On Similarity in Fuzzy Description Logics

Eva Armengol[1] Pilar Dellunde[1,2] and Àngel García-Cerdaña[1]
[1] IIIA - CSIC
[2] Universitat Autònoma de Barcelona
Campus UAB, 08193 Bellaterra, Spain
{eva,pilar,angel}@iiia.csic.es

Abstract

In this paper we explore the possibility of introducing the equality symbol in the languages of Fuzzy Description Logics (FDLs) interpreted as a similarity relation. In the talk we will present a state of the art concerning the notion of similarity in some fields of artificial intelligence, and we analyze some variants of the languages for the FDLs introduced in [10]. The goal is twofold: dealing with attribute-value representations at the domain objects level, and integrating the treatment of similarities inside the description languages and their corresponding knowledge bases.

Description Logics (DLs) are knowledge representation languages built on the basis of classical logic. DLs allow the creation of knowledge bases and provide ways to reason on the contents of these bases. A full reference manual of the field is [1]. The vocabulary of DLs consists of *concepts*, which denote sets of individuals, and *roles*, which denote binary relations among individuals. From atomic concepts and roles and by means of *constructors*, DL systems allow us to build complex descriptions of both concepts and roles. These complex descriptions are used to describe a domain through a knowledge base (KB) containing the definitions of relevant domain concepts or some hierarchical relationships among them (*Terminological Box* or $TBox$), and a specification of properties of the domain instances (*Assertional Box* or $ABox$). One of the main issues of DLs is the fact that in both the $TBox$ and the $ABox$ can be identified with formulas in first-order logic or an extension of it; therefore we can use reasoning to obtain implicit knowledge from the explicit knowledge in the KB.

Fuzzy Description Logics (FDLs) are natural extensions of DLs for dealing with vague concepts, commonly present in real applications (see for instance [21, 22, 20, 2] or [16] for a survey). Hájek [13] proposed to deal with FDLs taking as basis t-norm based fuzzy logics with the aim of enriching the expressive possibilities in FDLs and to capitalize on recent developments in the field of mathematical fuzzy logic. From this perspective, in [10] a family of DLs was defined. These languages include truth constants for representing truth degrees, thus allowing the definition of the axioms of the knowledge bases as sentences of a fuzzy predicate language in much the same way as in classical DLs.

Similarity has been a central issue for decades in different disciplines, ranging from philosophy (Leibniz's Principle of the Identity of Indiscernibles) and psychology (Tversky's stimuli judged similarity) to natural sciences (taxonomy) and mathematics (geometric similarity). Also in artificial intelligence similarity plays an important role since the analogy reasoning is behind some of the early machine learning methods. Particularly, case-based reasoning methods are based on the principle that "similar problems have similar solutions" where the notion of similarity has a capital importance (see [14]).

In the *fuzzy framework*, the notion of similarity was introduced by Zadeh in [24] as a generalization of the notion of equivalence relation (see [17] for a historical overview on the notion of t-norm based similarity). As Zadeh pointed out, one of the possible semantics of fuzzy sets is in terms of similarity: the degree of membership of an object to a fuzzy set can be seen as the degree of resemblance between this object and prototypes of the fuzzy set. Thus, an important logical issue is to define a logic of similarity that can account for the proximity between the boolean interpretations of a propositional language. Dubois and Prade developed this idea in [8]. Ruspini [18] presents a formal characterization of the notions of both implication and consequence of propositional fuzzy logic in terms of the similarity notion between pairs of possible worlds. Based on his work, Esteva et al. [9] consider graded entailments between sets of propositional formulas induced by a similarity relation on the set of interpretations. Hájek studies similarities in fuzzy predicate logics and applies the obtained results to the analysis of fuzzy control [12]. Bělohlávek [4] presents a general theory of fuzzy relational systems. Gerla [11] proposed a fuzzy predicate theory whose fuzzy models are plausible candidates for the notion of approximate similarity. For a reference about model-theoretic properties of algebras with fuzzy equalities see [6] and [4].

Similarity in DLs has been studied by Borgida et al. [3] and D'Amato et al. [7] among others, focusing on similarity measures between DL concepts, and from the logical point of view by Sheremet et al. [19]. D'Amato et al. take as starting point the idea that measures for estimating concept similarity have to be able to appropriately consider concept semantics in order to correctly assess their similarity value. In accordance with this goal the authors propose a set of properties that a semantic similarity measure should have, analyze different extensional-based and intensional-based similarity measures proposed in the literature, and show that these approaches lack some of the needed properties. Finally, they define a measure for complex descriptions in some DL languages that is compliant with all of these criteria. Sheremet et al. propose an integration of logic-based and similarity-based approaches in classical DLs. They use concept constructors such as 'in the a-neighborhood of *C*' where a is a positive rational number; or the operator $C \leftleftarrows D$ which is interpreted by the set of all points in the similarity space that are closer to the instances of *C* than to the instances of *D*. For example, it can be used to model statements like 'X resembles *C* more than *D*'. Since we want to study the similarity in FDLs, we explore the possible generalization of both approaches, [19] and [7], to the fuzzy framework.

