Fuzzy Ontologies for the Semantic Web

Elie Sanchez[1] and Takahiro Yamanoi[2]

[1] Université de la Méditerranée, LIF – Biomathématiques et Informatique Médicale
Faculté de Médecine, 13385 Marseille Cedex 5, France
elie.sanchez@medecine.univ-mrs.fr
[2] Hokkai Gakuen University, Division of Electronics and Information Engineering
Faculty of Engineering, Sapporo 064-0926, Japan
yamanoi@eli.hokkai-s-u.ac.jp

Abstract. It is presented several connections between Fuzzy Logic, the Semantic Web, and its components (Ontologies, Description Logics). It is then introduced and illustrated by an example ("Ontology of Art") a Fuzzy Ontology structure, Lexicon and Knowledge Base.

1 Introduction: Fuzzy Logic, Semantic Web and Ontologies

The field of Fuzzy Logic has been maturing for forty years. These years have witnessed a tremendous growth in the number and variety of applications, with a real-world impact across a wide variety of domains with humanlike behavior and reasoning. Fuzzy logic is now confronted with a new challenge, namely the vision of the Semantic Web. During recent years, important initiatives have led to reports of connections between Fuzzy Logic and the Internet [11,12]. Scattered papers were published on Fuzzy Logic and the Semantic Web, and a special session was organized during the previous IPMU conference [8]. Then, the first workshop on Fuzzy Logic and the Semantic Web (FLSW) [5] at Marseille was attended by European experts in the field. During BISC-SE 2005 at Berkeley, a panel [4, pp.27-30] discussed recent advances in these combined fields. A recently published volume [13] has shown the positive role Fuzzy Logic, and more generally Soft Computing, can play in the development of the Semantic Web. Finally, the Second Workshop on Fuzzy Logic and the Semantic Web (FLSW-II) will take place during IPMU 2006 at Paris. These are healthy symptoms that indicate, as we believe, that in the coming years, the Semantic Web will be a major field of applications of Fuzzy Logic.

The Semantic Web allows relational knowledge to be embedded as metadata in web pages enabling machines to use ontologies and inference rules in retrieving and manipulating data. *Ontologies* are a key component of the Semantic Web. There are several ways to describe the meaning of concepts (or classes of individuals or categories or types) and relationships between them. Ontologies facilitate a machine processable representation of information. They bridge an effective communication gap between users and machines.

There are many (descriptive) definitions of ontologies, depending also on communities. Basically, they are executable, formal conceptualizations with shared agreement

H. Legind Larsen et al. (Eds.): FQAS 2006, LNAI 4027, pp. 691 – 699, 2006.

between members of a community of interest. They can be viewed as "collections of statements written in a language such as RDF that define the relations between concepts and specify logical rules for reasoning about them. Computers *can understand* the meaning of semantic data on a web page by following links to specified ontologies" [3].

The most typical kind of ontology has a *taxonomy* and a *set of inference rules*. Note that besides the Semantic Web, ontologies have been studied in various domains, for ex. in knowledge engineering, natural language processing, knowledge management, information retrieval, digital libraries, electronic commerce, etc.

There are several types of ontologies, and one may consider, among others:

- *Upper-level (or generic or reference) ontologies.* They describe general concepts, like structure, space, time, state, substance, which are independent of a particular domain.

- *Domain ontologies*. They cover concepts in particular domains and in a *specific way* (for ex. human anatomy or E. coli) or in a *general way* (for ex. organs or gene function). They are the most common and agreed-upon types of ontologies.

- *Task (or application) ontologies.* They express conceptualizations relative to task models (for ex. reasoning processes for medical diagnosis).

Note that in biology, most ontologies are formed by a mixture of these three types.

