A Semantic Architecture for Knowledge Intensive CBR through Fuzzy Ontologies

Panos Alexopoulos[1, 3], Manolis Wallace[2], Konstantinos Kafentzis[1] and Dimitris Askounis[3]

[1] IMC Technologies S.A., 360A Kifissias Str. & 2 Marathonos Str., 15233, Athens, Greece, {palexopoulos, kkafentzis}@imc.com.gr
[2] Department of Computer Science and Technology, University of Peloponnese, End of Karaiskaki St., 22100, Tripolis, Greece, wallace@uop.gr
[3] School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou Str., 15773, Athens, Greece, askous@epu.ntua.gr

Abstract. In this paper we present a semantic architecture that allows developers of knowledge intensive CBR systems to effectively deal with the problem of knowledge imprecision and vagueness through the utilization of fuzzy ontologies. The architecture is based on a novel knowledge intensive CBR framework, called IKARUS-CBR, which we presented in previous work and which addresses the problem of knowledge imprecision through the integration of Fuzzy Algebra in the Ontology-Based CBR paradigm. Along with the architecture we also present a comprehensive methodology that defines the necessary steps for building and using CBR systems that follow the IKARUS-CBR framework.

Keywords: Case Based Reasoning, Fuzzy Ontologies

1 Introduction

Knowledge intensive CBR systems [1] are CBR systems that utilize, in some way, additional background knowledge about the domain in order to increase their effectiveness in terms of case retrieval. A typical KI-CBR system comprises some knowledge representation formalism that enables the explicit modelling of case and domain knowledge and some case similarity measure that utilizes this domain knowledge for retrieving cases. This combination enables the matching of cases to be performed based on the similarity of the cases' meanings (semantic similarity) which leads to more accurate and intuitive results compared to more superficial approaches which usually determine the syntactic similarity between the cases' features.

A prominent KI-CBR approach that has been proposed in the research community is Ontology-Based CBR [5] [10]. The distinctive characteristic of this approach is that its knowledge representation and reasoning mechanism is derived from the area of Ontologies. Ontologies have been developed and investigated for some time in Artificial Intelligence as the main way of facilitating knowledge sharing and reuse [11]. Technically speaking, an ontology is a formal description of the entities, relationships, and constraints that make a conceptual model. Depending on the expressiveness and the degree of formality of the underlying representation language, an ontology can range from a simple taxonomic hierarchy of concepts to a logic program utilizing first-order logic, modal logic, or description logics.

One advantage of using ontologies in developing KI-CBR systems is related to the process of domain knowledge acquisition. The latter is usually the bottleneck of the whole development process and, to that end, methods and techniques derived from the ontological engineering community [6] can help to make the acquisition process more efficient. Moreover, the reasoning capabilities that ontologies provide enhance significantly the effectiveness of the case retrieval process and the ontology-based representation of cases itself enables their reuse and adaptation in a variety of application scenarios.

Nevertheless, an issue that the Ontology-Based CBR paradigm has not yet addressed is that of knowledge imprecision. As Zadeh [12] argues, much of the world knowledge, namely the knowledge which humans acquire through experience, communication and education is perception-based and thus subject to imprecision and inaccuracy. Such knowledge, when not treated in some suitable way that is able to consider and convey its inherent imprecision, usually leads to poor effectiveness of the knowledge-based systems that use it. Current ontology-related tools and techniques cannot handle imprecision as they are mostly based on bivalent logic. Therefore new tools drawn from the area of Fuzzy Logic are needed.

To that end, we have developed, in previous work [3], a novel KI-CBR framework, called IKARUS-CBR (Imprecise Knowledge Acquisition Representation and Use), which can handle and exploit imprecise knowledge through the effective integration of Fuzzy Algebra in the ontology-based CBR paradigm. The approach we have followed differs from other Fuzzy CBR approaches [9] in that it uses ontologies as the "vehicle" for the introduction of fuzzy semantics to CBR. This difference, as shown in [3], makes our approach more effective and complete as far as management and exploitation of imprecision is concerned, the main reason being that traditional fuzzy CBR systems do not handle and exploit fuzzy ontological relations.

