Rough Ontology: Extension of Ontologies by Rough Sets

Syohei Ishizu, Andreas Gehrmann, Yoshimitsu Nagai, and Yusei Inukai

Aoyama Gakuin University,
5-15-1 Fuchinobe, Sagamihara City, Japan 229-8558
ishizu@ise.aoyama.ac.jp, gehrmann@yhc.att.ne.jp,
nagai@ise.aoyama.ac.jp, 5703014@ise.aoyama.ac.jp

Abstract. Ontology is widely used in the areas of knowledge engineering, web based data mining, etc. In rough set theory, accuracy of approximation of set and a concept of granularity are introduced. Rough set theory is very useful to define dependency among attributes and extract decision rules from the set. One of our main aims of this paper is to propose a concept of rough ontology. A concept of rough ontology is extended concept of rough set, and it enables us to use flexible information system in the form of ontology. And rough ontology is useful to introduce concepts rough set theory in to ontology. In this paper we formulate a concept of rough ontology, which is extended concept of rough set theory. We define upper and lower approximation, accuracy of approximation of preference, concept of granularity of preference. And we also show the property of rough ontology.

Keywords: Rough set theory, Ontology, OWL, information system, accuracy.

1 Introduction

With the development of the semantic web, methods for knowledge extraction and knowledge representation are becoming of great importance. However, the existing concepts of an ontology provides no means to capture incomplete knowledge about instances in a domain. Rough set theory is a natural choice for dealing with incomplete information. Incorporating Rough set theory into existing ontology concepts provide possibilities to quantify the degree of accuracy of knowledge.

Rough set theory was proposed by Pawlak [1,2] as an extension of the set theory in order to represent incomplete knowledge. It is useful to represent imperfect knowledge, to mine data, and to analyze attributes dependency. On the other hand ontologies are introduced for complete knowledge representation [9]. In this paper we propose a concept of a rough ontology, which utilizes ideas of both approaches for representation of incomplete knowledge.

In defining concepts usually two approaches are relevant, extensional definition and intensional definition of concepts used in an ontology. An extensional definition of a concept formulates its meaning by specifying its extension, that is, every object that falls under the definition of the concept is directly identified. An intentional definition gives the meaning of a term by giving all the properties required of something that falls under that definition; the necessary and sufficient conditions for belonging

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2007, LNCS 4557, pp. 456–462, 2007.

to the set being defined. We apply rough sets to grasp incomplete information on extential definition of concepts.

2 Concepts of Approximation and Accuracy of Approximation in Rough Set Theory

In this section, we overview the concepts of rough set theory and ontology, and the next section we define the concepts of rough preference. Relationships among the concept of rough set theory and rough preference will be shown in Table 1 in section 3. An information system is defined as 4-tuple $<U, Q, V, f>$ where U is finite set of objects, Q is finite set of attributes, $V = \bigcup\{V_q \mid q \in Q\}$ and $V_q = \{f(x,q) \mid x \in U\}$ is range of the attribute q, and $f : U \times Q \rightarrow V$ is called information function such that $f(x,q) \in V_q$ for every $q \in Q$ and $x \in X$. For every $P \subset Q$, $IND(P) = \{(x, y) \in U \times U \mid \forall q \in Q : f(x,q) = f(y,q)\}$ denotes equivalence relation. For any $P \subset Q$ and $x \in U$, $[x]_P = \{y \mid \forall q \in P : f(x,q) = f(y,q)\}$ denotes an equivalence class. For any $Y \subset U$ and $P \subset Q$, P-upper approximation $P^*(Y)$ of Y, and P-lower approximation $P_*(Y)$ of Y are defined as follows.

$$P^*(Y) = \bigcup\{x \in U \mid [x]_P \subset Y\}$$
$$P_*(Y) = \bigcup\{x \in U \mid [x]_P \bigcap Y \neq \phi\}$$

For every set $Y \subset U$, accuracy of approximation of Y by P is defined as follows.