The construction of an ontology implies the parallel construction of a vocabulary for it. As T. Gruber pointed out in [6], "pragmatically, a common ontology defines the vocabulary with which queries and assertions are exchanged among agents." But most of the information which relates to world knowledge is ill-structured, uncertain and imprecise. What is then needed is a collection of tools drawn from fuzzy logic, for example Zadeh's *PNL* (Precisiated Natural Language) [19,20]. Usually in the Semantic Web, knowledge is assumed to be crisply defined and no uncertainty or imprecision is allowed in the description of objects. The Semantic Web, as presented under *W3C* recommendations [17], deals with hard semantics in the description and manipulation of crisp data, like in "the Huveaune is a river." RDF based languages do not have the ability to represent soft semantics as in "the Huveaune is a *very_small* river." To process this type of information, fuzzy logic concepts and techniques are needed: "the Huveaune is a *very_small* river" can be translated into "length(Huveaune) is *very_ small*." It can then be encoded in RDF format with a triple

< Huveaune , length , very_small >,

where the term "very_small" is assumed to be the label of a fuzzy set *[note: the Huveaune is a 51 km long river that flows into the Mediterranean sea at Marseille].* It can be considered as a typed literal and an XML schema [18] can be defined to describe its membership function.

2 Fuzzy Ontologies

There has been different approaches to characterize or define *fuzzy ontologies*. In [16] the query refinement *PASS* System (Personalized Abstract Search Services) uses a fuzzy ontology of term associations to suggest alternative queries for searching for abstracts of research papers. The ontology is automatically generated. It provides information on sets of terms with *broader* and *narrower* semantic meaning.

The following chapters of [13], that we are going to survey, report an overview and recent advances on fuzzy ontologies. Our approach on Fuzzy Ontologies and related structures will be presented in section 4.

In *"On the Expressiveness of the Languages for the Semantic Web - Making a Case for 'A Little More'"*, Ch. Thomas and A. Sheth introduce the need for fuzzy-probabilistic formalisms on the Semantic Web, in particular within OWL. In *"Fuzzy ontologies for information retrieval on the WWW",* D. Parry uses fuzzy ontologies, and presents a broad survey of relevant techniques, leading up to the notions of fuzzy search and fuzzy ontologies.

The second section, "Fuzzy description logics for ontology construction," deals with fuzzy description logics in theoretical aspects and applications. In *"A Fuzzy Description Logic for the Semantic Web",* U. Straccia describes a fuzzy version of SHOIN(D), the corresponding Description Logic of the ontology description language OWL DL. He shows that its representation and reasoning capabilities go clearly beyond classical SHOIN(D). In *"What does mathematical fuzzy logic offer to description logic ?",* Petr Hajek proposes a fuzzy description logic based on fuzzy predicate logic, to deal with vague (imprecise) concepts. In "*Possibilistic uncertainty and fuzzy features in description logic. A preliminary discussion",* D. Dubois, J. Mengin, and H. Prade introduce another approach by injecting fuzzy features in Description Logics, this time based on fuzzy and possibilistic logic. In *"Uncertainty and Description Logic Programs over Lattices",* U. Straccia presents a Description Logic framework for the management of uncertain information. In this approach, sentences are certain to some degree, where certainty values are taken from a certainty lattice. Finally, in the last chapter of this section, *"Fuzzy Quantification in Fuzzy Description Logics",* D. Sanchez and A. Tettamanzi, introduce reasoning procedures for a fuzzy description logic with fuzzy quantifiers.

In *"Bottom-up Extraction and Maintenance of Ontology-based Metadata",* P. Ceravolo, A. Corallo, E. Damiani, G. Elia, M. Viviani and A. Zilli, present an approach to build fuzzy ontologies in a bottom-up fashion, by clustering documents, based on a fuzzy representation of XML documents structure and content.

In *"A fuzzy logic approach to information retrieval using an ontology-based representation of documents",* M. Baziz, M. Boughanem, G. Pasi and H. Prade, work on Information Retrieval using fuzzy ontologies. In *"Towards a Semantic Portal for Oncology using a Description Logic with Fuzzy Concrete Domains",* M. d'Aquin, J. Lieber and A. Napoli, present a work on encoding medical guidelines in fuzzy description logics and using that for a portal. The three systems that are presented are fully implemented within the KASIMIR oncology project. In *"Fuzzy Relational Oncological Model in Information Search Systems",* R. Pereira, I. Ricarte and F. Gomide, introduce another approach to Information Retrieval using fuzzy ontologies encoded by fuzzy relations.