Our goal in this paper is not so much to describe the IKARUS-CBR framework (this has been already done in [3]) but rather to present the architecture that realizes it and the way this architecture may be practically used for building and using fuzzy ontology CBR systems. With that in mind, the structure of the rest of the paper is as follows: In the next section we briefly describe the main components of IKARUS-CBR, focusing on the knowledge representation and retrieval mechanisms it provides. Subsequently, in section 3 we describe the functional and technical architecture of the platform that implements these mechanisms and we provide a methodology for building and using, by means of the platform, CBR systems that utilize imprecise and vague knowledge. Finally, in section 4 we list our concluding remarks and outline potential future work.

2 The IKARUS-CBR Framework

The integration of Fuzzy Algebra in Ontology-Based CBR is performed in IKARUS-CBR in two levels, the first having to do with the representation of imprecise knowledge itself and the second with the latter's exploitation for case retrieval. In the first level, the framework supports the representation of imprecise case- specific and domain-specific knowledge through a comprehensive fuzzy ontology framework while, in the second level, the retrieval of cases is enabled by a highly customizable fuzzy semantic similarity framework. The following paragraphs summarize the key features and characteristics of these two frameworks.

2.1 Case Representation

The representation framework is based on the premise that in Ontology-Based CBR cases are represented as concept instances and their attributes as ontology relations or properties. The values the relation attributes may take are instances defined within some domain ontology. Based on this, the representation of vagueness within a case is facilitated in two ways: i) by allowing the relation attribute values to be derived from some fuzzy domain ontology and ii) by allowing case attributes to be defined as fuzzy ontology relations or fuzzy ontology properties.

A fuzzy ontology can be defined as an ontology that comprises, in addition to concepts, relations and properties, also **Fuzzy Relations**, **Fuzzy Properties** and **Fuzzy Valued Properties**. A fuzzy relation is a relation that can be established between two concept instances and whose meaning may, in a given domain, context or application scenario, be interpreted as imprecise or vague (e.g. *isExpertAt*, *isNearTo* etc.). Any pair of instances that is linked through a fuzzy relation is accompanied by some fuzzy degree (real number between 0 and 1) which denotes the degree to which the relation should be considered true for the pair (e.g. *John is an expert in Knowledge Management to degree 0.8*). In a similar fashion, a fuzzy property is a concept property whose implied relation between instances and literal values may, in a given domain, context or application scenario, be interpreted as imprecise. As in fuzzy relations, a fuzzy property's literal value for a given instance is accompanied by some fuzzy degree.

Finally, a fuzzy valued property is a concept property whose potential values may be expressed through imprecise terms (e.g. tall, high, fierce etc.). The representation of such a property within a fuzzy ontology is typically performed through the fuzzy logic notion of Fuzzy Linguistic Variable [8]. Such a variable defines the linguistic terms that may act as the property's imprecise values and maps each term to a fuzzy set that defines its meaning by assigning to each of the property's precise values a fuzzy degree. This degree practically indicates the extent to which the precise value and the imprecise term should be considered to express the same thing.

In a more formal way, a Fuzzy Ontology may be defined as a tuple $O_F = \{C, I, FHR, FAR, FP, FLV, FVP\}$ where:

- C is a set of concepts.
- I is a set of instances. Each instance belongs to at least one concept.
- FHR and FAR are sets of fuzzy hierarchical and fuzzy associative relations.

Each fuzzy relation $fr \in \{FHR \cup FAR\}$ is a function $I^2 \rightarrow [0,1]$.

