Ontology Granularity and Rough Equality of Concepts

Pavel Klinov

Applied Computational Intelligence Laboratory
University of Cincinnati
Cincinnati, USA

Lawrence J. Mazlack

Applied Computational Intelligence Laboratory
University of Cincinnati
Cincinnati, USA

***Abstract* - Ontological structures play a major role in the Semantic Web; they became the target of an extensive research over the last decade. An important advance was the attempt to employ Fuzzy Sets and Fuzzy Logic techniques in representing ontologies for intrinsically vague domains of interest where most of human knowledge cannot be expressed in crisp logical formulas. Although the fuzzy approach alleviates the crispness problem, it does not deal satisfactory with the ambiguity caused by deficient discernibility of objects. We consider building rough approximations of fuzzy concepts and relations to measure the roughness of generated ontologies. The key objective of this work is to roughly approximate fuzzy concepts, entailments and sub-sumptions of a given ontology to have a better understanding of both ontology's quality and boundaries of usage. This will be useful for estimating appropriateness of existing ontologies for reasoning in incomplete domains**.

***Keywords* - Ontology, granularity, rough, fuzzy, semantic web**

1. INTRODUCTION

Ontological structures play a major role in the Semantic Web [1] and have been the target of an extensive research over the last decade. An important advance in this area was the attempt to employ Fuzzy Sets and Fuzzy Logic techniques for representing ontologies for intrinsically vague domains of interest where most of human knowledge cannot be expressed in crisp logical formulas. Although the fuzzy approach alleviates the crispness problem it does not deal satisfactory with the ambiguity caused by deficient discernibility of objects. We consider building rough approximations of fuzzy concepts and relations to measure the roughness of generated ontologies. This will be useful for estimating appropriateness of existing ontologies for reasoning in incomplete domains

Ontologies suffering from an inability to be expressed using crisp logical structures (such as OWL, OWL DL, etc) have received a lot of attention recently. It became clear that many domains of interest contain subjective and consequently vague knowledge with the result that rigid logical classification and reasoning is often done poorly. To support inferences based on imprecise ontologies a number of fuzzy extensions have been suggested. One of the significant achievements in this area was the extension of syntax and semantics of SHOIN(D) Description Logic to represent and reason about imprecise concepts [2]. As an example of infer-ence that would be impossible in crisp SHOIN(D) consider:

$$Tall = \exists Height. \geq 170$$

$$Giant = \exists Height. very(High)$$

Here "Tall" is strictly defined as anyone with height greater than 170. Then we define *fuzzy* concept "Giant" using fuzzy predicate "High" in conjunction with *fuzzy modifier* "very" [2]. It's natural to expect $Tall \subseteq Giant$ and f-SHOIN(D) allows that provided High(x) is a fuzzy membership function. Then we may say that $Tall \subseteq Giant$ to some degree that most of people would reasonably be expect to be close to 1.

Although f-SHOIN has proved to deal well with vague concepts, it still assumes that a reasoner possess complete information about individuals to build ABox of assertions [2]. However, in many cases there is a requirement to analyze domains involving both imprecise concepts and ill-described objects. This usually happens when some information about domain objects is missing or is unreliable (for example, patients with a lack of tests taken or if a lost disease history). Consequently, the questions that frequently arise are:
- Can we still use the ontological structures that have been built using datasets of fully described objects?
- Are they still appropriate or can possibly give fallacious results?" And most importantly
- What is the threshold of objects incompleteness after which we may not reason using existing ontologies?

To address these issues we use Rough Sets [3] and Rough-Fuzzy Sets [4] techniques. They have been investigated as an efficient instrument for building approximations of crisp and fuzzy sets [5] respectively when dealing with partially indiscernible objects. In most cases indiscernibility (the central concept in Rough Sets Theory [3]) is brought about by the lack of information which makes impossible to distinguish between certain objects. Partial indiscernibility is also more likely to be a rule rather than an exception in *granulated domains* <U, R> where relation ***R*** can be used to partition the universe of discourse ***U***. Consequently, the key objective of this work is to roughly approximate fuzzy concepts, entailments and subsumptions of a given ontology to have a better understanding of both ontology's quality and boundaries of usage.