Finally, in *"Evolving Ontologies for Intelligent Decision Support",* P. Gottgtroy, N. Kasabov and S. MacDonell, integrate soft computing techniques and ontology engineering. They investigate the rather different topic of evolving ontologies, presenting biomedical case studies.

3 Fuzzy Logic and Description Logics

Description Logics (DLs) [1] are a logical reconstruction of frame-based knowledge representation languages, that can be used to represent the knowledge of an application domain in a structured and formally well-understood way. For ex., here are some simple assertions:

"Human $\sqsubseteq$ Mammal"

"Woman $\sqcap$ Parent $\equiv$ Mother"

"(Human$\sqcap\neg$Male)$\sqcap\exists$married.Biologist$\sqcap$($\geq$3hasChild)$\sqcap\forall$hasChild.student" (denotes "a woman who is married to a biologist and has at least three children, all of whom are students.")

They are considered as a good compromise between expressive power and computational complexity. DLs are essentially the theoretical counterpart of the Web Ontology Language OWL DL [7], the state of the art language to specify ontologies. DLs can be used to define, integrate and maintain ontologies (see [2].) But these DLs embodied in Semantic Web languages do not allow a treatment of uncertainty and imprecision encountered in real-world applications. To this end, DLs have been extended with fuzzy capabilities, yielding *FDLs* (Fuzzy Description Logics) in which concepts are interpreted as fuzzy sets. In [14] U. Straccia has extended the DL $\mathcal{ALC}$, a significant representative of DLs. For example a *concept* C of the language $\mathcal{ALC}$ has an *interpretation* I, which is a pair I = (Δ^I, $\bullet^I$) consisting of a *domain* Δ^I and an *interpretation function* $\bullet^I$. In FDL, a concept C is interpreted as a fuzzy set and a statement like "a is C" has a truth-value in [0,1]. In this case, $\bullet^I$ is an *interpretation function* mapping C into a membership function C^I, $C^I : \Delta^I \rightarrow [0,1]$. Acting on concepts, the crisp operations of conjunction, disjunction, negation and quantification are naturally extended to their fuzzy counterparts [14]. Here are some examples of assertions, involving fuzzy sets [15]:

"YoungPerson = Person$\sqcap\exists$age.Young", where *Young* is the label of a fuzzy set;

"SportsCar = Car$\sqcap\exists$speed.very(High)", where *very* is a concept modifier and *High* is a fuzzy concrete predicate over the domain of speed expressed in km/h.

4 Fuzzy Ontology Structure, Lexicon and Knowledge Base

The following developments are an extension of work on crisp ontologies [10]. Now a *fuzzy ontology structure* and other associated structures will be introduced: *lexicon* and *knowledge base*.

Basically, a fuzzy ontology structure can be defined as consisting of concepts, of fuzzy relations among concepts, of a concept hierarchy or taxonomy, of non-hierarchical associative relationships and of a set of ontology axioms, expressed in an appropriate logical language.

Then, a lexicon for a fuzzy ontology will consist of lexical entries for concepts (knowledge about them can be given by fuzzy attributes, with context-dependent

values), of lexical entries for fuzzy relations, coupled with weights expressing the strength of associations, and of reference functions linking lexical entries to concepts or relations they refer to.
So, a Fuzzy Ontology structure can now be defined as follows.