- FP is a set of fuzzy properties. Each fuzzy property $fp \in FP$ is a function $I \rightarrow F(X)$, $F(X)$ being the set of all fuzzy sets in the universe of discourse X.
- FLV is a set of fuzzy linguistic variables. Each $flv \in FLV$ is a tuple $\{u, T, X, m\}$ in which u is the name of the variable, T is the set of linguistic terms of u that refer to a base variable whose values range over a universal set X and m is a semantic rule that assigns to each linguistic term $t \in T$ its meaning $m(t)$ which is a fuzzy set on X.
- FVP is a set of fuzzy valued properties. Each fuzzy valued property $fvp \in FVP$ is a function $I \rightarrow T$ where T is the set of the linguistic terms of a fuzzy linguistic variable $flv \in FLV$.

We can now define a Fuzzy Case Ontology as a subset of a fuzzy ontology where:

- $FCT \subseteq C$ is a set of fuzzy case types.
- $FC \subseteq I$ is a set of fuzzy cases. Each fuzzy case belongs to at least one fuzzy case type.
- $FRA \subseteq \{FHR \cup FAR\}$ is a set of case attributes defined as fuzzy relations. Each $fra \in FRA$ is a function $FC \times I \rightarrow [0,1]$.
- $FPA \subseteq FP$ is a set of case attributes defined as fuzzy properties. Each $fpa \in FPA$ is a function $FC \rightarrow F(U)$, $F(U)$ being the set of all fuzzy sets in the universe of discourse.
- $FVPA \subseteq FVP$ is a set of case attributes defined as fuzzy valued properties. Each $fvpa \in FVPA$ is a function $FC \rightarrow T$ where T is the set of the linguistic terms of a fuzzy linguistic variable $flv \in FLV$.

It should be noted that, since a crisp ontology is a special case of a fuzzy ontology in which all relation and property degrees are equal to 1, the above formalization retains the characteristics of the traditional ontology-based CBR paradigm. This means that all relevant methods and techniques that have already been developed for this paradigm are applicable within the IKARUS-CBR framework as well.

2.2 Case Retrieval

Case retrieval in CBR systems is performed through the assessment of the similarity between the requested case and those stored in the case base. This assessment is performed by determining their partial similarities based on the values of each of their attributes and then aggregating these similarities into a single case similarity score. The exact way these partial similarities are calculated depends on the nature and purpose of the attributes and particularly on the type and range of values these may take.

In IKARUS-CBR, the different types of attributes a case may have are three, namely Fuzzy Property Attributes, Fuzzy Valued Property Attributes and Fuzzy Relation Attributes. From these, only the second type has been the subject of other relevant works and in particular of efforts in the area of Fuzzy CBR and Fuzzy Decision Making [7]. Indeed, the common characteristic of these efforts is that they all use attributes with fuzzy values and a fuzzy pattern matcher for case similarity assessment. Yet, these approaches are incomplete since, as was shown above, imprecision in ontology-based CBR may be manifested in other ways as well.

To address these issues, we have developed in [3] methods for calculating, for each of these types, the similarity between the values they may take. The key characteristic of these methods is that they regard similarity as an application-specific and highly subjective notion that cannot be effectively assessed without taking in mind the application context. Therefore, the methods are accompanied by a context model which, when parameterized appropriately, can adapt the case retrieval process to the particular similarity requirements of the application scenario.

In particular, the idea of similarity subjectiveness is central to IKARUS- CBR, especially in regard to the similarity between ontology entities. In non-fuzzy ontology-based CBR, this similarity is generally assessed by utilizing the components of the domain ontology and particularly the relations that connect the entities through various semantic similarity measures. In our approach, where imprecision is important, instance similarity is also influenced by the imprecision contained in the relations that connect them, therefore, the fuzzy degrees of these relations need to somehow participate in the similarity assessment process a well. Nevertheless, determining which fuzzy ontology relations, in what way and to what degree should participate in the assessment of instance similarity is a highly subjective and application-dependent task as, in different application scenarios and among different users, the contribution of the same fuzzy relation to the similarity between two entities might be totally different [2].

This difference is what the IKARUS-CBR case retrieval framework is able to i) capture through a context model that models the expected role of the fuzzy ontology's relations in the entity similarity assessment process and ii) exploit through an ontology entity similarity assessment algorithm that utilizes the information contained in the fuzzy ontology and the context model. Due to space limitations, the user is referred to [3] for a detailed description of the context model and the similarity assessment algorithm.

