G Model
JOCS-641; No. of Pages 9
ARTICLE IN PRESS
Journal of Computational Science xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Journal of Computational Science
journal homepage: www.elsevier.com/locate/jocs

Variable decision knowledge representation: A logical description

Yanhui Zhai [a], Deyu Li [b,*], Jing Zhang [c]

[a] School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China
[b] Key Laboratory of Computational Intelligence, Chinese Information Processing of Ministry of Education, Taiyuan 030006, China
[c] School of Foreign Languages, Shanxi University, Taiyuan 030006, China

ARTICLE INFO

Article history:
Received 9 December 2016
Accepted 23 March 2017
Available online xxx

Keywords:
Variable decision implication
Non-redundancy
Completeness
Inference rules

ABSTRACT

Decision-making is important in all science-based professions, where specialists apply their knowledge to making valuable decisions. In Formal Concept Analysis, decision-making problem is handled within decision contexts and in the form of decision implications. In this paper, we introduce the notion of variable decision implication to generalize decision implications and *uncertain* decision implications. We describe the semantic aspect of variable decision implications by defining the notions of *follow*, *non-redundant*, *complete*, etc., and provide the syntactical description by presenting three inference rules and proving their soundness and completeness. This paper also provides a re-explanation of deduction of associative rules, and should be regarded as a starting point for effectively reducing the size of associative rules.

1. Introduction

In all science-based professions, specialists makes important decisions by removing unnecessary premises, synthesizing the remaining premises, applying their knowledge and then providing one or more choices. Although it is not clear what the psychological mechanism behind decision-making is, there are many inspirational results from the study of Artificial Intelligence [1–5]. In Formal Concept Analysis (FCA)[1] [7,8], a mathematical tool for software engineering [9], machine learning [10], information retrieval [8], social networks [11], cognition-based concept learning [12–16] and knowledge reduction [17–23], decision-making problem is handled in decision contexts and in the form of decision implications [23–34].

Decision implication is a formula of the form $A \rightarrow B$ expressing that each object having all *condition attributes* from A also has all *decision attributes* from B. In practical, decision implication can be extracted from decision contexts [24,26–29] or merely be discussed in the logical way [31,32]. Generally speaking, lots of decision implications may be generated, because there are *redundant* decision implications, which can be deduced from other decision implications. In order to remove such decision implications, Qu et al. [24] presented a special inference rule, called α-decision inference rule, to deduce other decision implications by enlarging premises and reducing consequences of decision implications. Li et al. [26–28,35,29] also studied this inference rule in decision context, incomplete decision context and real decision context based on the decision implications that are generated from constructed concept lattices instead of those that are generated directly from decision contexts [24]. In view of the complexity of generating concept lattices [7], it seems more suitable to generate decision implications from decision contexts rather than from concept lattices. Thus, many researchers adopted the logical way of studying decision implications [24,32,31].

Similar to the logical studies of functional dependency [36], attribute implication [7] and fuzzy attribute implication [37], the study of decision implications can be divided into two parts [32]: semantic aspect and syntactical aspect. The semantic aspect needs to answer the following questions: (1) Soundness of decision implications: is a decision implication is *valid*? (2) Redundancy of decision implications: is a set of decision implications compact? In other words, can a decision implication derive from the others? (3) Completeness of decision implications: can we derive a given set of decision implications from a set of decision implications? (4) Decision implication basis: how to derive a *non-redundant* and *complete* set of decision implications?

In the syntactical aspect, one begins with a set of decision implications and some *inference rules*, and then infers new decision implications from the given set by repeatedly applying the inference rules. This process brings forth the following questions concerning the semantic aspect: (1) Soundness of inference rules: is any decision implication that can be inferred by applying the

* Corresponding author.
[1] Interested readers may find a comprehensive survey on FCA in [6].

http://dx.doi.org/10.1016/j.jocs.2017.03.019
1877-7503/

inference rules valid? (2) Completeness of inference rules: can we obtain all possible valid decision implications by applying the inference rules? (3) Redundancy of inference rules: can we obtain one inference rule from the others?

Zhai et al. [32] described the semantic and syntactical aspects of decision implications. Specifically, in the semantic aspect, they introduced the notions of closure and unite closure, and presented some results to determine the completeness and non-redundancy of decision implications; in the syntactical aspect, they presented two inference rules, (Augmentation) and (Combination), and proved their completeness with respect to semantic aspect. Based on this work, Zhai et al. [33,34] presented the most compact set of decision implications, i.e., decision implication canonical basis (DICB), and proven that DICB is complete, non-redundant and optimal (i.e., DICB contains the least number of decision implications among all complete sets of decision implications) [33,34]. An effective algorithm for generating DICB has also been developed in [25]. In addition, Zhai et al. [31] introduced complete residuated lattice based fuzzy decision implication and constructed the semantic and syntactical aspects of fuzzy decision implications.

Both [31,32], however, only considered the *certain* decision implications that may be crisp or fuzzy; little work has been done on *uncertain* decision implications from a logical perspective. Here, a (crisp) decision implication may be $\{a, b\} \to \{c\}$, meaning that we are sure that if we have attributes a and b, then we have attribute c. A fuzzy decision implication may be $\{0.5/a, 0.5/b\} \to \{1/c\}$, meaning that we are sure that if we have attribute a to the degree 0.5 and attribute b to the degree 0.5, then we have attribute c to the degree 1. An uncertain decision implication may be $(\{a, b\} \to \{c\}, 0.5)$, meaning that we are sure to the degree 0.5 that if we have attributes a and b then we have attribute c; in other words, we are not sure if $\{a, b\} \to \{c\}$ is true.

