International Journal of Approximate Reasoning 55 (2014) 436–456

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

An update logic for information systems

Md. Aquil Khan [a,c,1], Mohua Banerjee [b,*], Roland Rieke [c]

[a] Indian Institute of Technology, Indore, India
[b] Indian Institute of Technology, Kanpur, India
[c] Fraunhofer SIT, Darmstadt, Germany

ARTICLE INFO

Article history:
Available online 26 July 2013

Keywords:
Information system
Lower and upper approximations
Modal logic
First order predicate logic

ABSTRACT

Updates in a knowledge base, given as an information system in rough set theory, may need to be made due to changes in (i) the set of attributes, (ii) attribute-values, or (iii) the set of objects (instances). In this article, we propose a logic for information systems which incorporates all these three aspects of updates. The logic can capture the flow of information as well as its effects on the approximations of concepts. A sound and complete deductive system for the logic is presented. The decidability issue is also discussed.

1. Introduction

Rough set theory, introduced by Pawlak in the early 1980s [35], offers an approach to deal with uncertainty inherent in real-life problems, more specifically that stemming from inconsistency or vagueness in data. The theory is based on the simple notion of an *approximation space*, considered by Pawlak to be a set W (domain) with an equivalence relation R. In an approximation space, one studies the definability of concepts, represented as subsets of the domain. *Lower* and *upper approximations*, defined with the help of equivalence classes in W due to R, are used to approximate the concepts maximally from 'within' and minimally from 'outside' (respectively).

Applications of rough set theory are mostly based on an attribute-value representation model, called (*deterministic/complete*) *information system*. It is a table which provides information about a domain W of objects regarding a set of attributes/properties. Initially, rough set theory was applied to information systems which are *complete*, in the sense that the value corresponding to each object-attribute pair is assumed to be known. But this need not always be the case: some information may be missing. Thus, to deal with missing information, the notion of *incomplete information system* was proposed (cf. [12,13]), and classical rough set theory has been extended to deal with such information systems. Another extension of rough set theory handles the situation where the 'knowledge base' or the involved information system needs to be *updated* with time, due to inflow of new information. For instance, information currently missing in an incomplete information system may be provided later. Information about new attributes may be obtained ([5,23,26]), new objects may be brought into consideration [39,41,15,24], or information about attributes may change [25,6]. This change in attribute information could be, for example, due to coarsening or refining of attribute values. In this article, our aim is to study rough set theory and information systems in the perspective of updates. In literature, one can find work in this direction (e.g. [39,5,41,15,23–25,6,26,44]), but most of these concentrate on the issue of incremental updates of approximations and decision rules. Here, we approach the issue from the viewpoint of formal logic. Our aim is to present a formal system of logic which can capture updates of information system made due to a flow of information. Moreover, as the approximations of concepts

* Corresponding author.
E-mail addresses: mdaquilkhan@gmail.com (Md.A. Khan), mohua@iitk.ac.in (M. Banerjee), roland.rieke@sit.fraunhofer.de (R. Rieke).

[1] This work was partly done during the author's tenure as an ERCIM "Alain Bensoussan" Fellow. The ERCIM programme is supported by the Marie-Curie Co-funding of Regional, National and International Programmes (COFUND) of the European Commission.

0888-613X/$ – see front matter
http://dx.doi.org/10.1016/j.ijar.2013.07.007

are determined through the information systems, a change in an information system due to inflow of information would possibly affect the approximation of concepts also. Therefore, we would like the logic to be strong enough to capture effects on the approximations of concepts as well.