3 The IKARUS Platform

The IKARUS Platform implements and provides to users the case and knowledge representation and retrieval capabilities of the IKARUS-CBR framework, described in the previous section. In the following paragraphs we provide a high-level description of the platform’s functional architecture and we outline a stepwise methodology for using its capabilities for implementing fuzzy ontology CBR systems.

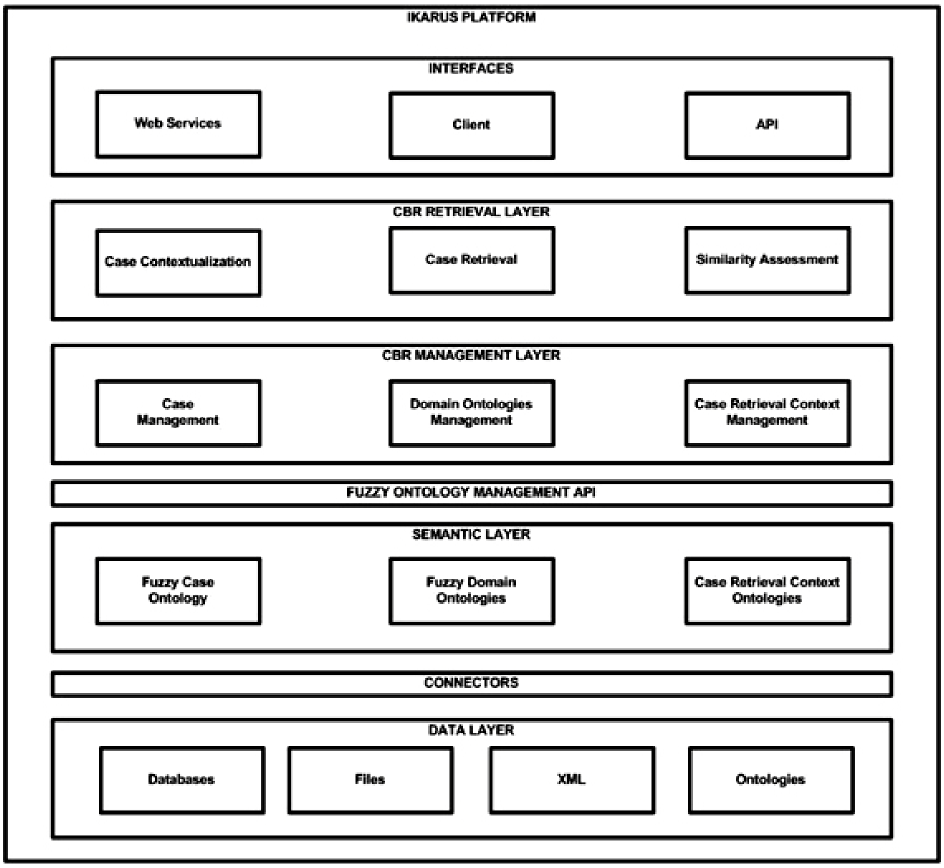

Fig. 1. IKARUS Platform Architecture.

3.1 Architecture

The capabilities that the IKARUS-CBR framework supports are delivered through the layered architecture of figure 1. Each of these layers is analyzed as follows:

- **Data Layer:** The data layer is responsible for the storage and low-level manipulation of the platform's content, namely cases and domain knowledge. This content may have the form of ¯les, databases, XML Schemas and/or ontologies. For each type, appropriate connectors are used in order to make the content accessible to the upper layers.
- **CBR Representation Layer:** The CBR representation layer comprises three ontological schemas through which the unified representation and processing of the platform's content is made possible. The first two schemas, namely the Fuzzy Domain Ontology and the Fuzzy Case Ontology, enable the

representation of fuzzy domain knowledge and of fuzzy cases respectively, according to the formalization of paragraph 2.1. The third is used for representing the context model for the retrieval of cases, as defined in [3]. All schemas are expressed in the Web Ontology Language (OWL) [4].