Definition 1. A *Fuzzy Ontology structure* is a quintuple

$$O := (C, R, T, A, X),$$

where

- C is a set of *(fuzzy) concepts* (or *classes* — cf. in OWL — of individuals, or *categories*, or *types*), for ex.: Mountain, Patient, Cell, Diabetes, Pneumonia, Fracture of neck of femur, etc. Concepts can be *primitive concepts* or *defined from other concepts*, for ex.: "*Prokaryotic cells* are cells that do not have a nucleus." Note the fuzziness, and vagueness, that can be inherent in the definition of concepts. For ex. a *hill* is "a landform that extends less than 600 metres above the surrounding terrain (the Encyclopædia Britannica requires a prominence of 2,000 feet — 610 m — for a *mountain*) and that is smaller than a mountain." Many hills are higher than 600 metres, but hills are generally smaller (note that in the definition of a *mountain,* one has: a *mountain* is generally much higher and steeper than a *hill*, but there is considerable overlap ...)."
- R is set of (*fuzzy) relations* (or *roles*, or *slots*) in C x C, for ex.: "the concept *Nucleus* has a *part-of* relationship with the concept *Cell*" or "a *Very Tall* person is *Tall*."
- T is a relation in C x C, called *Taxonomy* (or *concept hierarchy*). It organizes concepts into sub-(super-)concept tree structures, most commonly in *Specialisation* relationships (for ex. "an *enzyme* is_a *protein"* or "*cancer* is_a *disease"*) or in *Mereological* (or *Partonomic*) relationships (for ex. $T(C_1, C_2)$ means that C_1 is a subconcept of C_2, like in "*eukaryotic cells* are *cells* that have a nucleus").
- A is a set of non-taxonomic (fuzzy) *Associative* relationships that relate concepts across tree structures, for ex.:
  - *Naming* relationships, describing the names of concepts
  - *Locating* relationships, describing the relative location of concepts
  - *Functional* relationships, describing the functions (or properties) of concepts
- X is a set of *Ontology Axioms* (or *rules*), expressed in an appropriate logical language, for ex. asserting class subsumption, equivalence, more generally to (fuzzily) constrain the possible values of concepts or instances (for "*instances*," see below).

Now, a *lexicon* is a list of words in a language, a *vocabulary*, including some knowledge of how each word is used. Each word, or group of words, in a lexicon is described in a *lexical entry*. So, a lexicon can be viewed as an index that maps a written form of a word to information about that word. Let us now define a lexicon associated with a fuzzy ontology structure.

Definition 2. A *Lexicon* for the fuzzy ontology structure $O := (C, R, T, A, X)$ is a quadruple

$$L := (L^C, L^R, F, G)$$

consisting of:

- a set L^C of lexical entries for concepts; knowledge about them is given by (fuzzy) attributes, with context-dependent values
- a set L^R of lexical entries for (fuzzy) relations, from C to C, coupled with weights in [0,1] expressing the strength of associations
- two *reference functions* F and G ($F : L^C \rightarrow 2^C$ and $G : L^R \rightarrow 2^R$) that link lexical entries of L^C, resp. of L^R, to the set of concepts, resp. of relations, they refer to.

Because an ontology is a conceptualization of a domain, it is not supposed to contain instances, hence the following definition.

Definition 3. A *Fuzzy Knowledge Base* structure is a couple

$$KB = (O, I)$$

where $O := (C, R, T, A, X)$ is a Fuzzy Ontology structure and I is a set of *Instances* (or *Individuals*) associated with the ontology, i.e. 'objects' represented by a concept. For ex.: *Haptoglobin* is an instance of the concept *Protein*, or in *"Carol has Diabetes"*, *"Carol"* is an instance/individual and *"Diabetes"* is a concept/class.

Illustrative Example

(Fuzzy) Ontology Structure (see figure 1): Ontology of Art

$C = \{C_1,C_2,C_3,...,C_{24}\}$
$R = \{r_1,r_2,r_3,...,r_9\}$
$T(C_2,C_1)$, $T(C_3,C_1)$,..., $T(C_6,C_1)$, $T(C_{15},C_{16})$
$r_1(C_4,C_7)$, $r_1(C_4,C_8)$, $r_1(C_4,C_9)$
$r_2(C_4,C_{10})$, $r_2(C_4,C_{11})$, $r_2(C_4,C_{12})$
$r_3(C_4,C_{13})$
$r_4(C_4,C_{14})$, $r_4(C_{13},C_{14})$
$r_5(C_5,C_{15})$
$r_6(C_4,C_{17})$, $r_6(C_4,C_{18})$, $r_6(C_4,C_{19})$, $r_6(C_5,C_{17})$, $r_6(C_5,C_{18})$, $r_6(C_5,C_{19})$
$r_7(C_{19},C_{20})$
$r_8(C_5,C_{21})$, $r_8(C_5,C_{22})$, $r_8(C_5,C_{23})$
$r_9(C_{23},C_{24})$

