

Utilizing Imprecise Knowledge in Ontology-based CBR Systems by Means of Fuzzy Algebra

Panos Alexopoulos, Manolis Wallace, Konstantinos Kafentzis, and Dimitris Askounis

Abstract

Case Based Reasoning (CBR) is a problem-solving paradigm that uses knowledge of relevant past experiences (cases) to interpret or solve new problems. An evolution to this paradigm is ontology-based CBR, an approach that combines, in the form of formal ontologies, case specific knowledge with domain one in order to improve the effectiveness of the CBR process. This effectiveness is further improved if ontology-based CBR systems are able to utilize knowledge that is vague or imprecise; to that end, we present in this paper a novel CBR approach that manages and utilizes imprecise knowledge through the integration of Fuzzy Algebra in the ontology-based CBR paradigm. The approach has been applied in real life and constitutes the core of a portal that provides the public with intelligent access to knowledge assets.

Keywords: *Ontologies, Case Based Reasoning, Fuzzy Algebra.*

1. Introduction

This Case Based Reasoning (CBR) is a problem solving technique originating from Schank's concept of reminders [11]. The latter states that when people are thinking they are merely recalling past experiences that are somehow similar to their current situation. This statement applied in problem solving is translated into trying to solve new problems by comparing them to problems already solved [1, 15]. The underlying assumption is that if two problems are sufficiently similar, then their solutions are probably also similar.

This basic idea of problem comparison is formalized within the CBR paradigm through a cyclic process that involves organizing, storing, retrieving and updating

pairs of problems and solutions. In the CBR terminology these pairs are called *cases* while a collection of them is called a *case base*. Problem solving is performed by identifying the current problem situation, retrieving from the case base a similar to it past case and using the latter to suggest a solution to the current problem. The new problem-solution pair forms a new case that is added to the case base for future use.

The CBR paradigm has evolved throughout the years leading to a number of approaches and systems, each differing into the way it supports the activities of the CBR process. One may categorize these approaches in many ways, nevertheless an important categorization factor is the extent at which each approach utilizes additional knowledge that describes the cases' domain(s). Approaches that do utilize such knowledge are characterized as *knowledge intensive CBR* [1] and are generally more effective in retrieving similar cases than the ones which do not use any such knowledge.

Knowledge intensive CBR (KI-CBR) approaches typically comprise a) a knowledge representation formalism that enables the explicit modeling of case and domain knowledge and b) a reasoning method 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 recently proposed is Ontology-Based CBR [5]. 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 [12]. Technically speaking, ontologies are formal descriptions 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, ontologies 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

Corresponding Author: Panos Alexopoulos is with IMC Technologies S.A, 2 Marathonos St & 360 Kifissias St., 15233, Chalandri, Athens, Greece and with the School Of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou St., Zografou, 15773, Athens, Greece.

E-mail: palexopoulos@imc.com.gr

Manuscript received 19 Oct. 2009; revised 11 Feb. 2010; accepted 20 Feb. 2010.

KI-CBR systems is detected in 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 making 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 [17] 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 present in this paper a novel KI-CBR framework that may handle and exploit imprecise knowledge through the effective integration of Fuzzy Algebra in the ontology-based CBR paradigm. The approach we follow 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 we will show in the next sections, makes our approach more effective, complete and intuitive as far as management and exploitation of imprecision is concerned.

The integration of Fuzzy Logic in Ontology-Based CBR is performed in our framework 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 particular, our framework supports the representation of imprecise case-specific and domain-specific knowledge through a comprehensive fuzzy ontology framework while the retrieval of cases is enabled by a highly customizable fuzzy semantic similarity framework.

Given the above, the structure of the rest of the paper is as follows: In the next section we describe the key dimensions of our approach discussing related work while in section 3 we describe the key characteristics and components of our fuzzy ontology framework and we show how it may be used for the representation of cases and relevant domain knowledge. In section 4 we provide an analytical description of the framework’s case re-

trieval mechanism and in particular of the way the combination of ontologies and fuzzy logic enables the effective assessment of case semantic similarity. In sections 5 and 6 we describe a comprehensive case study which illustrates the applicability of our framework in a concrete CBR scenario and we evaluate the effectiveness of our approach, compared to other more conventional approaches, in both a qualitative and quantitative way. Finally, in section 7 we list our concluding remarks and outline potential future work.

2. Overview of our Approach and Related Work

As suggested in the previous section, our proposed KI-CBR framework operates on two axes, namely ontology-based representation of imprecise knowledge and utilization of this knowledge for effective case retrieval.

For the first axis, a number of approaches for defining fuzzy ontologies have been proposed including Fuzzy Description Logics [13, 14] and several other, non-standard and application-specific approaches [2, 3, 18]. However, while all these efforts focus on (different) formalisms that enable the representation of fuzziness, none of them deals with the issue of what this fuzziness actually means. Or, to be more precise, all approaches suggest that a fuzzy degree is a measure of imprecision but none of them actually attempts to define the nature and meaning of this imprecision in a comprehensible and application-independent manner.

The lack of such a definition restricts the usability and effectiveness of these frameworks as i) the fact that nobody knows what a fuzzy degree actually means makes the knowledge acquisition process problematic and the definition of effective fuzzy similarity measures difficult and ii) the fact that imprecision is not modelled in an application-independent manner means that the reusability of a fuzzy ontology is not possible [7].

For that, in this work, we focus on developing a fuzzy ontology framework that does not merely support the definition of fuzzy degrees but also gives them a specific meaning and role depending on where and how they are used within the ontology. In doing that, we first identify the different ways imprecision may appear within the components of an ontology as well as the different meanings this imprecision may take. Then, based on this analysis, we suggest for each component a comprehensive way for representing imprecision by means of Fuzzy Algebra elements.

For the second axis, namely case retrieval, relevant to ours work can be found in ontology-based CBR systems and Fuzzy CBR systems. More specifically, case retrieval in CBR systems is typically performed through a computational approach that involves assessing the

similarity between the requested case and those stored in the case base and retrieving the cases with the highest similarity. The assessment is performed by first determining the partial similarities of the two cases based on the values of each of their attributes and then aggregating these similarities in order to derive a single similarity value for the two cases.

In ontology-based CBR, the retrieval of cases involves the exploitation of the structure and content of the ontology for computing the semantic similarity between the attribute values and consequently the cases. For that, a number of ontology-specific similarity functions have been proposed, each utilizing ontological knowledge in a different manner [10]. None of these measures however, utilizes imprecise knowledge in any way.

On the other hand, Fuzzy CBR systems, utilize imprecise knowledge through the use of Fuzzy Logic for case representation and relevant fuzzy pattern matching techniques for similarity assessment. In our opinion, however, the fact that the representation of this knowledge is not in the form of ontologies restricts the effectiveness of these systems as the latter are not able to take advantage of the reasoning capabilities that ontologies provide, especially in regard to semantic relations.