The central hypothesis is that an ontology can be built once (automatically, semi-automatically or by domain experts) and then be used for reasoning when data comes from different sources with different degrees of data precision and data completeness. While an ontology might be viewed as an entirely dynamic structure, this approach is sometimes too costly and might lead to conflicting/inconsistent ontologies. This is especially the case when data sources contain incom-plete or mutually contradictory information. Consequently, before reflecting changes in the ontology one might wonder if

they are really needed or existing ontology might still be relevant *to a reasonable degree*. The question of how to measure the relevance of given ontology with respect to partially described objects is of the main interest in the context of this project.

Another interesting aspect of the same hypothesis is the phenomena of *granulated information systems* [6]. The term *granulated* stands for existence of certain partitioning of objects space such that objects are *indiscernible* within partitions (sometimes referred as granules). Granulated information systems are not unusual as an instrument of expressing human knowledge to be represented in the Semantic Web. For example, domains such as geographical knowledge [7] or approximate classification of animals are inherently granulated (because people are not necessarily interested in having *all* details to distinguish objects). Also some other domains can be roughly approximated by granulated object spaces in order to reduce the complexity of reasoning [8] and not only because of incomplete sources of data. Because of those obvious possible applications granulated information calculi [5,6] has been recently studied and our project might be particularly considered as an attempt to employ it for approximating fuzzy ontologies for the Semantic Web.

We believe that recent significant advances in the areas of Rough Sets, Rough-Fuzzy Sets and Granular Computing provide a strong basis for introducing approximations to the field of learning and analyzing ontologies. In our opinion there is a trend in moving ontological descriptions from crisp logical categories to more appropriate fuzzy and approximated structures. As long as fuzzy techniques made possible certain previously unfeasible inferences [2], it seems logical to make the next step now in the direction of rough sets theory and find which new possibilities can be revealed out of it.

2. OBJECTIVES

The objectives of this work can be summarized as follows:

1. Develop extensions for fuzzy SHOIN(D) description logic [2]. Extensions are especially desired for:
   a. *Fuzzy concepts*. Concepts that are central constituents of any ontology have been fuzzified and now can be extended to allow rough approximations. Approximations are particularly interesting for asserting membership of indiscernible objects.
   b. *Fuzzy concepts inclusion axioms*. As previously shown [2] fuzzy concepts can be subsumed by other concepts in some degree which is a function of their structural definitions and fuzzy membership. Subsumption relations can be approximated by low and upper bounds in case of indiscernible objects
   c. *Fuzzy roles inclusion axioms*. The same objective as for concepts inclusion but a little bit less obvious – approximation of role inclusions means establishing bounds in which fuzzy degree of inclusion can vary. More formally: let's consider role inclusions axioms be of the form: $<\alpha \le n>, <\alpha < n>, <\alpha > n>, <\alpha \ge n>$, where α is a SHOIN(D) role inclusion axiom [2]. The objective is to approximate n by a pair <l, u>, $0 \le l \le u \le 1$ basing on the degree of discernibility of domain objects.

   d. *Fuzzy concept and role assertions*. All assertions can also be approximated by inf and sup of fuzzy membership of all objects that are indiscernible with respect to the relation used to partition the universe of discourse.

   The main outcome is the extended syntax of f-SHOIN(D). An important place in the new syntax is going to be occupied by the equivalence relation R that is used to partition the universe. For example, the fuzzy concept "ExpensiveCar" in f-SHOIN(D) will be become fuzzy-rough concept $<\underline{A}, \overline{A}>$ where $\underline{A}, \overline{A}$ - fuzzy sets approximating A- fuzzy set of all expensive cars. Then, if we split all cars on granules and calculate fuzzy membership for all of them, we will be able to measure a degree to which some car can be *definitely* or *possibly* classified as expensive.

2. Develop qualitative measures of roughness of a given ontology for classifying incompletely described objects. In other words we aim at being able to answer the question: "How rough is the ontology" given its fuzzy-rough SHOIN (D) (fr-SHOIN(D) henceforth) and <U,R> - granulated information system.

3. Develop techniques to measure the strength and the nature of the connection between **R** and the roughness of the ontology. This is particularly important because ontology's roughness implicitly reflects its appropriateness for reasoning. Thus a one may decide if she wants to update the ontology or choose a different **R** for partitioning basing of its impact on roughness. We are also interested in computational aspects of measuring roughness as well as in choosing of α(β)-cuts [2] for that.

