hence, it is of no surprise that normal modal systems were focussed upon, during investigations on logics for rough sets. In particular, in case of Pawlak rough sets, the two approximation operators clearly obey all the S5 laws. With the evolution of rough set theory with time, more expressive logics were required to be introduced to reason about approximations (cf. e.g. [2,5]). Consequently, in literature, one can find complete formal systems for reasoning based on rough set theory in multi-agent systems (e.g. [4,13,15,28]). The languages of most of these logics have 'agent-constants', and one or more binary operations are used to build the set of terms. The modal operators are indexed with these terms, which, in turn, are used to capture approximations relative to the knowledge base of individual or groups of agents. A binary relational symbol $\Rightarrow$ on the set of terms is also used in [28]. The expression $s\Rightarrow t$ reflects that "the classification ability of agent t is at least as good as that of agent s". In [13,15], the first order logic feature of quantification is also incorporated into the proposed logic. This enables the logic to capture quantification over the knowledge bases of the systems. With this feature, the logic presented in [13,15] is able to capture the notions of strong/weak approximations proposed in [13]. In this article, we will work on an extension of this logic so that the extended logic will have the expressibility power of the logic proposed in [15], and, in addition to this, it will also be able to capture the preference ordering on the set of agents. This will enable the logic to capture the notions of approximations proposed in this article.