Given the above, our approach towards case retrieval involves combining the reasoning capabilities of classical ontologies with the ones of fuzzy systems in order to create a powerful hybrid reasoning mechanism that may be effectively used for the assessment of semantic similarity between imprecisely defined cases.

3. Fuzzy Ontology Framework for Case Representation

A. Ontology Components and Imprecision

An ontology may be defined as a set of concepts, instances, properties and relations. A concept represents a set or class of entities within a domain while the entities that belong to a concept are called instances of this concept. A relation in turn links a concept instance to another instance while a property links an instance to a standard data type such as string, integer, float, boolean etc.

Imprecision in an ontology is primarily detected in relations and properties. More specifically, an ontology relation connecting two concept instances is regarded as imprecise when this connection may, in certain domains, contexts and application scenarios, be considered to contain some vague characteristic [16]. The existence and meaning of this vagueness depends on the intended meaning of the relation and given the vast range of possible such meanings, an exhaustive mapping between relation and imprecision types is a very difficult task. Nevertheless, we can identify two broad kinds of ontol-

ogy relations in which imprecision may be assigned a comprehensible meaning: hierarchical relations and associative ones.

Hierarchies organize entities in a tree-structure by defining a partial ordering of them according to some relation. Common types of hierarchical relations are taxonomies and mereologies. A taxonomical relation is the well-known is-a relation which associates an entity of a certain type to another entity of a more general type. This association can be regarded as imprecise as, while it denotes that the meaning of an entity is more specific than the one of another, the actual level of this specialization, is unclear. Therefore, imprecision in a taxonomy reflects the absence of information on how “close” the meaning of the child entity to that of the parent is.

A mereological relation on the other hand is based on the is-part-of relation which associates parts with their wholes. In general, part-of relations are not by default transitive therefore a mereological relation can be defined as a part-of relation that is considered transitive in a given context.

The nature and possible meaning of imprecision in a mereological relation may be derived from the fact that the part is always “less” than the whole. This “less”, when not determined explicitly in a qualitative or quantitative way, is the imprecise aspect of the relation. What this imprecision actually stands for can be determined by the fact that, in general, a part-of relation attempts to take into account the degree of differentiation of the parts with respect to the whole.

Finally, associative relations relate entities in a non-hierarchical way. Imprecision in such a relation reflects the lack of accurate information on the strength of the association between the entities, either due to the relation’s inherent vagueness or the absence of accurate information on some relation’s characteristic.

As far as properties are concerned, these may be regarded as imprecise in two cases:

- When the property may relate its concept’s instances to literal values in an imprecise way.
- When the literal values to which the property relates its concept’s instances may be expressed in an imprecise way.

Imprecision in the first case is similar to the one of associative ontology relations as the only difference between a property and a relation is that the latter relates instances to each other instead of instances and literal values. Consider for example the property *category* of the concept *Book* that takes as values strings denoting subject categories. The property practically links book instances to relevant specific categories. Since relevance is an inherently vague notion, the property may be regarded as imprecise.

For the second case one may consider the example of

the property *height* of the concept Person. The property normally takes as a value a number denoting the height in some metric system. Yet, uncertain application scenarios, it might be that values of height are expressed through the terms “short”, “average” and “tall”, the reason being that its exact numerical value is unknown. This kind of representation, however, is vague as it is not possible (or desirable) for any of these terms to be defined in a clear and precise way. Therefore *height* is considered to be an imprecise property.

Within our framework we consider important to enable representation of all the above types of imprecision as all of these may come up in practical application scenarios. Therefore in the following two paragraphs we define a generic fuzzy ontology representation model that supports all these types and we show how it may be used for the representation of fuzzy cases.

B. Fuzzy Ontology Representation Model

Given the analysis of the previous section, imprecise ontology components may be mapped to fuzzy algebra constructs as follows:

- An imprecise ontology relation may be seen as a fuzzy relation linking concept instances at certain degrees.
- An ontology property that relates instances to literal values in an imprecise way may be seen as a property whose literal values form a fuzzy set.
- An ontology property whose literal values can be imprecisely expressed can be represented as a property whose values are terms defined within some fuzzy linguistic variable.

Thus, 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 $flv \in FLV$.

C. Fuzzy Case Representation

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, our framework enables representation of imprecision within a case in two ways:

- By allowing case attributes to be defined as fuzzy ontology relations or fuzzy ontology properties.
- By allowing the relation attribute values to be derived from some fuzzy domain ontology.

More formally, we 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 $flv \in FLV$.

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 been developed for this paradigm are applicable within our framework as well.

4. Fuzzy Semantic Similarity Framework for Case Retrieval

As suggested in section 2, 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. Therefore, any CBR

framework should recognize the different attribute types it may contain, define corresponding similarity measures for each of them and provide meaningful operators for partial similarity aggregation.

In our framework, 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 [8]. 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 we have explained above, imprecision in ontology-based CBR may be manifested in other ways as well. For that, in this work, we focus on defining a case similarity framework that covers all of the aforementioned types of fuzziness.

Our approach towards such a definition includes the development of methods for assessing, for each fuzzy attribute type, the similarity between the values these 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 imprecision and the application's context. Therefore, for each fuzzy attribute type we first determine how imprecision affects similarity and what kind of contextual information (if any) is needed and then we proceed to suggest a proper similarity measure.

On the other hand, the establishment of methods for the aggregation of different attribute-level similarities into a single similarity value is out of the scope of this work, as any such method from the already existing CBR approaches may be applied directly in our proposed framework without need for special adaptation.

A. Value Similarity for Fuzzy Property Attributes

For the purpose of our analysis, we consider in this paragraph a generic case retrieval scenario in which the stored cases are characterized by one fuzzy property attribute and have respective fuzzy sets of literals as values while the requested case consists of a crisp set of literals regarding the same attribute. In this scenario, the similarity between a stored case and the requested one, based solely on the values of their fuzzy property attribute, is calculated by comparing their respective literal sets.

In traditional CBR, where the two sets are crisp, this comparison is generally performed in two levels, namely the literal level and the set level. The first involves comparing each pair of the two sets' literals and finding their similarity while the second involves aggregating these pair similarities into a single value denoting the

similarity of the two sets. Literal-level similarity is usually assessed in a binary way, that is it is 1 if the literals are the same and 0 if not. Nevertheless, depending on the literal type, other similarity measures may be used as well. Set-level similarity in turn may be calculated in a number of ways depending on the application scenario. For example, in one scenario it might be desirable that the requested case is fully similar to a stored one when its value set is a subset of the latter's set while in another scenario complete similarity might be desirable when the intersection of the two sets is a non-empty set. In any case the set-level similarity between the two value sets can be considered to be derived from a formula of the form

$$\text{sim}(RCVS, SCVS) = \text{Agg}(\{\text{sim}(rl_i, sl_j)\}) \quad (1)$$

where Agg is a similarity aggregation function, $RCVS$ and $SCVS$ are the requested and stored cases' (crisp) value sets respectively and $\text{sim}(rl_i, sl_j)$ is the similarity between the i_{th} literal value of $RCVS$ and the j_{th} literal value of $SCVS$. The latter is usually 1 when the two literals are equal and 0 otherwise.