4. As a consequence of previous objective, we intend to investigate a special case of object space: *object collections*. Object collections imply the existence of a set of attributes describing each object and this is very common way of representing knowledge [9]. Missing attributes is one of the most common reasons for indiscernibility which can cause the decreasing of granularity and therefore increasing of roughness [2] of an existing ontology. Establishing a relation between existence (absence) of certain attributes and the degree of appropriateness of the ontology is also an important goal of this project. As the result we would like to see which attributes weigh more for proper classification using existing ontologies and which less (can potentially be neglected to simplify the analysis)

5. Finally, increasing the roughness of ontology may have an impact on its internal structure as well as on correlations to other ontologies for the same or overlapping domain. It has been shown that fuzzy sets can be considered to be *roughly equal* [2] in a contrast to being *fuzzily equal* when their respected rough approximations happen to be the same. Hence, two previously distinct (and possibly related) concepts might appear to be roughly equal when the degree of roughness increases. Consequently, ontology's designer might decide to merge them in order to reduce the complexity of overall structure. Therefore, our last major objective is to automatically recognize concepts and relations that might potentially be merged (split) when overall roughness increases (decreases).

3. SIGNIFICANCE

The important issue addressed in this work is the insufficient capabilities of existing methods for constructing and maintaining ontologies to deal with imprecise concepts and relationships. There are two important sources of imprecision and vagueness that are inherent in representing of human knowledge.

- First is the imprecision caused be *subjective characteristics of concepts* that were invented by humans and for humans without any though about intelligent machines in mind. The obvious examples given in any books on Fuzzy Logic are such concepts as "tallness", "beautifulness", etc. These issues have been traditionally (since 60s) addressed by Fuzzy Sets and Fuzzy Logic and these techniques are used in ontological languages. [2]
- Second is *"the ambiguity caused by limited discernibility of objects in the domain of discourse"* [5]. For example, any crowd of people always contains people of different heights so a one can classify them as tall or short (it doesn't really matter whether tall and short are crisp or fuzzy characteristics). But if that crowd is observed from 15th floor people are *not* tall or short simply because a one's vision cannot *distinguish* between a tall man and a short man (provided that discernible traits are the same). In this case we say that people are *indiscernible* given their height only. But as long as objects (people) are still to be classified we would like at least to be able to figure out how *rough* our classification will be in the worst case, in the best case and in average.

This project aims at complementing previous work [2] concentrated on bringing fuzzy techniques in the area of the Semantic Web. The positive result of it would open several significant possibilities to both, humans (mostly ontology designers) and intelligent agents (ontology's users):

1. Ontology designers and people responsible for ontologies maintenance (for example, in case of semi-automated modifications) will be able to estimate roughness of ontologies for their particular domains. Depending on the analysis, ontologies might be partially corrected or even fully rebuilt to improve the precision. For example, if measured approximations for a given concept do not meet designer's expectations she may decide to merge it with some other concept or even re-learn it from the new corpus (use other available attributes instead of missing).

2. If the ontology is proved to be reliable and/or is too expensive to rebuild, experts may try to *re-partition* the domain. Roughly speaking if given partitioning leads to *too rough* ontology and the ontology is fixed at that point, the partitioning might be considered to change. For example, if we're unable to make a difference between a tall man and a short man, we may think of using their hair colors as another source of information. And if there's a concept "HairColor" (fuzzy or crisp) people may get classified *less roughly* then using the original ontology. The value of "improveness" is simply the decrease of roughness.

3. Intelligent machines can take advantage of this new technique when choosing between several ontologies for the same information system <U, R>. As long as quality of reasoning is connected with the preciseness of classification (or, annotations of Web-content) they may improve by seeking for the least rough ontology (if available).

