On the Expressiveness of the Languages for the Semantic Web - Making a Case for 'A Little More'

Christopher Thomas and Amit Sheth[*]
Large Scale Distributed Information Systems lab,
University of Georgia, Athens, GA USA

"So far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain, they do not refer to reality."
Albert Einstein

Abstract

The best pair of shoes you can find are the ones that fit perfectly. No inch too short, no inch too wide or long. The same, of course, holds for applications in all fields of computer science. It should serve our needs perfectly. If it does more, it usually comes with a tradeoff in performance or scalability. On top of that, for logic based systems, the maintenance of a consistent knowledge base is important. Hence a decidable language is needed to maintain this consistency computationally. Recently, the restriction of the Semantic Web standard OWL to bivalent logic has been increasingly criticized for its inability to semantically express uncertainties. We will argue for the augmentation of the current standard to close this gap. We will argue for an increased expressiveness at different layers of the cake and we want to show that only a spiced up version of some of the layers can take the blandness out of it. We want to show that it is possible to have a mixture that can account for reasoning based on uncertainties, possibilistic measures, but also for other epistemologically relevant concepts, such as belief or trust.

Keywords: Semantic Web, Fuzzy Logic, Probability Theory, Knowledge Representation

1. Introduction

During the 2004 WWW conference, Tim Berners-Lee, the spiritual leader of the semantic web community, made a statement that might impact further research on semantic web foundations. Asked whether it is necessary to have a representation for uncertainty on the layered semantic web stack, his answer was basically "no". Knowing about the impact of a person like Berners-Lee, this view might have great impact on the community and hence on unfortunate students and researchers funded for the development of uncertainty calculi for the semantic web. But aside from personal considerations, this answer can be challenged on many levels. After all, history has shown many times that "the revolution eats its children".

[*] Corresponding author. Tel.: +1-706-542-2310; fax: ++1-706-542-4771.
E-mail addresses: chaos@uga.edu (Christopher Thomas), amit@cs.uga.edu (Amit Sheth).

When talking about extending the kinds of logical inference used in local or global knowledge bases, we have to be aware of the different ways in which this can be done. The semantic web standard assumes a monotonic bivalent logic. Bivalent means that statements are either true or false; no third possibility, such as unknown, exists, nor anything in between true and false. In the remainder of the paper we will refer to these bivalent logics as FOL (first order logics).

Research in logic has produced a plethora of different formalisms which are more or less used in real-world applications. There are many-valued logics, continuous-valued logics with different semantics of their value assignments, such as fuzzy logics and probabilistic logics [12; 14], non-monotonic logics [36], paraconsistent logics [4] and combinations thereof [9]. One can make a case for the use of any of these logics for different applications. The question for semantic web researchers is, whether to augment the famous layer cake by adding a different layer for non-bivalent and/or non-monotonic logics or whether these augmentations should exist locally, for groups of agents that agree to their own representation of non-FOL statements.

Berners-Lee's main argument for the rejection of a representation of uncertainty at the very basis of Semantic Web formalisms was their lacking scalability. This is a strong argument. For example, in order to do complete probabilistic inference on dependent objects in a Bayesian Network or any other technology for probabilistic inference, potentially every path through the network has to be taken into account. The same holds for any kind of non-monotonic logic. Given that, even in the case of FOL, such as SHIQ [17] or similar Description logics, inference cannot be done in a computationally efficient way, it seems mandatory to keep the evaluation of statements local.

The major problem that monotonic knowledge bases face is that of inconsistency. In a monotonic logic, such as each of the various flavors of bivalent Description Logics and hence OWL, it is assumed that if a true statement can be derived from a set S of facts and rules, then it can also be derived from every larger set S' that contains S. This is an appealing assumption to make, because it allows reasoning to be local and to only take into account the rules and facts that are immediately necessary to deduce a new statement. But it is also an unrealistic assumption to make, because the world, even the formalized one, is full of contradictions. And the more this world of formal statements grows, the more likely it is that a contradiction occurs. For this reason, the CYC® Ontology is partitioned into consistent micro theories or contexts [30]. Inconsistencies with other micro theories are explicitly stated. To give an example of how easily inconsistencies can occur when we get to know more about a domain, we cite an example that is widely used in AI textbooks. It involves birds. We will state the knowledge base informally:

- Birds fly ($\forall x$: bird(x) $\rightarrow$ fly(x))
- Penguins are birds ($\forall x$: penguin(x) $\rightarrow$ bird(x))
- Joe is a penguin (penguin(Joe))

We can deduce now, that Joe is a bird and hence can fly. But there's one more rule that, for the sake of accuracy, could be added to the domain representation.