In our CBR framework, where imprecision is present, the fuzziness of the set $SCVS$ is expected to play a role in the above computation. Our goal is to determine in what way should this fuzziness be incorporated in the above formula so that it covers all possible application scenarios. In doing that we consider the nature of the aggregation function Agg as well as the meaning of $SCVS$'s fuzziness as defined in section 3.

More specifically, any manifestation of the function Agg regards the similarity $\text{sim}(rl_i, sl_j)$ as a partial similarity between the sets $RCVS$ and $SCVS$ and consequently between the stored and the requested case. When the set $SCVS$ is fuzzy the strength of the association between any of its literals and the stored case is not fixed but it is determined by the corresponding fuzzy degrees. This means that in order for $\text{sim}(rl_i, sl_j)$ to be considered as a partial similarity between the two cases it needs to be "adjusted" according to the fuzzy membership degree of the literal sl_j . Such an adjustment could take the following form:

$$\text{sim}(RCVS, SCVS) = \text{Agg}(\{\text{sim}(rl_i, sl_j) * SCVS(sl_j)\}) \quad (2)$$

where $SCVS(sl_j)$ is the fuzzy membership degree of the j_{th} literal value of $SCVS$.

Nevertheless, the actual decision of whether this adjustment should be made depends on the given application scenario and particularly on whether the imprecision of the literal-case association is considered to play a role in similarity in that scenario. Therefore, in order to deal with the subjectiveness of the imprecision's role in simi-

larity assessment we regard it as contextual information that needs to be captured and utilized within the similarity calculation formula.

More specifically, given a set of fuzzy property attributes FPA we define Fuzzy Property Attribute Similarity Context as a function $fpasc : FPA \rightarrow \{0,1\}$. If $fpa \in FPA$ then $fpasc(fpa)$ determines whether the imprecision of the attribute should be considered when comparing the value sets of a requested and a stored case for this attribute. A value of 0 denotes no consideration while a value of 1 the opposite. In the case of no consideration the set $SCVS$ is considered crisp. Given that, the similarity assessment formula between a stored case SC and a requested one RC based solely on the values of the attribute fpa becomes:

$$sim(RCVS, SCVS) = Agg(\{sim(rl_i, sl_j) * \max(SCVS(sl_j), 1 - fpasc(fpa))\}) \quad (3)$$

B. Value Similarity for Fuzzy Valued Property Attributes

Similarly to above, we consider in this paragraph a generic case retrieval scenario in which the stored cases are characterized by one fuzzy valued property attribute and have respective single linguistic terms as values while the requested case consists of a single linguistic term regarding the same attribute. In such a scenario, the similarity between a stored case and the requested one, based solely on the values of their fuzzy valued property attribute, is calculated by comparing their respective linguistic terms.

As suggested above, this type of comparison has been extensively examined in the literature; for example Chen and Hwang's *crisp score method* for defuzzifying fuzzy sets [4]. This and similar approaches could be used within our framework as well.

C. Value Similarity for Fuzzy Relation Attributes

For the last type of case attributes we consider again a retrieval scenario in which the stored cases are characterized by one fuzzy relation attribute and have respective ontology instances as values while the requested case consists of a crisp set of ontology instances regarding the same attribute. The instances are considered to be derived from some fuzzy domain ontology. In such a scenario, the similarity between a stored case and the requested one, based solely on the values of their fuzzy relation attribute, is calculated by comparing their respective ontology instance sets.

The only difference between this comparison and the one performed in the case of a fuzzy property attribute is that the compared set elements are ontology instances rather than literals. That is because for a single stored case the fuzzy related instances may be well regarded as

a fuzzy set. This means that the similarity assessment formula of paragraph 4.A may be applied here as well, reducing thus the problem into assessing the similarity between pairs of ontology instances.

In traditional ontology-based CBR, this is generally performed by utilizing the components of the domain ontology and particularly the relations that connect the instances. That is because most of these connections may, in certain contexts, indicate some kind of similarity between the instances. In our approach, where imprecision plays a central role, we claim that 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.

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. That is because in different application scenarios and among different users, the contribution of the same fuzzy relation to the similarity between two instances might be totally different. Therefore, relevant contextual information needs to be modeled.

With the above in mind, we define as parts of our semantic similarity framework the following components:

- The Fuzzy Ontology Relation Similarity Context (FORSC), a context model that enables the effective modeling of information regarding the expected role of the fuzzy ontology's relations in the instance similarity assessment process.
- An algorithm for the assessment of the similarity between any two fuzzy ontology instances based on the information contained in the ontology and the respective similarity context.

In particular, within an ontology an instance may be connected to another instance through a number of relations or compositions of them. The aim of the Fuzzy Ontology Relation Similarity Context is to define whether and to what extent each of these connections should be interpreted as similarity ones.

More formally, given a fuzzy ontology $O_F = \{C, I, FHR, FAR, FP, FLV, FVP, FORSC(O_F)\}$, defines:

- How each hierarchical relation $R \in FHR$ should be used for computing similarity between two concept instances.
- How the reverse relation of each hierarchical relation $R \in FHR$ should be used for computing similarity between two concept instances.
- How each associative relation $R \in FAR$ should be used for computing similarity between two concept instances.
- How the reverse relation of each associative relation $R \in FAR$ should be used for computing similarity

between two concept instances.

- How the composition of n relations $R_1, R_2, \dots, R_n \in FAR \cup FHR \cup FAR^{-1} \cup FHR^{-1}$ should be used for computing similarity between two concept instances. The sets FAR^{-1} and FHR^{-1} contain the reverse relations of the sets FAR and FHR respectively.

To do that, FORSC comprises five distinct subcontexts that correspond to each of the above cases:

- The Fuzzy Hierarchical Relation Similarity Context which is defined as a function $fhrsc : FHR \rightarrow [-1,1]$.
- The Fuzzy Reverse Hierarchical Relation Similarity Context which is defined as a function $frhrsc : FHR \rightarrow [-1,1]$.
- The Fuzzy Associative Relation Similarity Context which is defined as a function $farsc : FAR \rightarrow [-1,1]$.
- The Fuzzy Reverse Associative Relation Similarity Context which is defined as a function $frarsc : FAR \rightarrow [-1,1]$.
- The Fuzzy Composite Relation Similarity Context which is defined as a function $fcrsc : (FHR \cup FAR \cup FHR^{-1} \cup FAR^{-1})^n \rightarrow [-1,1]$ where n is the number of composed relations.