$$\alpha_P(Y) = card(P_*(Y)) / card(P^*(Y))$$

A set of attributes $P \subset Q$ depends on a set of attributes $P' \subset Q$, denoted $P \rightarrow P'$, if $IND(P) \subset IND(P')$

3 Concepts of Accuracy of Approximation in Rough Ontology and Rules Generation

3.1 Ontology Information System

In this section we introduce the concepts of rough ontology. In Rough set theory an Information system of rough set is based on the decision table, which is theoretically same as relational database. Ontology has more flexible information structure. One of our main aims of this paper is to propose a concept of rough ontology. A concept of rough ontology is extended concept of rough set, and it is enable us to use flexible information system by the form of ontology. And rough ontology is useful to introduce concepts of accuracy of approximation of class, a concept of granularity, defining dependency among attributes, and extraction of decision rules.

Ontology information system is defined as $<U, Q, C, Dom, Range, rel>$. Where U is a finite set of individuals, Q is a finite set of property names. Let C be a finite class of subsets of U. This means if $c \in C$, then $c \subset U$. Each property name has domain, range, and relation, e.g., $Dom: Q \to C$, $Range: Q \to C$, $rel(p) \subset Dom(p) \times Range(p)$ for any $p \in Q$. Note $rel(p)$ may not be function, but relation.

Ontologies are frequently represented by RDF expression. RDF expression of the takes a form of a set of triples $RDF \subset U \times Q \times U$, where $(x, p, y) \in RDF \leftrightarrow (x, y) \in rel(p)$.

In an ontology, class hierarchy is important concept. Class hierarchy is easily represented by the form of $<C, \subseteq>$, where $\subseteq$ is set theoretical inclusion.

In order to compare usual rough set theory with ontology system directly, we introduce some conditions on rough ontology

1) Every property has same domain.
2) Individuals are divided by domain and union of the ranges of properties.
3) Every relation is functional.

Under the conditions, the following notations are introduced. Since every property has same domain, so $U = Dom(q)$ where $q \in Q$. Let $V_q = Range(q)$ where $q \in Q$, then V_q is finite set of individuals. Let $V = \bigcup\{V_q \mid q \in Q\}$. Let $f: U \times Q \to V$ be defined as $f(x, q) = \{y \mid (x, y) \in rel(q)\}$. Since every relation is functional, then f is function. Then $<U, Q, V, f>$ can be regarded as information system.

Conversely rough ontology is free from these conditions, and the conditions 1), 2), 3) may characterize flexible features of ontology information system comparing with information system in rough set theory. In ontology information system, domain of properties may not same. There may be some individuals, which involved in both range and domain of properties. For some $p \in Q$, $rel(p)$ may not be function, and $rel(p)(x)$ may be null or multiple values.

Since usual ontology information system may not satisfy the conditions 1), 2), 3), we can not apply rough set theory directly to ontology system. We need the new rough set theory extended to ontology system. From the definition of ontology system, we introduce ontology information system. From the nature of relation rel, i.e. rel is not functional, we introduce extended information function. $\tilde{f}: U \times Q \to Pow(U)$ next as follows. $\tilde{f}(x, q) = \{y \mid (x, y) \in rel(q)\}$. Where $Pow(U)$ denote power set of U. We define ontology information systems as $<U, Q, C, \tilde{f}>$.

3.2 Rough Ontology and Accuracy of Approximation

Next we introduce extended concepts of rough ontology and accuracy of approximation by a set of properties. $IND(P)$ for any $P \subset Q$, for any $P \subset Q$ and $x \in U$, $[x]_P$, for any $Y \subset U$ and $P \subset Q$, P-upper approximation $P^*(Y)$ of Y , and P-lower approximation $P_*(Y)$ of Y . For every $P \subset Q$, $IND(P) = \{(x, y) \in U \times U \mid \forall q \in Q : \tilde{f}(x, q) = \tilde{f}(y, q)\}$ denotes equivalence relation. For any $P \subset Q$ and $x \in U$, $[x]_P = \{y \mid \forall q \in P : \tilde{f}(x, q) = \tilde{f}(y, q)\}$ denotes an equivalence class.