- **CBR Management Layer:** The CBR Management layer provides the necessary functionality for the comprehensive management of the ontological models of the previous layer. In particular, the Ontology Manager, the Case Manager and the Context Manager modules enable the definition of new (and the manipulation of existing) fuzzy domain knowledge models, fuzzy case schemas and case retrieval context models respectively.
- **CBR Retrieval Layer:** The CBR Retrieval layer provides the necessary functionality for the retrieval of cases from the platform's case base. In particular, the Similarity Calculator module defines the similarity measures that are to be used during the case retrieval process while the Case Retrieval Contextualizer module in turn is responsible for adapting these measures according to the parameters of the context model. Finally, the Case Retriever module provides high-level case retrieval functionality such as query structuring, retrieval execution, result manipulation etc.
- **Interfaces:** The Interfaces layer provides access to the functionality of the two previous layers (CBR Management and CBR Retrieval) through a comprehensive and well-documented API, through web services and through a web-based graphical user interface. The latter is particularly appropriate for the tasks of case and knowledge modeling while the API and the web services allow the integration and interoperability of the platform to any external system. Figure 2 is a snapshot of the platform's graphical user interface and depicts the editor for the management of cases (please note it is still under development).

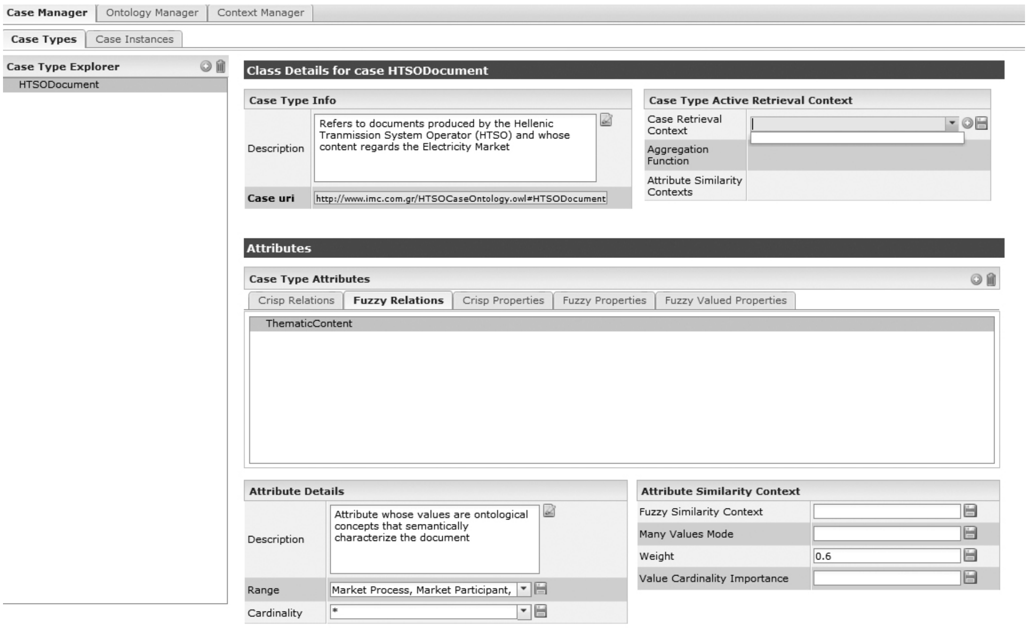

Fig. 2. IKARUS Platform Case Manager.