The exact meaning of each context is the following:

- If $R \in FHR$ then $fhrsc(R)$ is the degree at which the relation $Tr^t(R)$ should be considered to denote similarity between all entities $a, b \in I$ for which $[Tr^t(R)](a, b) \neq 0$. $Tr^t(R)$ is the sup-t transitive closure of the relation R .
- If $R \in FHR$ then $frhrsc(R)$ is the degree at which the relation $[Tr^t(R)]^{-1}$ should be considered to denote similarity between all entities $a, b \in I$ for which $[Tr^t(R)]^{-1}(a, b) \neq 0$.
- If $R \in FAR$ then $farsc(R)$ is the degree at which the relation R should be considered to denote similarity between all entities $a, b \in I$ for which $R(a, b) \neq 0$.
- If $R \in FAR$ then $frarsc(R)$ is the degree at which the relation R^{-1} should be considered to denote similarity between all entities $a, b \in I$ for which $R^{-1}(a, b) \neq 0$.
- If $R_1, R_2, \dots, R_n \in FAR \cup FHR \cup FAR^{-1} \cup FHR^{-1}$ then $fcrsc(R_1, R_2, \dots, R_n)$ is the degree at which the relation $[R_1^* \circ^t R_2^*, \circ^t \dots, \circ^t R_n^*]$, $R_i^* = R_i$ if $R_i \in FAR \cup FAR^{-1}$ and $R_i^* = Tr^t(R_i)$ if $R_i \in FHR \cup FHR^{-1}$, should be considered to denote similarity between all entities $a, b \in I$ for which $[R_1^* \circ^t R_2^*, \circ^t \dots, \circ^t R_n^*](a, b) \neq 0$.

The values of all the above degrees might range from -1 to 1. A degree of -1 denotes that the relation or the

pair of relations should not be considered at all in measuring similarity. A degree of 1 denotes the exact opposite, namely two concepts connected with this relation should be considered identical. Any degree between -1 and 1 denotes an intermediate situation.

The utilization of FORSC for the context-based assessment of instance semantic similarity is performed through the following process: Given a fuzzy ontology O_F and a corresponding similarity context $FORSC(O_F) = \{fhrsc, frhrsc, farsc, frarsc, fcrsc\}$ we define the similarity context operator sco as follows:

$$sco(R(a, b), f) = \begin{cases} R(a, b)^{1-f(R)}, & 0 \leq f(R) \leq 1 \\ R(a, b) \times (1 + f(R)), & -1 \leq f(R) < 0 \end{cases} \quad (4)$$

where $R \in FAR \cup FHR \cup FAR^{-1} \cup FHR^{-1}$ and $f \in FORSC(O_F)$. This operator is used for applying to each ontology relation its corresponding similarity context and changing its degrees. The basic idea is that after a fuzzy ontology relation has been ‘‘contextualized’’ through its corresponding similarity context and the operator, then the relation’s ‘‘adjusted’’ degrees are considered as similarity ones.

More analytically, the contextualization of the fuzzy domain ontology is performed through the following steps:

- $\forall R \in FHR$ such that $fhrsc(R) \neq -1$ we compute the fuzzy relation $R'_h = sco(Tr^t(R), fhrsc)$ (that is we apply the hierarchical relation similarity context to all the hierarchical relations that, according to the context, denote some kind of similarity).
- $\forall R \in FHR$ such that $frhrsc(R) \neq -1$ we compute the fuzzy relation $R''_h = sco(Tr^t(R), frhrsc)$ (that is we apply the reverse hierarchical relation similarity context to all the hierarchical relations whose reverse relation, according to the context, denotes some kind of similarity).
- $\forall R \in FAR$ such that $farsc(R) \neq -1$ we compute the fuzzy relations $R'_a = sco(R, farsc)$ (that is we apply the associative relation similarity context to all the associative relations that, according to the context, denote some kind of similarity).
- $\forall R \in FAR$ such that $frarsc(R) \neq -1$ we compute the fuzzy relations $R''_a = sco(R, frarsc)$ (that is we apply the reverse associative relation similarity context to all the associative relations whose reverse relation, according to the context, denotes some kind of similarity).
- $\forall R_1, R_2, \dots, R_n \in FAR \cup FHR \cup FAR^{-1} \cup FHR^{-1}$ such that $fcrsc(R_1, R_2, \dots, R_n) \neq -1$ we compute the fuzzy relation $R_{comp} = sco([R_1^* \circ^t R_2^*, \circ^t \dots, \circ^t R_n^*], fcrsc)$ where $R_i^* = R_i$ if $R_i \in FAR \cup FAR^{-1}$ and $R_i^* = Tr^t(R_i)$ if

$R_i \in FHR \cup FHR^{-1}$ (that is we apply the composite relation similarity context to all the relations whose composition, according to the context, denotes some kind of similarity).

In the end of the above procedure we compute the fuzzy relation $R_c = R_h' \cup R_h'' \cup R_a' \cup R_a'' \cup R_{comp}$, namely the fuzzy union of all the resulting relations of the algorithm's steps. Using R_c we can determine the semantic similarity between two instances $a, b \in I$ simply by getting the fuzzy degree of $R_c(a, b)$.

5. Case Study: The Electronic Library of HTSO

The approach presented in this paper has been the core upon which the development of the electronic library of the Hellenic Transmission System Operator S.A. (HTSO) has been based. This is a real life, fully deployed knowledge portal that provides the public with intelligent access services to knowledge regarding the Greek electricity market.

The knowledge has the form of legal and technical documents which, prior to the development of the library, were accessible merely by means of full-text search. In the context of the library these documents are modelled as cases and are handled and processed according to our proposed framework by means of a corresponding Fuzzy Semantic CBR engine. The engine manages to capture and exploit not only the documents' semantic content but also the imprecision that characterizes it, enabling thus the provision of fundamentally more effective semantic search services to the portal's users.

In the next paragraphs we provide the details of how this was made possible by describing i) the engine's architecture, ii) the way the engine's case base was specified and populated according to our fuzzy ontology representation framework and iii) the way the engine's case retrieval mechanism was adapted and utilized in the context of the specific application scenario.

A. CBR Engine Architecture

The architecture of the library's CBR Engine, depicted in figure 1, comprises the following components:

- **Case Repository:** An ontology repository which stores the library's cases and their retrieval context in the form of two OWL ontologies. The case ontology models section's 3 representation framework while the context ontology models the contexts defined in paragraphs 4.A and 4.C.
- **CBR Storage Manager:** A system that implements all the necessary functionality for accessing and managing the case repository's contents.