Lexicon (see figures 2 and 3)

L^C = {artist, musician, singer, painter, sculptor, dancer, ..., marble, region}
L^R = {paints, uses, is_influenced_by,…, creates, comes_from}
F(artist) = $\{C_1\}$, F(musician) = $\{C_2\}$, F(singer) = $\{C_3\}$, F(painter) = $\{C_4,C_{13}\}$, …, F(region) = $\{C_{24}\}$
$T(C_2,C_1)$ = is_a, $T(C_3,C_1)$ = is_a, ..., $T(C_{15},C_{16})$ = is_a
G(paints) = $\{r_1\}$, G(uses) = $\{r_2, r_8\}$, G(creates) = $\{r_5\}$,…, G(comes_from) = $\{r_9\}$.

Relations in R (r_i's) can be associated with weights, expressing the strength of relations, or linguistic quantifiers. For ex. "a painter (C_4) paints (r_1) *mostly* still life (C_9)" or "he prefers using (r_2) water colors (C_{10}), with a *weight of preference* $w_{4;10}$ in [0,1]" or "he is influenced_by (r_3) a *famous* painter (C_{13}), with weight $w_{4;13}$."

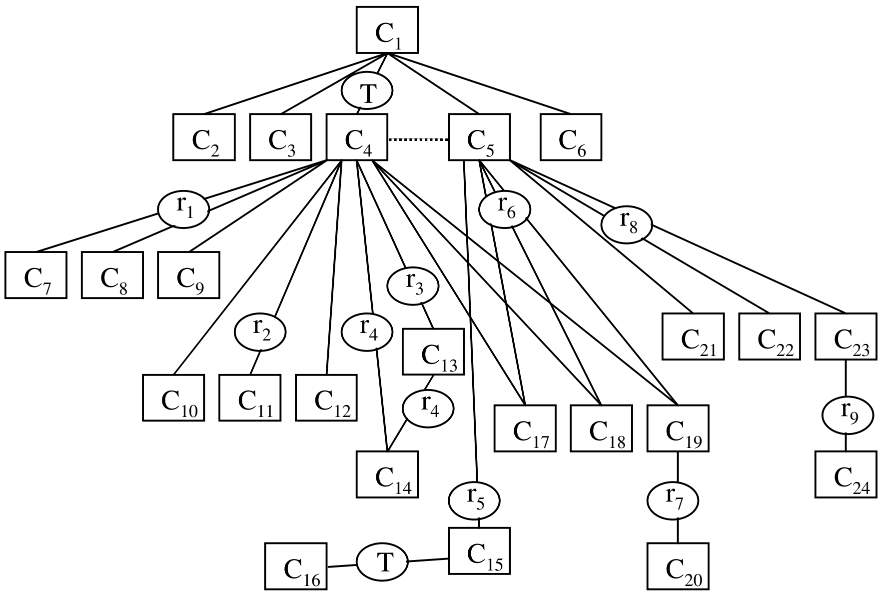

Fig. 1. Ontology of Art

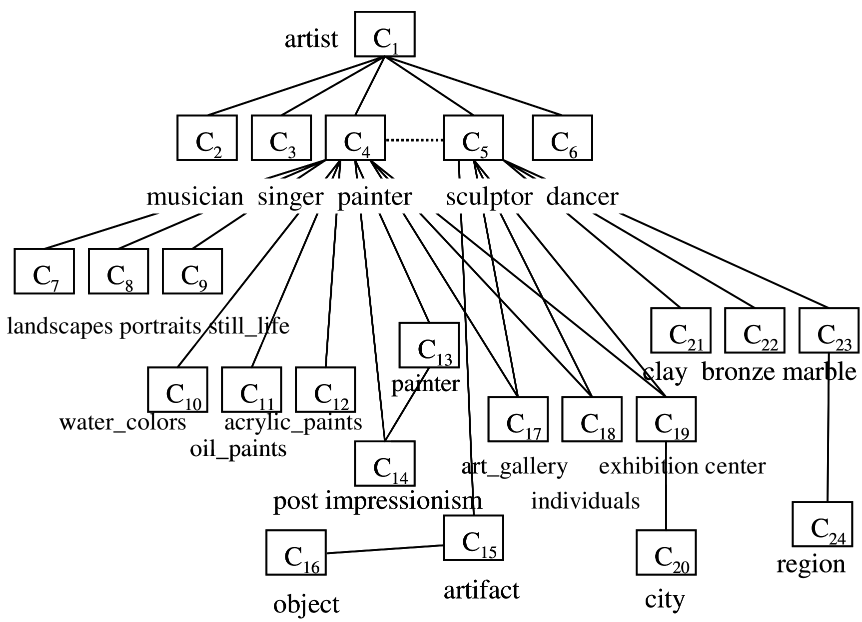

Fig. 2. Illustration of L^C

Epistemic (knowledge-directed)Lexicons
A concept C_i can be associated with an *epistemic lexicon* [19] $K(C_i)$, expressing *world knowledge* about it. $K(C_i)$ is organized into relations, with entries defined as (fuzzy) distribution-valued attributes of $K(C_i)$ that are context dependent. Ex. for a painter C_4: degree of *notoriety*, *usual* residence, etc.