Finally, we may expect a positive impact on the performance of reasoning in the Semantic Web because of possible reductions in the complexity of objects descriptions. Roughness of ontologies is one of the main indicators of how reasonable is the level of granularity chosen for the given domain of interest. Sometimes the granularity might be unreasonably fine taking into account too many attributes to differentiate between objects. For some tasks it can be redundant and the jump to less granulated system may lead to a performance speed up. For example, to perform tasks concerned with people's age an agent will unlikely be worried about their skin color, religious beliefs, etc. So after estimating roughness on its ontology it may apparently drop additional information, organize people into bigger granules (age-based categories) and finally accomplish its tasks faster. The answer on the question of *how* faster essentially depends on the nature of the specific task and is beyond the scope of this project.

4. BACKGROUND

There has been an extensive effort on investigating how different soft computing techniques can aid in area concerned with Knowledge Representation, Learning, Pattern Recognition, etc. Most significant and promising results have been achieved in such areas as Fuzzy Logic and Fuzzy Sets [9], Rough Sets [3] and Granulation Calculi [5,6]. In the context of the Semantic Web most influential contributions have been made by fuzzy researches that first showed the need of soft computing in SW and how it can be addressed.

Although the fuzzy model can be applied also for representing *uncertainty* [10] it is not directly related to this project which focuses on vagueness and indiscernibility. Vagueness and subjectivism in representing knowledge is the area which fuzzy theory was originally designed for, so it's only natural that fuzzy extensions for SW started appearing regularly. They are usually concerned with enriching the syntax and semantics of knowledge representation methods to deal sufficiently with imprecise and vague concepts.

The first target to extend was RDF (Resource Description Framework), the recent W3C Candidate for representing knowledge as a collection of facts about resources. Mazzieri [11] extends both RDF syntax and semantics by introducing a positive real number $n \in [0,1]$ such that:

- Any RDF-triple <s, p, o> where s: subject, p: property, o: object is extended to a couple <t, n> where t is the original triple.
- Semantics is extended by using fuzzy interpretations for each RDF-triple. The interpretation satisfies <t, n> iff the degree of membership of both interpretation of the subject and the object belong to the interpretation of the predicate to the degree greater than n.

Although the extension of RDF is an important step towards "fuzzification" of the Semantic Web, it doesn't help much in managing impreciseness in ontologies. More interesting attempts have been concerned with providing fuzzy capabilities to OWL and OWL DL languages that are most frequently used to define ontological structures. Two successful works in this direction:

1. Stoilos [12] provided useful fuzzy extensions to deal with fuzzy concepts and fuzzy axioms (both concepts inclusion and roles inclusion). They have also defined fuzzy

interpretation to adequately expand the semantics of fuzzy DL. Finally issues of reasoning in the case of such fuzzy operations as t-norm (intersection), t-conorm (union) and implication are briefly discussed.

2. Straccia [2] gives almost identical to Stoilos [12] results for f-SHOIN(D) but in addition he defines such fuzzy modifiers as *very, much,* etc. which are frequently used to modify degree of membership functions for given concepts.

Another approach to dealing with the imperfect knowledge was suggested by Pawlak [3] in his Rough Sets Theory in early 80s. The main idea is to measure ambiguity not by using any extra information (like degree of membership in Fuzzy Sets) but by boundary regions of every set of objects. The main difference is that membership of an object to a set is defined not by assigning degree by an expert (agent, machine, software, whatever) but by the portion of objects from the same granule that do belong to the set. Therefore if a one can distinguish between every pair of objects in the universe, all rough sets automatically become crisp. Because original sets can be fuzzy as well as crisp, Rough Sets Theory is really complement to Fuzzy Sets Theory and is obviously not a counterpart of it.

Works focused on taking advantage from both approaches to handling vagueness are Banerjee and Pal [4] and Yao [13] which establish a formal model of approximating a fuzzy set by a pair of rough sets. Roughly speaking, the idea is to build a "pessimistic" and "optimistic" view on approximating a set assuming that all objects are split on equivalence classes. The resulting mathematical model appeared to be capable of defining such terms as *measure of roughness* of a fuzzy set with respect to so called α-, β-cuts that are thresholds of definiteness and possibility in membership of objects to the set [4]:

$\rho_A^{\alpha,\beta} = 1 - \underline{A}_\alpha \div \overline{A}_\beta$ where:

$\rho_A^{\alpha,\beta}$ - measure of roughness of rough-fuzzy set $<\underline{A},\overline{A}>$
w.r.t. to α, β