- Penguins don't fly ($\forall x$: penguin(x) $\rightarrow$ $\neg$fly(x))

This statement causes an inconsistency in the knowledge base. If we follow the path Joe is a penguin, penguins are birds, birds fly, then Joe flies. If we follow the path Joe is a penguin, penguins don’t fly, then Joe doesn’t. The addition of a rule made the knowledge base inconsistent, even though it was a sensible rule to add. And locally, the knowledge base was also correct before this rule was added. It was just missing some information and it still is missing information. Complete formalized knowledge of a domain is an illusion.

It is no question that inconsistencies arise when combining knowledge from multiple sources. It is also no question that most real-world phenomena cannot be sufficiently described in a bivalent logic. This holds in two cases. We can look at this from a scientific point of view, where we have to deal with uncertainty, randomness and partial knowledge. In this case even extremely accurate measurements result in uncertainties about the implications of the obtained data. The second case is that of knowledge representation and computation based on human perception as described in [38]. Here, non-bivalence and inconsistencies occur because of a fundamental inability to conduct completely accurate measurements on the one hand but the corresponding human ability to accurately reason with this inaccurate knowledge. The question for the Semantic Web research is, though, whether the basis of Semantic Web standards should be the bivalent monotonic case and the uncertain and/or non-monotonic cases are a specialization thereof, or whether the basis should be the uncertain case with bivalent monotonic logics as a special case.

1.1. Science

In essence, it is also a question about the extent to which the use of ontologies will have an impact on how humans gain and evaluate knowledge. The time we are living in is partially signified by the specialization of the scientific fields. The time in which an Aristotle could have all scientific knowledge of his cultural realm is long gone. It is even impossible to know about all current developments in adjacent fields. Over time, science evolved into a hydra who’s heads are autonomous and quite ignorant about what happens in most of it’s other heads, a giant distributed apparatus that is striving towards progress without the ability to take a comprehensive snapshot of the current state. Hypotheses are generated, but are mostly evaluated against the knowledge of a small field. A formalization of these hypotheses could allow the scientific communities to extensively evaluate their findings. Agents could roam the web and cross check hypotheses with data obtained by other labs and compare the results and hypotheses generated from the data to find similarities and contradictions.

In a scientific environment or in general in fields that are mainly driven by empirically closing the gaps of incomplete knowledge of their domain, it is necessary that the knowledge bearing agent is aware of the partiality of its knowledge. In a localized system, that is easily accomplished. If we want to share information across the internet, then the formalism that encapsulates this knowledge needs to reflect this partiality, so it can be propagated to other agents. Otherwise it is like telling rumors. I heard from my not very trustworthy acquaintance Paul that Peter and Mary broke up recently. While I might still have propagated it as a rumor that I did not really trust to Cathy, she might have told it as a fact to Marc who was Peter’s best friend but always had his eyes on Mary. After asking her out on a date, his Friendship with Peter was shattered. If every link in this gossip chain had kept the information that it was just a rumor, Marc and Peter would be drinking beer now.

1.2. Information Retrieval

In information retrieval, we are dealing with all sorts of noisy data. Web search engines and proposed question answering systems that are based on information available on the internet or in local ontologies will on average yield results that are not totally relevant or reliable. The annotation of data with metadata will improve this situation, but not fundamentally change it. Hence, an estimate of relevance or truth, delivered together with the answer or answer set, would be desirable.

Recently, more statistical techniques have emerged in internet search engines. While Google's page rank [7] measures the popularity of a page and makes the implicit assumption that popularity is somewhat linked to importance, other search engines use clustering to order their search results. The Vivisimo search engine [42] builds a hierarchy for the search results using a hierarchical clustering technique.