Knowledge Base
Considering the illustrative example above, a *(Fuzzy) Knowledge Base* can be constructed with a collection of *Instances.* Ex. for a painter C_4: "Paul Cézanne", who *painted* landscapes C_7 ("Mount Sainte-Victoire"), still life C_9 ("Apples and Peaches"). He *belonged to art movement* C_{14} ("postimpressionism"). He was *influenced by* C_{14} ("Gustave Courbet" or "Edouard Manet" or "Camille Pissarro", etc.).

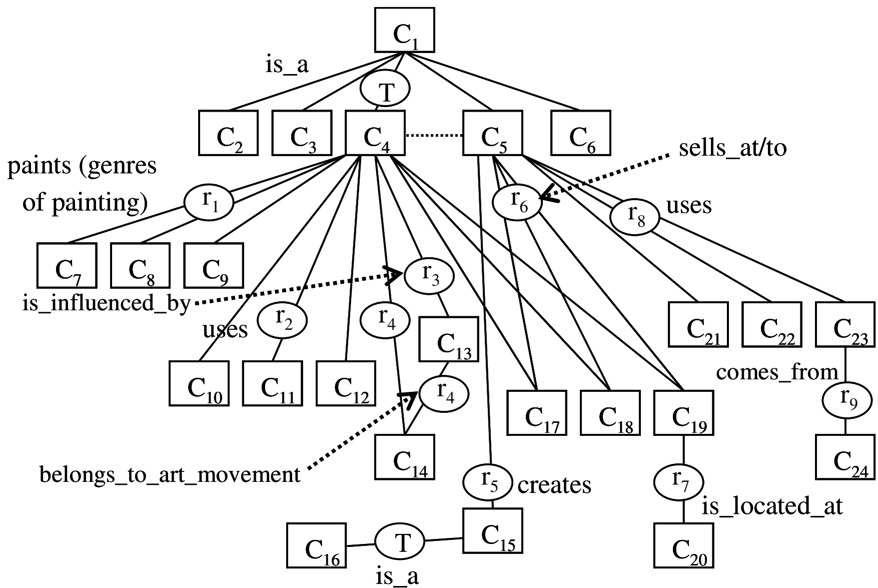

Fig. 3. Illustration of L^R

5 Concluding Remarks

"The success of the deployment of the Semantic Web will largely depend on whether useful ontologies will emerge, allowing shared agreements about vocabularies for knowledge representation." [9]

The vision of a Semantic Web Wave is attracting much attention in the scientific world. Design, implementation and integration of ontologies will be crucial in the development of the Semantic Web. A Semantic Web, in which Fuzzy Logic, and more generally Soft Computing, will certainly have a positive role to play.