For any $Y \subset U$ and $P \subset Q$, P-upper approximation $P^*(Y)$ of Y , and P-lower approximation $P_*(Y)$ of Y are defined as follows.

$$P^*(Y) = \bigcup\{x \in U \mid [x]_P \subset Y\},\ P_*(Y) = \bigcup\{x \in U \mid [x]_P \bigcap Y \neq \phi\}$$

For every set $Y \subset U$, accuracy of approximation of Y by P is defined as follows.

$$\alpha_P(Y) = card(P_*(Y)) / card(P^*(Y))$$

A set of attributes $P \subset Q$ depends on a set of attributes $P' \subset Q$, denoted $P \to P'$, if $IND(P) \subset IND(P')$. We extended the concepts of rough set theory, and proposed ontology system and rough ontology, and accuracy of approximation. Table 1 shows the relationships among the concepts of rough set and rough ontology. As shown in Table 1, each concept of rough set theory is naturally extended to rough ontology.

3.3 Rule Generation

Rule generation is one of the important roles in rough set theory. Let Y be $Y \subset U$. If $\alpha_P(Y) = 1$, then rules can be generated and represented by attributes only P . Now we consider the rules of ontology information system. First we define the condition of rule represented by P as $\varphi_P = \{(q, X_q) \mid q \in Q\}$. A set of rules represented by P is defined as R_P . We also define set of individuals which satisfies the condition φ_P as $U_{\varphi_P} = \{x \in X \mid \forall q \in Q : \tilde{f}(x, q) = X_q\}$. Next we define the concept of applicability of rules to Y .

Definition 1
Let Y be a sub set of U . Let R_P be a set of rules represented by P . If the following condition is satisfied then R_P is called applicable to Y .

(1) $\forall \varphi_P \in R_P : U_{\varphi_P} \subset Y$

(2) $\forall y \in Y \bigcap U : \exists \varphi_P \in R_P : y \in U_{\varphi_P}$

Proposition 1

Let Y be a sub set of U, and R_P .is a set of rules. R_P is applicable to Y, if and only if $\alpha_P(Y)=1$.

Table 1. Relationships among the concepts of rough set and rough ontology

Rough set	Rough ontology
Information system $<U, Q, V, f>$ $f: U \times Q \to V$ $V = \bigcup\{V_q \mid q \in Q\}$ $V_q = \{f(x,q) \mid x \in U\}$	**Ontology system** $<U, Q, C, Dom, Range, rel>$ **Ontological information system** $<U, Q, C, \tilde{f}>$ $\tilde{f}: U \times Q \to Pow(U)$
Indiscernibility relation $IND(P) = \{(x,y) \in U \times U \mid$ $\forall q \in P: f(x,q) = f(y,q)\}$	**Indiscernibility relation** $IND(P) = \{(x,y) \in U \times U \mid$ $\forall q \in P: \tilde{f}(x,q) = \tilde{f}(y,q)\}$
Equivalence class $[x]_P = \{y \mid \forall q \in P:$ $f(x,q) = f(y,q)\}$	**Equivalence class** $[x]_P = \{y \mid \forall q \in P:$ $\tilde{f}(x,q) = \tilde{f}(y,q)\}$
Rough set $Y \subset U$ **Upper and lower approximation** $P^*(Y) = \bigcup\{x \in U \mid [x]_P \subset Y\}$ $P_*(Y) = \bigcup\{x \in U \mid [x]_P \cap Y \neq \phi\}$	**Rough ontology** $Y \subset U$ **Upper and lower approximation** $P^*(Y) = \bigcup\{x \in U \mid [x]_P \subset Y\}$ $P_*(Y) = \bigcup\{x \in U \mid [x]_P \cap Y \neq \phi\}$
Accuracy and dependency $\alpha_P(Y) = card(P_*(Y)) / card(P^*(Y))$ $IND(P) \subset IND(P')$	**Accuracy and dependency** $\alpha_P(Y) = card(P_*(Y)) / card(P^*(Y))$ $IND(P) \subset IND(P')$