$\underline{A}_\alpha \equiv \{x \in U; \underline{A}(x) \geq \alpha\}; \underline{A}(x) \equiv \{\underline{A}(X_i); \forall x \in X_i\}$

$\overline{A}_\beta \equiv \{x \in U; \overline{A}(x) \geq \beta\}; \overline{A}(x) \equiv \{\overline{A}(X_i); \forall x \in X_i\}$

$0 \leq \beta \leq \alpha 1$

Furthermore, approximating of fuzzy sets provide enough mathematical basis for defining of *rough equality* of fuzzy sets which means that their lower and upper approximations are respectively equal [4]. And finally Banerjee and Pal [4] show that the relation of rough equality holds for *any* chosen α-, β- thresholds:

$$A \approx B \Leftrightarrow \underline{A} = \underline{B} \wedge \overline{A} = \overline{B} \quad \rho_A^{\alpha,\beta} = \rho_A^{\alpha,\beta} \ \forall \alpha, \beta : 0 \leq \beta \leq \alpha$$

This suggests that we can guess if some fuzzy sets are roughly equal without caring two much about choosing ideal α and β.

These results inspired us to apply rough-fuzzy approximations on fuzzy ontologies.

5. RELATIONSHIP TO OTHER WORKS IN PROGRESS

Other research projects are mostly concentrated on providing fuzzy extensions to traditional crisp sets and logic methods for building ontologies (and knowledge representation in general). Our research can be considered complement to that because we expect our approach to be applied as a generic approximation method to a broad range of possible fuzzifications of ontological descriptions. We briefly outline most promising in our opinion attempts and research directions that are currently under active investigations:

- Straccia [2] and Sanchez et al. [15] are working on further enhancements for his f-SHOIN(D) description logic. One of his next objectives is introducing *fuzzy quantifiers* to use within DL syntax. Quantifier like "most", "usually", "likely" can bring more flexibility for defining complex, intrinsically vague concepts. For example: [2]

  $TopCustomer = Customer \cap (usually) buys.ExpensiveItem$

  $ExpensiveItem = Item \exists \Pr ice.High$

- Banerjee and Pal [4] concentrate of further investigating of the general machinery for approximating. The main direction is to generalize the idea of partitioning the domain of interest and extend it to *fuzzy partitions*. This result (if achieved) can be used in approximating ontologies for domains where objects are *fuzzily granulated.*

- Jensen and Shen [9] are exploring the area of objects collections and computing attributes reducts. Next step in their research is to investigate the issue of *fuzzy reducts* "which would allow attributes to have a varying possibility of becoming a member of the resulting reduct". We are interested in this research because it's important for measuring ontologies roughness when the corpus can be represented as an object set.

To sum up, these attempts aim at providing and extending of the mathematical model for handling vague concepts. From this point of view our work cannot be considered as a competitor to them. Instead we may extend the idea of rough approximation to deal with many fuzzy novelties in ontologies. For example, when fuzzy quantifiers come up we will be ready to provide methods for approximating fuzzy-rough concepts involving quantifiers in their definitions. This particular feature can be plugged in our method if fuzzy membership function for, say, "TopCustomer", would take into account fuzzy quantifier "usually". The same is true for other relevant works, like fuzzy partitions which would give us more freedom in choosing a proper partition of the domain to make ontologies as rough as we need.

6. PLAN

We will use Rough Sets Theory (RST) and Fuzzy-Rough Approximation Techniques (FRAT) as the main machinery for achieving declared objectives. The starting point for this project is extending f-SHOIN(D) to enable rough approximations and rough reasoning with ontologies. This is mandatory goal and only after it has been achieved we may come to advanced issues of analyzing ontologies with respect to its rough description and specific properties of certain domain of knowledge. Therefore, the outline of our approach and methodology for attaining project's goals can be summarized as following:

- Provide fuzzy-rough extensions to f-SHOIN(D) and run experimental classifications of individuals in a given dataset. We may use any available ontology (such as OpenCyc[1] or any ontology created by automated tool similar to Text2Onto[2]) for testing approximations of its concepts and axioms with respect to an equivalence relation. The relation used to partition the domain (we call it equivalence

relation) can be indistinguishability, similarity, proximity or functionality [14]. First we will try simple equivalence relation based on *attribute reductions* [9]. Taking this approach we assume that all [1] objects' attributes can be reduced to a subset used to distinguish between objects:

x, y – two objects represented as n-tuples (attributes and their values)

$$x = \bigcup_i < a_i, c_i >, y = \bigcup_j < a_j, c_j >; x, y \in D \subseteq U$$