It is expected that the structures that have been introduced in this paper around the notion of Fuzzy Ontology will enrich the ingredients that will contribute to the real success of the Semantic Web.

References

1. F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P. Patel-Schneider, "The Description Logics Handbook: Theory, Implementation and Applications", Cambridge university Press (2003).
2. F. Baader, I. Horrocks, U. Sattler, "Description Logics as Ontology Languages for the Semantic Web", in D. Hutter, W. Stephan (Eds.), Festschrift in honor of Jörg Siekmann. Lecture Notes in Artificial Intelligence, Springer-Verlag (2003).
3. T. Berners-Lee, J. Hendler, O. Lassila, "The Semantic Web", *Scientific American*, May 2001.
4. BISC-SE 2005, BISC Special Event in Honor of Prof. Lotfi A. Zadeh, "Forging New Frontiers", Proceedings: M. Nikravesh (Ed.), Memo No. UCB/ERL M05/31, Nov. 2, 2005.
5. "Fuzzy Logic and the Semantic Web" Workshop, Extended abstracts available at: http://www.lif.univ-mrs.fr/FLSW , Marseille, France, 2005.
6. T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing". Originally in N. Guarino and R. Poli (Eds.), *Int. Workshop on Formal Ontology*, Padova, Italy (1993). Published in Int. J. of Human-Computer Studies, Vol.43, Issue 5-6 (1995) 907-928.
7. I. Horrocks, P.F. Patel-Schneider, F. van Harmelen, "From SHIQ and RDF to OWL: the making of a Web Ontology Language", *J. of Web Semantics* 1, 1 (2003) 7-26.
8. IPMU 2004, Special Session "Fuzzy Logic in the Semantic Web: a New Challenge", Proc. pp.1017-1038, IPMU04@dipmat.unipg.it, Perugia, Italy, 2004.
9. O. Lassila and D. L. McGuinness: " The Role of Frame-Based Representation on the Semantic Web ", Knowledge Systems Laboratory Report KSL-01-02, Stanford University (2001).
10. A. Maedche, "Ontology Learning for the Semantic Web", Kluwer Academic Publishers (2002).
11. "New directions in Enhancing the Power of the Internet" (Proceedings UCB/ERL, Berkeley, Memo N° M01/28, August 2001) and "Enhancing the Power of the Internet", M. Nikravesh, B. Azvine, R. Yager and L.A. Zadeh (Eds.), Springer Verlag, 2004.
12. NAFIPS-FLINT 2002 Int. conf., IEEE SMC Proceedings 02TH8622, New-Orleans, 2002.
13. E. Sanchez (Editor), "Fuzzy logic and the Semantic Web", Elsevier (2006).
14. U. Straccia, "Reasoning with Fuzzy Description Logics", *J. of Artificial Intelligence Research* 14 (2001) 137-166.
15. U. Straccia, "A Fuzzy Description Logic for the Semantic Web", in E. Sanchez (Ed.), "Fuzzy logic and the Semantic Web", Elsevier (2006) 73-90.
16. D. H. Widyantoro, J. Yen, "Incorporating Fuzzy Ontology of Term Relations in a Search Engine", *2001 BISC Int. Workshop on Fuzzy Logic and the Internet,* Proceedings (2001) 155-160.
17. www.w3.org
18. www.w3.org/XML/Schema
19. L. A. Zadeh, "Web Intelligence and Fuzzy Logic – The concept of Web IQ (WIQ)", Invited talk at the 2003 IEEE/WIC Int. Conference on Web Intelligence (WI 2003), Halifax, Canada, available at: www.comp.hkbu.edu.hk/IAT03/InvitedTalk1.htm
20. L.A. Zadeh, "From Search Engines to Question-Answering Systems — The Problems of World Knowledge, Relevance, Deduction and Precisiation", in E. Sanchez (Ed.), "Fuzzy logic and the Semantic Web", Elsevier (2006) 163-210.