With the inception of rough set theory, the question of a suitable logic for rough sets and information systems attracted the attention of many researchers. One can find a number of proposals of logics with different features, which can be used to study the properties of concepts related to rough set theory. We refer to [8,2] for detailed surveys on these logics. It is not difficult to observe that in a logic for rough sets and information systems, one would like to have the two following features. (i) The logic should be able to describe aspects of information systems such as attribute, and attribute-values. (ii) It should also be able to capture concept approximations induced by different sets of attributes. In literature, one can find logics with the first feature (e.g. [32,33,42,1]) as well as logics with the second feature (e.g. [31,34,32,33,38,43,10,11]). In [18], a logic with a sound and complete deductive system having *both* features, was introduced. This proposal also paved the way for the introduction of a dynamic logic for rough set theory [19], in the line of dynamic epistemic logic [9]. The logic for rough set theory and information systems was extended by adding information operators carrying information. The notion of update of information system was introduced to give the semantics for the dynamic logic, and reduction axioms were provided to obtain its axiomatization.

The dynamic logic proposed in [19] has two deficiencies. It does not cover information systems where the attribute-values range over an infinite set. It also cannot capture information asking for the introduction or deletion of objects to or from the domain. With time, we may get information about new objects, and introduction of these objects might affect the concept approximations. Moreover, sometimes, we may wish to view only a part of the domain instead of the whole, while taking decisions. For instance, at some point, there may be insufficient information about certain objects with respect to some attributes, and we may wish, at that point, to ignore those objects while making decisions. As evident from the work in [39,41,15,24], any study of rough set theory in the perspective of update cannot be complete without touching these aspects of updates. In the current work, we shall propose an update logic for rough set theory and information systems which overcomes these limitations of the dynamic logic of [19].

The remainder of this article is organized as follows. In the next section, basic concepts related to rough set theory and information systems are given. We also introduce in it, the notion of *extended information system*, which is crucial for the later development. Section 3 introduces the update logic ULIS of information systems. In Section 4, we shall show how the language of ULIS can be used to reason about the properties of rough sets in the perspective of updates. Section 5 presents a deductive system for the logic ULIS, and corresponding soundness and completeness theorems are proved in Section 6. Section 7 raises some decidability issues. Section 8 concludes the article.

2. Information systems

One can generalize the notion of an approximation space in rough set theory, by considering a relation that is not necessarily an equivalence. Thus, for instance, we have the notion of a *tolerance approximation space* [40], where the relation involved is a tolerance relation instead of an equivalence. Given a (generalized) approximation space (W, R), any concept represented as a subset (say) X of the domain W may be approximated by its *lower* and *upper approximations* given as $\underline{X}_R := \{x\colon\ R(x)\subseteq X\}$ and $\overline{X}_R := \{x\colon\ R(x)\cap X \neq \emptyset\}$ respectively. $R(x)$ denotes the set $\{y \in W\colon\ (x, y) \in R\}$.

Information systems induce approximation spaces. We shall be interested here in the following kinds of information systems.

Definition 1. A *deterministic information system* $\mathcal{S} := (W, A, Val, f)$, comprises a non-empty set W of objects, a non-empty set A of attributes, a non-empty set *Val* of attribute values, and an assignment $f : W \times A \to Val$.

Observe that the above notion of deterministic information system gives a complete description of the objects of the domain regarding the attributes. But, as mentioned in the Introduction, we may have a situation where some attribute values for an object may be missing. To indicate the absence of attribute values for an object, a distinguished value $*$ is used. Thus, we have the following definition.

Definition 2. (See [20].) An *information system* is a tuple $\mathcal{S} := (W, A, Val, f)$, where W, A, f are as in Definition 1, and *Val* is a non-empty set of attribute values containing the distinguished value $*$.

An information system which satisfies $f(x, a) = *$ for some $x \in W$ and $a \in A$ is called an *incomplete information system*.

Thus, a deterministic information system can be viewed as an information system with $f(x, a) \neq *$ for all $x \in W$ and $a \in A$.

Note 1. In literature, an alternative way of defining a (deterministic/incomplete) information system involves a collection of attribute-value sets, one for each attribute. That is, an information system is taken to be a tuple $\mathcal{S} := (W, A, Val, f)$, where $Val := \bigcup_{a\in A} Val_a$, and f satisfies the condition $f(x, a) \in Val_a$, for $a \in A$. In this article, we shall use the notion of information system given by Definition 2. In fact, the current work can easily be modified to fit the above-mentioned notion of information system.