3.2 Usage Methodology

The typical steps required for designing and using a new fuzzy ontology CBR system in the IKARUS platform are depicted in figure 3 and are the following:

- **Define the (fuzzy) case schema:** The case schema defines the structure of the cases (case types, attributes) according to the formalism of section 2.1 and its creation is facilitated by the Case Manager module. Fuzziness is optional as, apart from fuzzy attributes, crisp ones may be defined as well.
- **Develop or import the (fuzzy) domain ontology:** The platform's domain ontology represents knowledge about the application scenario's domain and consequently describes the cases' semantic content. Such an ontology is required when the case schema includes attributes defined as ontology relations, either fuzzy or crisp (as attributes' values are instances of concepts of the ontology). Also, depending on the domain, the ontology might or might not be fuzzy (following section's 2.1 formalism in case it is fuzzy) and it can be developed from scratch or by reusing existing ontologies. The required functionality for that is provided by the Ontology Manager module.
- **Populate the case base:** This step refers to the creation of the cases (if not already available) and the assignment of values to their attributes. This can be done manually, through the Case Manager module, or

(semi)automatically through semantic annotation techniques that the platform supports. In both cases, the appropriate connectors with the storage medium of the cases need to be determined and parameterized.

- **Assign values to the parameters of the case retrieval context:** The context parameters regard various aspects of the case retrieval process including attribute weights, similarity measures and aggregation functions. Of particular importance are the parameters that "adjust" the behavior of the domain ontology in the case similarity assessment process, as suggested in paragraph 2.2. The detailed description of the role and meaning of these parameters as well as a comprehensive example of their how they influence similarity assessment may be found in [3].
- **Perform case retrieval queries:** Having performed the previous four steps, the platform is ready to be used for case retrieval. This is done through the Case Retriever module which provides the required functionality for construction of query cases and retrieval of similar to them cases.

Finally, it is noted that all the above steps can be performed through all the available interfaces, namely the API, the Web Services and the Graphical Editor.

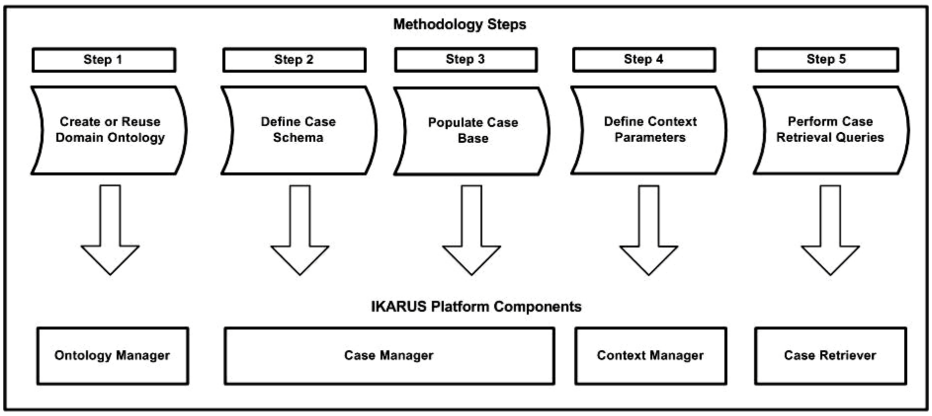

Fig. 3. Methodology for Building CBR Systems with the IKARUS Platform

4 Conclusions and Future Work

In this paper, we presented a semantic architecture that implements a novel knowledge intensive CBR framework, called IKARUS-CBR, that addresses the problem of knowledge imprecision and provides the necessary tools for handling and exploiting the latter in a comprehensive and effective way. In particular, the approach that IKARUS-CBR follows involves the integration of Fuzzy Algebra in the Ontology-Based CBR paradigm at the levels of knowledge representation and knowledge-based case retrieval by means of two frameworks; a fuzzy ontology framework for imprecise case-specific and domain-specific knowledge representation and a fuzzy semantic similarity framework for case retrieval based on this knowledge. The architecture implements these two frameworks in a unified and seamless way for the user and provides all the necessary functionality for applying them in real application scenarios. The exact way the architecture can be used is described by a comprehensive methodology that includes all the necessary steps for building CBR systems according to the IKARUS-CBR framework. The key benefits that the IKARUS platform provides to its users can be summarized as follows:

- It enables users to model a domain's imprecision and vagueness in a much more expressive way than traditional fuzzy CBR systems through the use of fuzzy ontologies.
- The case retrieval mechanism of the platform is highly customizable due to the IKARUS-CBR context mode, thus enabling users to deal effectively with a given application scenario's particular requirements in terms of case similarity.
- The representation of all the platform's schemas (case, domain and context) is done by means of the OWL language which is a W3C standard for ontology representation, thus enabling the semantic interoperability of the platform with a large variety of external systems that contain knowledge and content in that form.
- The platform can be used as well in scenarios and domains where there is no need for dealing with imprecision and vagueness. That is because, as suggested in section 2.1, a crisp ontology is a special case of a fuzzy ontology and therefore all the fuzzy-related functionality that has been implemented in the platform can be used for crisp content as well.

Both the architecture and the methodology have been applied so far in a real life application in the context of the electronic library of the Hellenic Transmission System Operator S.A. (HTSO), a deployed knowledge portal that provides the public with intelligent access services to knowledge regarding the Greek electricity market. The details regarding this application may be found in [3].

Finally, the directions that our future work should take in order for the IKARUS platform to be a complete Fuzzy Ontology CBR solution include i) the development of an imprecise knowledge acquisition methodology that will enable knowledge engineers to develop fuzzy ontologies in an effective way and ii) the development of a detailed methodology for parameterizing and adapting the IKARUS-CBR context model to the particular requirements of various application scenarios.

References

1. Aamodt, A., Plaza, E.: Case-based Reasoning: Foundational Issues, Methodological Variations and Systems approaches. AI-Communications, vol. 7, no. 1, pp. 39-59, (1994)
2. Alexopoulos, P., Wallace, M., Kafentzis, K.: A Fuzzy Ontology Framework for Customized Assessment of Semantic Similarity. In: 3rd International Workshop on Semantic Media and Adaptation (SMAP), Prague, Czech Republic (2008)

3. Alexopoulos, P., Wallace, M., Kafentzis, K., Askounis, D.: Utilizing Imprecise Knowledge in Ontology-based CBR Systems through Fuzzy Algebra. International Journal of Fuzzy Systems, Special Issue on Fuzzy Approaches for Ontology Applications and Adaptive Web Services, Vol. 12, No. 1 (2010)
4. Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D., Patel Schneider, P., Stein, L.A.: OWL Web Ontology Language Reference. W3C Recommendation 10 (2004)
5. Diaz-Agudo, B., Gonzalez-Calero, P. A.: Knowledge Intensive CBR through Ontologies. In: Proceedings of the 6th UK CBR Workshop (2001)
6. Gomez-Perez, A., Corcho, O., Fernandez-Lopez, M.: Ontological Engineering. Springer-Verlag London Limited (2004)
7. Kahraman, C.: Fuzzy decision-making applications (Special issue). Journal of Approximate Reasoning, vol. 44, no. 2 (2007)
8. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic, Theory and Applications. Prentice Hall (1995)
9. Quanming, Z., Lingling, L., Zhigang, L., Jiannan, W., Fengguo, L.: Fuzzy CBR based on Pattern Recognition and Its Application. In: Proceedings of 2nd IEEE International Conferences on Cybernetics & Intelligent Systems and Robotics, Automation & Mechatronics, Bangkok, Thailand (2006)
10. Recio-Garcia, A. J., Diaz-Agudo, B., Gonzalez-Calero, P., Sanchez, A.: Ontology based CBR with jCOLIBRI. Applications and Innovations in Intelligent Systems XIV, Proceedings of AI-2006, the Twenty-sixth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp. 149-162, Cambridge, United Kingdom, Springer (2006)
11. Smith, B.: Ontology (Science). In: Formal Ontology in Information Systems. Proceedings of FOIS 2008, pp. 21-35, Amsterdam New York: ISO Press (2008)
12. Zadeh, L.A.: From search engines to question-answering systems - the need for new tools. In: The 12th IEEE International Conference on Fuzzy Systems 2003, Volume 2, pp. 1107 – 1109 (2003)