C − set of all attributes. $R \subseteq C$

x, y are partially indiscernible, if $x_R E y_R$,

but whether $x_C E y_C$ is unknown

- Under this assumption we will measure the roughness of the ontology that has been built by learning from a corpus of fully described objects (with all attributes). To do so we need:

  1. Split objects space on partitions (granules with respect to the equivalence relation.
  2. Compute lower and upper approximations of every concept (or most valuable concept for a particular classification). This implies computing of lower and upper approximations of membership degree for every object in the corpus basing on the membership of other objects in the same granule.
  3. Choose α- and β-cuts ($0 \leq \beta \leq \alpha 1$) to calculate the measure of roughness. Roughness $\rho_A^{\alpha,\beta}$ depends greatly on α, β, it can be shown [5] that if β is kept fixed and α increases, $\rho_A^{\alpha,\beta}$ decreases (and vice versa). So by varying α, β we may notice how roughness changes for given domain and given equivalence relation that would help us establish formal definition of how good are α, β for a specific problem.

- The next step in investigating roughness of ontology is to investigate its correlation with the equivalence relation. First, we intend to experiment with different subsets of attributes used to distinguish between objects. This task is somewhat similar to the well-known problem of finding *reducts* [9]. Reduct R is a subset of the set of all attributes C such that objects can be uniquely identified basing on R only. This problem has been extensively studied in machine learning and has many applications (for example, in relational databases). In our case, we do not really look for reducts but rather would like to know how strong is the influence of certain attributes on roughness of the ontology. Nonetheless we can still use techniques such as QuickReduct [9] to compute the minimal reducts and then investigate their subsets. In addition we propose using QuickReduct-like methods to compute so called *core* – intersection of all reducts for the given information system. We expect that the attributes involved in the core are going to be the most valuable with respect to the roughness of the ontology for the given system and would like to get experimental results substantiating this expectation.

- The final experiments in this project will be concerned with *rough equality* of concepts in the ontology. As stated

above we expect that roughness will increase as objects get less distinguishable (so partitions host more objects). It easy to see that in ultimate cases rough equality of any two concepts is intuitive:

o All objects are distinguishable and no granule contains more than one object. Therefore roughness of any single concept is 0 (which is the roughness of any crisp set) and therefore concepts are either equal (if objects' degree of membership coincides in both) or not (if the degree of membership doesn't coincide for at least one object).

o No objects are discernible (all domain is viewed as one gigantic granule). Then all concepts are equal because this granule is lower and upper approximation of any concept.

Between these extreme cases we may compute the degree of subsumption of one concept by another and base our judgment about equality on that. By varying the equivalence function we intend to investigate if it is possible to find a good tradeoff between fine-granularity of objects of universe and the complexity of reasoning with a given ontology. For example, different intelligent agents can use the same ontology for different datasets and different purposes. Then it may appear reasonable for one of them to reduce the complexity of the ontology by merging concepts if they become roughly equal due to limited discernibility of objects the domain. Furthermore even if concepts are not equal it is worth to see how possible changes in the equivalence relation would affect potential equality of certain concepts and consequently the quality of reasoning. Semi-supervised (we do not expect fully autonomous agents to be able to make such decisions) agents may get back to domain experts asking for the permission to re-partition the domain, make the corresponding changes in the ontology and then return to the search for a solution. But they are responsible for presenting convincing arguments (based on computation of roughness and possible equalities/inequalities) that it *might be worth* to alter the equivalence relation.

7. SUMMARY

Ontological structures play a major role in the concept of the Semantic Web; they became the target of an extensive research over the last decade. An important advance in this area was the attempt to employ Fuzzy Sets and Fuzzy Logic techniques for representing ontologies for intrinsically vague domains of interest where most of human knowledge cannot be expressed in crisp logical formulas. Although the fuzzy approach alleviates this problem, it does not deal satisfactory with the ambiguity caused by deficient discernibility of objects. We are considering building rough approximations of fuzzy concepts and relations to measure the roughness of generated ontologies. The key objective is to roughly approximate fuzzy concepts, entailments and subsumptions of a given ontology to have a better understanding of both ontology's quality and boundaries of usage. This will be useful for estimating appropriateness of existing ontologies for reasoning in incomplete domains.