For comparison purposes, we will also consider associative rules [1,38,39], which are *uncertain attribute implications* that satisfy the constraints of Confidence and Support. Zaki et al. [1] presented some semantic results on associative rules, but left the syntactical aspect untouched. Balcazar [39] found more useful and important results by showing the equivalence of several redundancies [40,41] and proved the completeness of the three inference rules:

rR

$$\frac{A \Rightarrow B,\ B' \subseteq B}{A \Rightarrow B'}$$

rA

$$\frac{A \Rightarrow B}{A \Rightarrow A \cup B}$$

lA

$$\frac{A \Rightarrow B' \cup B}{A \cup B' \Rightarrow B}.$$

Based on the results, Balcazar [39] derived a complete and minimal set of associative rules. In addition, Balcazar [39] also studied a more general redundancy (closure-based redundancy) and presented the complete inference rules and optimum-size basis for closure-based redundancy.

However, since the literature [39] did not take the case of decision into account, they did not provide inference rules for *decision associative rules* (or uncertain decision implications). In fact, when considering uncertain decision implications, both (lA) and (rA) are invalid in the case of decision associative rules, and (rR) is not enough to be complete with respect to the semantic aspect.

This paper intends to extend decision implications [32] to the uncertain case and presents a logical way to describe the semantic and syntactical aspects of uncertain decision implications. Compared with [39], in the semantic side, this approach gives more fundamental definitions and deduction and, thus generalizes the semantic aspect of [39]. In the syntactical side, we provide three inference rules and show their soundness and completeness with respect to semantic aspect. In addition, the work also provides an extensible framework for fuzzy attribute implications [37] and fuzzy decision implications [31], since both of them did not take uncertainty into account.

2. Variable decision implication

We first recall the notions of decision implication and decision context, and then introduce the definition of variable decision implication.

Definition 1 *([32])*. Let C, D be two universes, called the condition set and the decision set respectively. A *decision implication between C and D* is of the form $A \to B$ satisfying $A \subseteq C$ and $B \subseteq D$. In this case, A is the premise and B the consequence.

We denote by $\mathcal{I}(C, D)$ the set of all decision implications between C and D.

Definition 2 *([42,32,24])*. A decision context is defined as a quadruple $K = (G, C, D, I)$ such that $C \cap D = \emptyset$ and $I = I_C \cup I_D$, where G is the object set, C the condition set, D the decision set, $I_C \subseteq G \times C$ the set of condition incidence relations, and $I_D \subseteq G \times D$ the set of decision incidence relations.

Next, we introduce the definition of *variable decision implication*. Different from [39], we define variable decision implication with *valid degree*, a value from a complete lattice, to provide more flexible and fundamental deduction.

Definition 3. Let L be a complete lattice. A *variable decision implication* is of the form $(A \to B, x)$, where $A \to B$ is a decision implication and $x \in L$ is the *valid degree of $A \to B$.*

Note that one cannot say whether a variable decision implication is valid before providing *model* for the variable decision implication. Moreover, to include associative rules as a special case and allow to choose different measures, we introduce the notion of *decision implication function*.

Definition 4. Let L be a complete lattice, $P(C \cup D)$ be the power set of $C \cup D$ and $T \subseteq P(C \cup D)$. A *decision implication function α_T w.r.s T* is a mapping $\alpha_T : 2^C \times 2^D \mapsto L$ such that, for $B_1, B_2 \in 2^D$, $B_1 \subseteq B_2$ implies $\alpha_T(A, B_1) \geq \alpha_T(A, B_2)$, where $\geq$ is the partial order over L.

Remark 1. The mapping α_T also defines an L-fuzzy set on $\mathcal{I}(C, D)$, $\alpha'_T : \mathcal{I}(C, D) \mapsto L$, where $\alpha'_T(A \Rightarrow B) = \alpha_T(A, B)$ is called the *valid degree of $A \to B$ w.r.s α_T* and the pair $(A \Rightarrow B, \alpha'_T(A \Rightarrow B))$ called a *variable decision implication w.r.s α_T*. From the perspective of fuzzy set theory, α'_T can be written as a classical set of variable decision implications w.r.s α_T, i.e.,

$$\alpha'_T = \{(A \Rightarrow B, \alpha'_T(A \to B)) | A \Rightarrow B \in \mathcal{I}(C, D)\}$$

or more often,

$$\alpha'_T = \sum_{A \Rightarrow B \in \mathcal{I}(C,D)} \frac{\alpha'_T(A \Rightarrow B)}{A \Rightarrow B}$$

In this case, $\alpha'_T(A \Rightarrow B)$ is also the membership of $A \to B$ in the L-fuzzy set α'_T.

Decision implication function generalizes some classical measures of associative rules and thus allows to choose different types of valid degrees in applications.

Example 1. Let $L = [0, 1]$ and K be a decision context. Define $T_K = \{g^C \cup g^D | g \in G\} \subseteq P(C \cup D)$, where for $O \subseteq G$, O^C and O^D are defined as follows [32]:

$$O^C = \{m \in C | (g, m) \in I_C, \text{ for all } \ g \in O\}$$