- **CBR Retriever:** A system that implements all the necessary functionality for calculating the similarity between cases. This involves implementation of the processes defined in paragraphs 4.A, 4.B and 4.C.

B. Case Specification and Population

The process of specifying and populating the system's case practically involved the use of the fuzzy ontology representation model provided in section 3 for defining the structure and content of the system's fuzzy case ontology. This definition was performed with the help of several experts in the domain of the Electricity Market and its result was the **HTSO Fuzzy Case Ontology**.

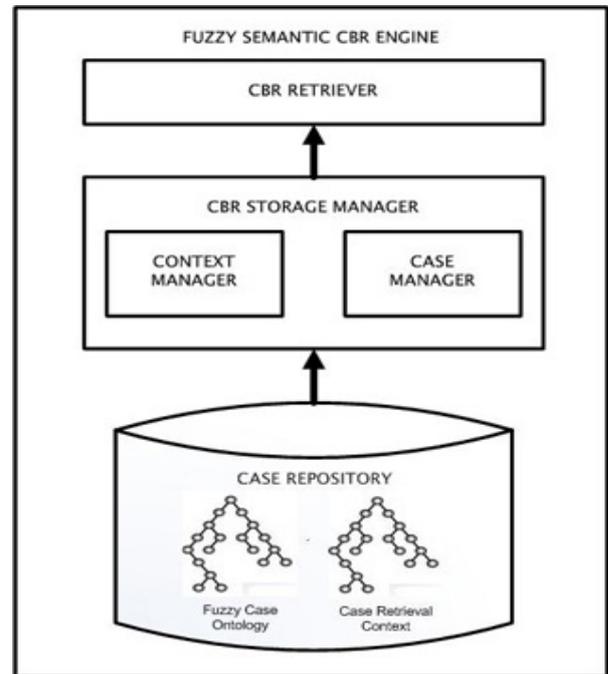


Figure 1. Fuzzy Semantic CBR Engine Architecture.

The ontology consists of a single fuzzy case type, called **HTSODocument**, which in turn comprises the following case attributes:

- **Title:** The title of the document.
- **Date:** The document's publication date.
- **Contributors:** The people or organizations that have contributed to the writing of the document.
- **ThematicContent:** Keywords that characterize the document's semantic content.
- **RelevantExternalResources:** Links to relevant to the document resources that are not contained in the library

The first two attributes are practically non-fuzzy therefore they are represented as crisp property attributes (i.e. fuzzy property attributes with all their value degrees equal to 1). On the other hand, the attributes **Contributors** and **RelevantExternalResources** are considered to

be fuzzy and therefore are represented as fuzzy property attributes. Fuzziness in the values of the Contributors attribute reflects how much a certain individual or organization has contributed to the document's creation while fuzziness in the values of **RelevantExternalResources** attribute reflects a given source's degree of relevance to the document.

Table 1. Electricity Market Ontology Concepts.

Concept	Description
Market Process	Refers to any process or procedure that is part of the overall operation of the electricity market.
<i>Market Participant</i>	Refers to individuals and organizations that play some role within the market.
<i>Market Action</i>	Refers to actions performed by market participants in the context of specific market processes
<i>Market Right</i>	Refers to rights of the market participants in the context of the market operation.
<i>Market Obligation</i>	Refers to obligations of the market participants in the context of the market operation.
Market Rule	Refers to rules that govern the operation of the market.
Market System	Refers to equipment, facilities and systems that are used in the context of the market operation.
<i>System Specification</i>	Refers to standards and specifications that characterize the market's systems. moment
<i>Market Information Source</i>	Refers to any piece of market related information such as manuals, document templates, archives, registries etc.
Market Extent	Refers to parameters, data, factors, elements and virtually any extent that is relevant with the electricity market.

Finally, the **ThematicContent** attribute is modelled as a fuzzy relation attribute relating cases to keywords that represent their meaning. These keywords have the form of ontology instances belonging to the concepts of a fuzzy ontology that covers the domain of the Electricity Market. These concepts, shown in table 1, are interrelated through 18 relations, both crisp and fuzzy. A sample of these relations is shown in table 2.

The population of the case repository was performed by creating instances of the concept HTSODocument and by assigning specific values to its attributes. Again this was done with the help of the domain experts.

C. Case Retrieval Context Specification

The specification of the case retrieval context involved defining for each fuzzy case attribute its corresponding type-specific context. Given the model of the previous section, this meant that we needed to define:

- A fuzzy property attribute similarity context for the **Contributors** attribute.
- A fuzzy property attribute similarity context for the **RelevantExternalResources** attribute.

- A fuzzy ontology relation similarity context for those relations of the electricity market ontology that relate instances that are potential values of the **ThematicContent** attribute

Table 2. Electricity Market Ontology Fuzzy Relations.

Relation	Description	Meaning of Fuzziness ^a
<i>isPartOfProcess</i>	Hierarchical relation that links processes with the other processes that are part of them	Fuzzy Parthood
<i>isPartOfAction</i>	Hierarchical relation that links actions with the sub-actions that might comprise them	Fuzzy parthood
<i>isKindOfParticipant</i>	Hierarchical relation that links types of participants	Fuzzy is-a
<i>isKindOfRight</i>	Hierarchical relation that links market rights	Fuzzy is-a
<i>isKindOfInformationSource</i>	Hierarchical relation that links types of market information sources	Fuzzy is-a
<i>isPartOfSpecification</i>	Hierarchical relation that links system specifications with the sub-specifications that might comprise them	Fuzzy parthood
participatesIn	Associative relation that links market participants with the processes they participate in	Extent of participation within the process
<i>isPerformedBy</i>	Associative relation that links market actions with the participants that they perform them	Extent of participation within the action
<i>isPerformedIn-TheContextOf</i>	Associative relation that links market actions with the processes they are part of.	Importance of action within the process
regardsProcess	Associative relation that links rights, obligation and rules with the processes they refer to.	Crisp Relation
concernsParticipant	Associative relation that links rights and obligations with market participants.	Crisp Relation
isUtilizedIn	Associative relation that links information sources with the processes they play some role in.	Crisp Relation
isFoundIn	Associative relation that links extents with information sources they are referred within.	Crisp Relation _r
<i>hasSpecification</i>	Associative relation that links units and systems with their respective specifications.	Crisp Relation

For the two first attributes the respective context values were chosen to be $fpasc$ (Contributors)=1 (i.e. the fuzziness of the attribute should be considered in similarity assessment) and $fpasc$ (Relevant External Resources)=0 (i.e. fuzziness is not important). For the the-

matic content attribute the context values that are not equal to -1 are shown in table 3. Note that a context value of 0 was assigned to the attribute ThematicContent (which is in principle an associative relation), practically denoting that the attribute's fuzziness should be considered when assessing case similarity (similar role as fuzzy property similarity context defined above).