The central hypothesis is that an ontology can be built once (automatically, semi-automatically or by domain experts) and then is used for reasoning when data comes from different

---

[1] http://opencyc.org
[2] http://ontoware.org/projects/text2onto

sources with different degree of data precision and data completeness. While an ontology might be viewed as an entirely dynamic structure, this approach is sometimes too costly and might lead to conflicting/inconsistent ontologies.

While an ontology might be viewed as an entirely dynamic structure, this approach is sometimes too costly and might lead to conflicting/inconsistent ontologies. Another interesting aspect of the same hypothesis is the phenomena of granulated information systems. The term *granulated* stands for existence of certain partitioning of objects space such that objects are *indiscernible* within partitions (sometimes referred as granules). Granulated information systems are not unusual as an instrument of expressing human knowledge to be represented in the Semantic Web.

We believe that recent significant advances in the areas of Rough Sets, Rough-Fuzzy Sets and Granular Computing provide a strong basis for introducing approximations to the field of learning and analyzing ontologies.

The important issue addressed in this work is the insufficient capabilities of existing methods for constructing and maintaining ontologies to deal with imprecise concepts and relationships. There are two important sources of imprecision and vagueness that are inherent in representing human knowledge:
• Imprecision caused be subjective characteristics of concepts.
• Ambiguity caused by limited discernibility of objects in the domain of discourse

REFERENCES

[1] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, vol. 284, pp. 34-43, 2001

[2] U. Straccia, "Towards A Fuzzy Description Logic For The Semantic Web," ISTI-CNR, Pisa 2004

[3] Z. Pawlak, "Rough Sets," *International Journal Of Computer And Information Sciences*, vol. 11, pp. 341-356, 1982

[4] M. Banerjee and S. K. Pal, "Roughness Of A Fuzzy Set," *Information Science*, vol. 93, pp. 235-246, 1996

[5] L. A. Zadeh, "Fuzzy Sets," *Information And Control*, vol. 8, pp. 338-353, 1965.

[6] Y. Yao, "Information Granulation And Rough Set Approximation," *International Journal of Intelligent Systems*, vol. 16, 2001

[7] M. Duckham, K. Mason, J. Stell, and M. F. Worboys, "A Formal Ontological Approach To Imperfection In Geographic Information," presented at GIS Research UK National Conference, York, 2000

[8] Q. Liu, "Granules And Reasoning Based On Granular Computing," presented at IEA/AIE 2003, 2003

[9] R. Jensen and Q. Shen, "Fuzzy-Rough Attributes Reduction With Application To Web Categorization," *Fuzzy Sets And Systems*, vol. 141, pp. 469-485, 2004

[10] M. Holi and E. Hyvönen, "A Method For Modeling Uncertainty In Semantic Web Taxonomies," presented at WWW (Alternate Track Papers & Posters), 296-297, 2004

[11] M. Mazzieri, "A Fuzzy RDF Semantics To Represent Trust Metadata," presented at 1st Italian Semantic Web Workshop on Semantic Web Applications and Perspectives (SWAP), Ancona, Italy, 2004

[12] G. Stoilos, G. Stamou, V. Tzouvaras, J. Z. Pan, and I. Horrocks, "Fuzzy OWL: Uncertainty And The Semantic Web," presented at International Workshop On OWL: Experience And Directions (OWL-ED2005), 2005

[13] Y. Y. Yao and Y. H. Chen, "Rough Set Approximations In Formal Concept Analysis," presented at 2004 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2004), 2004

[14] Y. Y. Yao, "Granular Computing Using Neighborhood Systems," presented at Advances in Soft Computing: Engineering Design and Manufacturing, The 3rd On-line World Conference on Soft Computing (WSC3), 1998

[15] D. Sanchez and G. B. Tettamanzi, "Generalization Quantification In Fuzzy Description Logic," in *8th Fuzzy Days*. Dortmund, Germany, 1996