Our approach. In this paper we introduce the equality symbol in the language of FDLs and in their knowledge bases interpreted as a similarity relation. We also define variants of the languages for the FDLs introduced in [10]. These variants allow to deal with attribute-value representations at the domain objects level. The attribute-value representation is commonly used in artificial intelligence. In this representation, domain objects are sets of pairs attribute-value, where the value of an attribute may be qualitative or quantitative (see [23] to clarify the classes of data types). The *global* similarity between objects has to be seen as an aggregation of the *local* similarities of the features (see [15] and [5] for a collection of similarity and aggregation measures, respectively). To integrate the treatment of similarities inside knowledge bases, we can take profit of the presence of truth constants in FDL languages, as it is done in [10]. This fact allows to state graded notions at the syntactical level for both similarity between objects and similarity between concepts. As Ruspini suggests in [18], the degree of similarity between two objects A and B may be regarded as the degree of truth of the vague proposition “A is similar to B”. Thus, similarity among objects can be seen as a phenomena essentially fuzzy. Following this idea, the capabilities of a FDL language to deal with truth degrees at the syntactical level are specially relevant in which concerns to a useful treatment of the notion of *similarity* in the DLs framework. Therefore, the basic lines of our approach are:

- to represent domain objects by means of sets of pairs attribute-value as it is commonly done in artificial intelligence. Our goal is to integrate such representation in the conceptualization of the domain by means of FDLs.
- to explore the possibility of defining a Similarity Fuzzy Description Logic, introducing similarity roles and new concept constructors, the former interpreted as similarities between objects, the later as similarities between concepts. From our point of view the similarity between objects can be seen as an aggregation of the local similarities between attributes. Also, the similarity between concepts could be defined from similarities among objects.
- the use of the truth degrees included in the languages as similarity degrees between both objects and concepts.

References

[1] F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi, and P.F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003.

[2] F. Bobillo, M. Delgado, J. Gómez-Romero, and U. Straccia. Fuzzy description logics under Gödel semantics. *International Journal of Approximate Reasoning*, 50(3):494–514, 2009.

[3] A. Borgida, T. Walsh, and H. Hirsh. Towards measuring similarity in description logics. In *Procs. of the International Description Logics Workshop, Vol. 147 of CEUR*, 2005.

[4] R. Bělohlávek. *Fuzzy relational systems: foundations and principles*. Number 20 in IFSR International Series on Systems Science and Engineering. Kluwer Academic / Plenum Press, 2002.

[5] T. Calvo, G. Mayor, and R. Mesiar, editors. *Aggregation operators: new trends and applications*. Physica-Verlag GmbH, Heidelberg, Germany, Germany, 2002.

[6] P. Cintula and P. Hájek. Triangular norm based predicate fuzzy logics. *Fuzzy Sets and Systems*, 161:311–346, 2010.

[7] C. D'Amato, S. Staab, and N. Fanizzi. On the influence of description logics ontologies on conceptual similarity. In *EKAW '08: Proceedings of the 16th international conference on Knowledge Engineering*, pages 48–63, Berlin, Heidelberg, 2008. Springer-Verlag.

[8] D. Dubois and H. Prade. The three semantics of fuzzy sets. *Fuzzy Sets Syst.*, 90(2):141–150, 1997.

[9] F. Esteva, P. Garcia, and L. Godo. *Similarity-based reasoning*, volume 57 of *Stud. Fuzziness Soft Comput.*, pages 367–396. Physica, Heidelberg, 2000.

[10] À. García-Cerdaña, E. Armengol, and F. Esteva. Fuzzy description logics and t-norm based fuzzy logics. *International Journal of Approximate Reasoning*, doi:10.1016/j.ijar.2010.01.001, 2010.

[11] G. Gerla. Approximate similarities and Poincaré paradox. *Notre Dame J. Formal Logic*, 49(2):203–226, 2008.

[12] P. Hájek. *Metamathematics of Fuzzy Logic*, volume 4 of *Trends in Logic. Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 1998.

[13] P. Hájek. Making fuzzy description logic more general. *Fuzzy Sets and Systems*, 154(1):1–15, 2005.

[14] E. Hüllermeier. *Case-Based Approximate Reasoning*. Springer-Verlag, 2007.

[15] T. W. Liao, Z. Zhang, and C. Mount. Similarity measures for retrieval in case-based reasoning systems. *Applied Artificial Intelligence*, 12(4):267–288, 1998.

[16] T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *Journal of Web Semantics*, 6(4):291–308, 2008.

[17] J. Recasens. *A Small Course on Indistinguishability Operators*, 2007. Manuscript.

[18] E. H. Ruspini. On the semantics of fuzzy logics. *International Journal of Approximate Reasoning*, 5:45–88, 1991.

[19] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. A logic for concepts and similarity. *Journal of Logic and Computation*, 17(3):415–452, 2007.

[20] G. Stoilos, G. Stamou, J. Z. Pan, V. Tzouvaras, and I. Horrocks. Reasoning with very expressive fuzzy description logics. *Journal of Artificial Intelligence Research*, 30(8):273–320, 2007.

[21] U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.

[22] U. Straccia. A fuzzy description logic for the semantic web. In Elie Sanchez, editor, *Fuzzy Logic and the Semantic Web*, Capturing Intelligence, chapter 4, pages 73–90. Elsevier, 2006.

[23] Y. Yang, G. I. Webb, and X. Wu. Discretization methods. In *The Data Mining and Knowledge Discovery Handbook*, pages 113–130. 2005.

[24] L. A. Zadeh. Similarity relations and fuzzy orderings. *Information Sciences*, 3:177–200, 1971.