D. Case Retrieval

The HTSO fuzzy case ontology and its corresponding retrieval context defined in the previous paragraphs enabled the CBR retriever system to assess the similarity between the cases' attribute values by implementing the methods described in section 4. This assessment of similarity is more effective than other CBR approaches as it considers both the imprecision and subjectiveness that characterize the domain and application scenario. To illustrate this we consider a number of attribute similarity assessment examples and contrast their results to those that would be derived from other CBR approaches.

In particular, for fuzzy property attributes the similarity assessment formula of paragraph 4.A applies. This means that in a specific example in which the parameters of this formula are as follows

- Attribute: Contributors
- Requested Case Value Set (RCVS): {Christoforos Zoumas}
- Stored Case Value Set (SCVS): {Christoforos Zoumas /0.6, Konstantinos Petsinis /0.1, Vasilis Ziogas /0.9}.

- Agg: $\frac{1}{|m|} \cdot \frac{1}{|n|} \sum_1^m \sum_j^n$

the similarity between the two value sets is 0.6 (given that $f_{pasc}(\text{Contributors})$ is 1). In a traditional CBR system where the set SCVS would be crisp the same similarity would be 1.0. The difference lies in the fact that in the first as the fuzzy participation degree of contributor Christoforos Zoumas is taken in mind adjusting accordingly the two sets' similarity.

On the other hand, if we consider a similar example for the attribute `relevantExternalResources` the two similarity values would be the same. That happens because $f_{pasc}(\text{RelevantExternalResources})=0$ meaning that in the specific scenario the attribute's fuzziness is not supposed to influence similarity.

The two examples highlight the unique characteristic of our framework to allow for selective utilization of fuzziness in similarity assessment. This is a very important feature that no other CBR framework has and which makes the framework easily adaptable to multiple application scenarios in each of which the same kind of fuzziness might play a different role as far as similarity is concerned.

Table 3. Fuzzy Ontology Relation Similarity Context for HTSO Library.

Parameter	Value
$f_{pasc}(\text{ThematicContent})$	0
$f_{hrsc}(\text{isPartOfProcess})$	1
$f_{hrsc}(\text{isPartOfAction})$	1
$f_{hrsc}(\text{isKindOfParticipant})$	1
$f_{hrsc}(\text{isKindOfRight})$	1
$f_{hrsc}(\text{isKindOfInformationSource})$	1
$f_{hrsc}(\text{isPartOfSpecification})$	1
$f_{pasc}(\text{participatesInProcess})$	0
$f_{pasc}(\text{performsAction})$	-0.2
$f_{pasc}(\text{hasRight})$	0
$f_{pasc}(\text{hasObligation})$	0
$f_{pasc}(\text{isPerformedInTheContextOf})$	-0.5
$f_{pasc}(\text{isPerformedInTheContextOf})$	-0.95
$f_{pasc}(\text{regardsProcess})$	-0.5
$f_{pasc}(\text{regardsProcess})$	-0.7
$f_{pasc}(\text{foundInInformationSource})$	-0.2
$f_{pasc}(\text{foundInInformationSource})$	-0.8
$f_{pasc}(\text{participatesInProcess, isPartOfProcess})$	0

For fuzzy relation attributes, the CBR Retriever applies again the similarity assessment formula of paragraph 4.A but this time it utilizes the method described in paragraph 4.C for comparing the sets' elements (as these are ontology instances). More specifically, during the initialization of the system, the Fuzzy Ontology Relation Similarity Context of table 3 was utilized for "contextualizing" the relations of the fuzzy electricity market ontology through the similarity context operator and the 5-step process described in paragraph 4.C. Given the data of table 3 this meant that the operator was applied:

- To the reverse relations of the hierarchical relations isPartOf Process , isPartOfAction , $\text{isKindOfParticipant}$, isKindOfRight , $\text{isKindOfInformationSource}$ and $\text{isPartOfSpecification}$ (step 2 of the contextualization process).
- To the associative relations $\text{participatesInProcess}$, performsAction , hasRight , hasObligation , $\text{isPerformedInTheContextOf}$, regardsProcess and $\text{foundInInformationSource}$ (step 3 of the contextualization process).
- To the reverse relations of the associative relations regardsProcess , $\text{isPerformedInTheContextOf}$ and $\text{foundInInformationSource}$ (step 4 of the contextualization process).
- To the composition of the relations $\text{participatesInProcess}$ and isPartOfProcess (step 5 of the contextualization process).

Subsequently, the resulting fuzzy relation R_C of the above process was stored within the system and was used for computing the similarity between ontology instances and consequently between fuzzy relation attributes.

To illustrate how this works we consider a snapshot of

the electricity market ontology, shown in table 4, depicting instances of the hierarchical relation *isPartOfProcess* and the associative relation *participatesInProcess*. For this snapshot and given the relation similarity context values of table 3, table 5 depicts the computed similarities for a number of different value comparison examples regarding the attribute ThematicContent. For the computation we used the same \$Agg\$ function as above as well as the product t-norm for the algorithm of paragraph 4.3.

Table 4. Electricity Market Ontology Relation Instances.

FuzzyRelation	Degree
<i>isPartOfProcess</i> (IssueOfDispatchInstructions, DispatchProcedure)	0.5
<i>isPartOfProcess</i> (SendingOfDispatchInstructions, DispatchProcedure)	0.5
<i>participatesInProcess</i> (DispatchableUnitProducer, IssueOfDispatchInstructions)	0.9

Table 5. Similarity Values for Thematic Content Attribute.

Requested Case Value Set (RCVS)	Stored Case Value Set (SCVS)	Similarity ^a
DispatchProcedure	IssueOfDispatchInstructions/1.0	1.0
IssueOfDispatchInstructions	DispatchProcedure/1.0	0.0
Dispatchable UnitProducer	DispatchProcedure/1.0	0.45

The first row of table 5 suggests that the similarity between a requested case that needs to have as thematic content the market process **DispatchProcedure** and a stored case that is characterized by the process **IssueOfDispatchInstructions** is 1.0. This is why the reverse similarity context of the *isPartOfProcess* relation is 1.0 indicating that for all requested cases characterized by a market process the system should return as identical all the stored cases that are characterized by a corresponding sub-process. On the other hand, as indicated by the table's second row, the opposite does not apply, namely stored cases characterized by hierarchically higher processes should not be returned as relevant results.

Finally, the third row illustrates the usage of a composite relation for similarity assessment; the requested case's thematic content refers to some market participant and the system returns as (partially) relevant all cases characterized by processes (and super-processes of them) in which this participant plays some role. This kind of behaviour is determined by the relevant context of table 3.

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert

it into the text after your paper is styled.