Table 2. Ontology information system for sneakers and decision attributes

	Color (q_1)	Model (q_2)	Closing mechanism (q_3)	Material (q_4)	Decision attributes(q_5)
sneaker1(x_1)	white	low	lace	Textile	o
sneaker2(x_2)	black	high	lace, Velcro fastener	textile, leather	o
sneaker3(x_3)	colorful	low	ace	textile	×

Table 2. *(continued)*

sneaker4(x_4)	black	low	-	Textile	×
sneaker5(x_5)	white	high	lace	textile, leather	×
Sneaker6(x_6)	colorful	low	-	leather	o

Table 3. Analysis of accuracy of approximation

P	Accuracy of approximation
q_1,q_2,q_3,q_4	1.00
q_1,q_2,q_3	1.00
q_1,q_2,q_4	1.00
q_1,q_3,q_4	1.00
q_2,q_3,q_4	0.67

P	Accuracy of approximation
q_1,q_2	0.67
q_1,q_3	0.67
q_1,q_4	1.00
q_2,q_3	0.33
q_2,q_4	0.33
q_3,q_4	0.17

P	Accuracy of approximation
q_1	0.00
q_2	0.00
q_3	0.17
q_4	0.17

Table 4. Decision rules

P	attribute1(q_1)	attribute2(q_4)
q_1,q_4	white	textile
	black	textile, leather
	colorful	leather

(a) Decision rules for selection good "o"

P	attribute1(q_1)	attribute2(q_4)
q_1,q_4	white	textile, leather
	black	textile
	colorful	textile

(b) Decision rules for selection bad "×"

In Tables 2-4, we show a simple example of rough ontology for sneakers. We show ontology information system (Table 2), the accuracy of approximation (Table 3) and resulted decision rules (Table 4).

4 Summary

One of our main aims of this paper is to present concepts of rough ontology which is an extended concept of from rough set theory. We generalize the concept of an ontology by using rough set theory to represent incomplete knowledge about concepts given by extentional definitions to the concepts of ontology information system, rough ontology, and accuracy of approximation. We define a set of rules which applicable to Y, and show the properties of rough ontology. Proposition 1 shows that if a set of rules R_P is applicable to Y, if and only if $\alpha_P(Y)=1$. We also show an example, which illustrates the concepts of rough ontology. By the example we show how to generate rules is demonstrated.

References

1. Gehrmann, A., Ishizu, S.: Representation of Evaluation Spaces over Attributes Structures - a practical example-. In: Proc. of SCI 2001. Information Systems, vol. 2, pp. 139–145 (2001)
2. Gehrmann, A., Ishizu, S.: Minimal evaluation structure for inconsistent multi-attributes decision-making – selection and addition method of evaluation attributes, vol. 14(1), pp. 85–98 (2005)
3. Ishizu, S.: An Attributes Structure on Evaluation Space. Advances in Systems Study II, 33–38 (1995)
4. Ishizu, S., Amakasu, H.: Basic Concepts and Algorithms for Construction of Evaluation Structures. JASMIN 12(2), 21–36 (2003)
5. Klir, G.J.: Fuzzy Measure Theory. Kluwer Academic, Plenum Publishers (1993)
6. Nicoletti, M.D.C., Uchoa, J.Q., Baptistini, M.T.Z.: Rough relation properties. Int. J. Appl. Math. Comput. Sci. 11(3), 621–635 (2001)
7. Pawlak, Z., Slowinski, R.: Rough set approach to multi-attribute decision analysis. European J. of OR. (72), 443–459 (1994)
8. Pawlak, Z.: Rough set. Int J. Inf. Comp.Sci. 11(5), 341–356 (1982)
9. Wolterink, T.: What is an Ontology? http://referaat.cs.utwente.nl/new/papers.php?confid=1