Through this examples, we can clearly see that the similarities derived by our approach are intuitively more accurate than those that would be derived from other ontology-based CBR approaches. The main reason for that is that similarity calculation in these approaches is based merely on the ontology's hierarchies and the relative location of concepts within them while our framework does additionally the following:

1. It includes in the calculation the ontology's associative relations as well as any possible composition of these. In that way it exploits a much bigger part of the domain knowledge.
2. It includes in the calculation the relation's fuzzy degrees wherever these are considered to influence similarity. For hierarchical relations this means that the semantic distances between the various elements of the hierarchy can be calculated in a more accurate way while for associative relations it means that the latter's relative strength is taken in mind.

5. Method Evaluation & Comparative Study

Whilst the qualitative evaluation has been presented in the previous section, this section describes the quantitative and comparative performance analysis of our proposed approach in a real life setting. Essentially, the system presented in section 5 is used as the base of our experimental setting and using it our approach's responses to specific user queries are examined against "ground truth" in order to evaluate retrieval performance. In what follows we outline the methodology followed to construct the ground truth, carry out the experiments and analyze the results.

A. Experimental Settings

Given the system presented in section 5, the first step was to develop the ground truth against which all retrieval results had to be compared in order to measure retrieval performance. The ground truth in general included a set of predetermined test queries and the corresponding sets of "ideal" system responses for each query. There are two distinct and equally important steps involved in this process:

1. Since we are using a real life system, its case base is quite extended. In order for the experimental results to be representative, the test queries were selected in a manner that allowed them to span the majority of topics and content represented in the case base. At the same time, they were selected taking into consideration that each one of them should be answered by a sufficient number of cases in the base so that the evaluation is meaningful.
2. Once the set of test queries was specified, the "ideal" set of responses (list of cases to be selected) had to be

specified with corresponding degrees of matching, or equivalently ranked.

Subsequently, the test queries were fed into the system and corresponding responses were recorded; these were automatically tested against ground truth in order to make comparisons. The latter were performed according to the following evaluation criteria and measures:

- *Precision p*: Precision is defined as the number of retrieved relevant items over the number of total retrieved items..
- *Recall r*: Recall is defined as the number of retrieved relevant items over the total number of relevant items.
- *Effectiveness e*: Effectiveness is a measure designed to serve as a single number indicator of system performance. It is given by the following formula.

$$e = \frac{1}{(1/2p) + (1/2r)}$$

In order to evaluate how our proposed approach performs with respect to the current state of the art, we performed an additional set of experiments with a modified version of the system presented in section 5. Specifically, we have removed all degrees from the utilized ontology and we have not considered the contextual information of section 5.3 nor the algorithm of section 4.3. Instead we used conventional ontology concept similarity functions, taken from [10], thus reducing the system to the current state of the art in ontology driven CBR.

B. Experimental Results

Using the procedure described above 25 distinct text queries were specified as part of the ground truth, together with their corresponding ideal system responses.

Table 6. Test Query 1: Dispatch Procedure (10 relevant documents in the ground truth).

Approach	Retrieved	Relevant ^a	p	r	e
Conventional;	14	8	57%	80%	67%
Proposed	12	9	75%	90%	82%

Table 7. Test Query 2: Frequency Demand Disconnection (15 relevant documents in the ground truth).

Approach	Retrieved	Relevant ^a	p	r	e
Conventional;	13	7	54%	47%	50%
Proposed	18	13	72%	87%	79%

Table 8. Test Query 3: Cold Reserve Unit Producer (35 relevant documents in the ground truth).

Approach	Retrieved	Relevant ^a	p	r	e
Conventional;	12	10	83%	29%	43%
Proposed	38	28	74%	80%	77%

Table 9. Overall results.

Approach	\bar{p}	\bar{r}	\bar{e}
Conventional;	83%	29%	43%
Proposed	74%	80%	77%

All test queries in the ground truth were issued to both versions of the system and the effectiveness was measured. Selected individual results are shown in tables 6, 7 and 8 while overall results are shown in table 9.

As one may easily see, in all individually presented queries the proposed approach outperforms the one that does not consider the imprecision of the domain knowledge nor the application scenario's context. Intuitively, this happens in each of the three queries for the following reasons:

- In query 1 the user wishes to retrieve cases whose thematic content regards the electricity market process **Dispatch Procedure**. For that our approach utilizes the *isPartOfProcess* fuzzy relation along with its corresponding similarity context of table ?. The conventional approach on the other hand treats the relation as a crisp one and calculates similarity based on the semantic distance between the instances in the processes taxonomy. Therefore, the conventional approach retrieves more results than our approach but less of them are actually relevant.
- In query 2 the user wishes to retrieve cases whose thematic content regards the electricity market action **Frequency Demand Disconnection**. For that our approach utilizes the *isPartOfAction* and *isPerformedInTheContextOf* fuzzy relations along with their corresponding similarity contexts of table ?. The conventional approach on the other hand treats the first relation as a crisp one and ignores completely the second relation. As a result, the conventional approach retrieves less results with even fewer of them being actually relevant to the query.
- In query 3 the user wishes to retrieve cases whose thematic content regards the electricity market participant **Cold Reserve Unit Producer**. For that our approach utilizes a lot of relations (*isKindOfParticipant*, *participatesInProcess*, *performsAction*, *hasRight*, *hasObligation*) as well as the composition of the relations *participatesInProcess* and *isPartOfProcess*. The conventional approach on the other hand utilizes only the crisp version of the relation *isKindOfParticipant*. As a result it retrieves significantly less results than our approach.

As far as the overall results are concerned, it is obvious that the proposed approach greatly outperforms the conventional one in the specific experiment. Moreover, as the reasons for this enhancement in performance have been well documented in the above, it is justified to ex-

pect that the proposed approach will have similar gains with respect to the current state of the art in any other application domain as well.

6. Conclusions & Future Work

In this paper, we proposed a novel knowledge intensive CBR framework that addresses the problem of knowledge imprecision and provides the necessary tools and methodologies for handling and exploiting the latter in a comprehensive and effective way. The approach we followed involved 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 fuzzy ontology framework differs from other similar approaches in that it focuses on assigning a meaning to the fuzziness of the ontology's components. This is a very important characteristic as it makes the fuzzy ontology's imprecision explicit facilitating thus more effective knowledge acquisition and ontology reuse. Furthermore, it enables the definition of more intuitive and effective semantic similarity measures which the CBR paradigm requires for facilitating case retrieval.

Such measures are defined within our proposed fuzzy semantic similarity framework. The latter has the unique characteristic that it combines the reasoning capabilities of ontologies with the ones of fuzzy systems in order to assess the similarity of imprecisely defined cases. To our best of knowledge no other CBR approach does that as Fuzzy CBR systems do not use ontologies and Ontology-Based CBR systems do not use Fuzzy Algebra.

Another important feature of the similarity framework is that it treats similarity assessment as a highly subjective and application dependent task that needs to be performed with the help of some kind of contextual information. For that it defines a comprehensive context model that may be used for adapting the similarity assessment to the requirements of the application scenario in hand.

The applicability and effectiveness of our framework is illustrated by its utilization in a real life application in the context of the electronic library of HTSO, a deployed knowledge portal that provides the public with intelligent access services to knowledge regarding the Greek electricity market. We have provided details regarding the implementation of the framework in the form of a fuzzy semantic CBR engine as well as the way it was prepared and utilized for case representation and retrieval in the context of the specific application scenario.

These details exemplified the generic and adaptable character of our framework and at the same time the directions that our future work should take in order for our framework to be a complete Fuzzy Ontology CBR solution. These directions include the development of an imprecise knowledge acquisition methodology that will enable knowledge engineers to develop fuzzy ontologies in an effective way and the development of an analytical methodology for parameterizing and using the semantic similarity framework in various application scenarios.

References

- [1] A. Aamodt and E. Plaza, "Case-based Reasoning: Foundational Issues, methodological variations and systems approaches," *AI-Communications*, vol. 7, no. 1, pp. 39-59, 1994.
- [2] G. Acampora, C. S. Lee, and M. H. Wang, "FML-Based Ontological Agent for Healthcare Application with Diabetes," *Web Intelligence/IAT Workshops*, pp. 413-416, 2009.
- [3] S. Calegari and E. Sanchez, "A Fuzzy Ontology-Approach to improve Semantic Information Retrieval," *Proceedings of the Third ISWC Workshop on Uncertainty Reasoning for the Semantic Web - URSW'07*, vol. 327, 2007.
- [4] S. J. Chen and C. L. Hwang, *Fuzzy Multiple Attribute Decision Making, Methods and Applications*, Springer Verlag, 1992.
- [5] B. Diaz-Agudo, and P. A. Gonzalez-Calero, "Knowledge Intensive CBR through Ontologies," In B. Lees, editor, *Procs of the 6th UK CBR Workshop*, 2001.
- [6] A. Gomez-Perez, O. Corcho, and M. Fernandez-Lopez, *Ontological Engineering*, Springer Verlag London Limited, 2004.
- [7] K. Kafentzis, M. Wallace, P. Georgolios, P. Alexopoulos, and G. Mentzas, "Semantics Enabled Problem Based Brokering of Organizational Knowledge," In I. Maglogiannis, K. Karpouzis, J. Soldatos, M. Wallace M. (Eds.), *Emerging Artificial Intelligence Applications in Computer Engineering*, IOS Press, 2007.
- [8] C. Kahraman, "Fuzzy decision-making applications (Special issue)," *Journal of Approximate Reasoning*, vol. 44, no. 2, 2007.
- [9] Z. Quanming, L. Lingling, L. Zhigang, W. Jiannan, and L. Fengguo, "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, June 7-9, 2006.
- [10] A. J. Recio-Garcia, B. Diaz-Agudo, P. Gon-

zalez-Calero, and A. Sanchez, "Ontology based CBR with jCOLIBRI," *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pp. 149-162, Cambridge, United Kingdom, Springer, 2006.

- [11] R. C. Schank, *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*, Cambridge University Press, 1982.
- [12] B. Smith, "Ontology (Science)," In C. Eschenbach and M. Gruninger (eds.), *Formal Ontology in Information Systems. Proceedings of FOIS 2008*, pp. 21-35, Amsterdam/New York: ISO Press, 2008.
- [13] G. Stoilos, G. Stamou, V. Tzouvaras, J.Z. Pan, and I. Horrocks, "Fuzzy OWL: Uncertainty and the Semantic Web," *International Workshop of OWL: Experiences and Directions*, Galway, 2005.
- [14] U. Straccia, "Towards a Fuzzy Description Logic for the Semantic Web (Preliminary Report)," In *Proceedings of the 2nd European Semantic Web Conference (ESWC-05)*, 2005.
- [15] D. Walker, *Similarity Determination and Case Retrieval in an Intelligent Decision Support System for Diabetes Management*, Master of Science Thesis, Ohio University, 2007.
- [16] M. Wallace, "Ontologies and Soft Computing in Flexible Querying," *Control and Cybernetics*, vol. 2, 2009.
- [17] L. A. Zadeh, "From search engines to question-answering systems the need for new tools," *Fuzzy Systems, The 12th IEEE International Conference*, vol. 2, pp. 1107-1109, 2003.
- [18] J. Zhai J. Y. Liang, J. Jiang, and Y. Yu, "Fuzzy Ontology Models Based on Fuzzy Linguistic Variable for Knowledge Management and Information Retrieval," *Intelligent Information Processing*, vol. IV, pp. 58-67, 2008.



Panos Alexopoulos received his degree in electrical and computer engineering from the National Technical University of Athens in Greece in 2004 and the MSc in advanced computing from Imperial College London in 2005. Currently he is pursuing a PhD degree in knowledge engineering in National Technical University of Athens while working as a

knowledge engineer and researcher for IMC Technologies (Athens, Greece). His research interests include ontology engineering, semantic technology, fuzzy systems and knowledge management.



Manolis Wallace acquired his degree in electrical and computer engineering from the National Technical University of Athens in Greece in 2001 and PhD in intelligent knowledge-based systems in uncertain environments from the same university in 2005. Dr. Wallace, formerly assistant professor and chair of the department of computer science at the University of Indianapolis, Athens Campus, currently holds a visiting lecturer position at the University of Peloponnese, while at the same time serving as a project coordinator for the Foundation of the Hellenic World. He is a recipient of a 2002-2005 HERACLETUS national fund for basic research. His research interests evolve around semantics, data mining, fuzzy systems, uncertainty and practical applications of the above. Dr. Wallace is an associate editor of the *International Journal of Fuzzy Systems* and the co-founder and co-chair of the Semantic Media Adaptation (SMAP) Initiative.



Konstantinos Kafentzis acquired his degree in electrical and computer engineering from the National Technical University of Athens in Greece in 2001. Currently he is Managing Director of IMC Technologies S.A. (Athens, Greece) while in the past he has worked as a researcher of ICCS-NTUA and as an R&D Director in IMC Technologies, participating and leading several research projects in the areas of knowledge management, eParticipation and semantic technologies.



Dimitris Askounis received his degree in electrical and computer engineering from the National Technical University of Athens in Greece in 1986 and PhD from the same university in 1991. Currently he is an Associate Professor in the School of Electrical and Computer Engineering of National Technical University of Athens. with more than 20 years research experience in the areas of Management Information Systems, Decision Support Systems and Intelligent Information Systems.