The Vivisimo approach is only the first of many search engines that give a hierarchical structure of their search results. We will also encounter situations in which we want to build ontologies or topic maps on the fly from sets of documents that have not yet been classified1 [20, 28]. Any automatic approach to ontology generation, be it statistical, using pattern based, NLP or other techniques, must be uncertain in nature and the map must reflect its own uncertainty about the inferred knowledge.

In the Semantic Web context, this uncertainty cannot just be stored within the generating or retrieving agent, but must be propagated through the deductions or other reasoning steps that are performed with the information. In a future situation, when a derived fact is annotated with the trace of its derivation (i.e. its provenance information), this will become clear. In community based knowledge accumulation, such as an augmented "Semantic Wikipedia" project [40], the community could assign degrees of belief to the formalized facts. Each agent in turn can assign degrees of belief to selected ontologies or single facts within ontologies. The latter is an individual assignment and doesn't require a formalization, but the former needs to be made available to general purpose reasoners, if we want the Semantic Web vision of connectivity between heterogeneous knowledge sources become a reality.

For the next generation of web search engines, question answering machines, it will be crucial to have a powerful representation of vagueness, because the terms humans use to describe the world will be vague and ambiguous [27, 29, 38].

1.3. Semantic Web

Semantic heterogeneity is an inherent reality of the internet. It is possible to encompass all types of syntactic, schematic and semantic heterogeneity that have been discussed in research on heterogeneous databases [21, 23, 32]. This diversity makes it the most useful, ubiquitous and unbiased collection of knowledge that has ever existed. The challenge to a computational use of this smorgasbord of opinions is to keep the diversity while still having a sound basis for computation. Inability of first order to model heterogeneity even in relational databases was conclusively discussed [25]. In the real world realization of the Semantic Web, very few ontologies will be consistent with each other. These inconsistencies can arise for multiple reasons. The easiest resolution is that

[1] The term ontology usually implies a rigid knowledge representation which meets formal criteria of inheritance or soundness/completeness issues. The term topic map has a less formal connotation.

of ambiguity in terms and actual formalization while still agreeing on the conceptualization as such. But ontologies can also be fundamentally inconsistent, because the ontological and metaphysical basis is different, i.e. the ontologies don't share the same view of the world or their domain, or the factual knowledge used to populate them come from different sources whose observations are inconsistent with each other.

This paper is meant to show that, despite computational disadvantages, many signs point towards incorporating the representation of more powerful semantics into the Semantic Web layer cake. While it is computationally impossible to logically track every statement on the web and use it in every logical inference, a restriction to deduction within local ontologies based on FOL will only exploit the local knowledge, but never help acquiring new knowledge. Hence the Semantic Web formalisms need to account for the possibilities that inductive and abductive inferences offer alongside with the ability to combine ontologies that, in FOL, would result in inconsistent and hence unusable knowledge.

The discussion about having representations of uncertainty, fuzzy reasoning and continuous representations in general is less concerned about whether this representation is necessary, but where it is necessary. Looking at the problems that non-classical logics pose to inference mechanisms, the placement has to be done carefully. The easy way out is to argue that each community that relies on a non-classical representation, does so in their own ontology with their own agents processing this information. There is a lot of merit to this argument, but we are going to show that it will not be sufficient in the long run, if the Semantic Web is to be exploited to its full potential.

In a current scenario, in which we see the mere beginning of the application of Semantic Web technologies, ontologies tend to be a convenient way of porting database information. This is a large step by itself, but while it allows machines to read and maybe process data from multiple sources, it doesn't give us a way of reliably putting derived knowledge from these multiple sources back on the web. With human intervention, we filter information that we think wasn't derived in a scientifically viable way and we have the ability to trust and distrust sources. In most current applications, there is only one step of computational inference between two tasks accomplished by humans. If now we are trying to have an automated chain of inference, the agents involved need to tell each other about the confidence they have in the truth of their derived information, which is in turn a function of the reliability of their sources and their inference methods. An internal procedure, even a very sophisticated one that allows an agent to filter out information that according to its criteria does not exceed a certain threshold of reliability will not be of any use, if it propagates the remaining inferred information as completely reliable information or knowledge.

The Semantic Web has the potential to make an immense impact on human society in general and on science in particular. The decisions made in the beginning will be difficult to revise at a later stage. We see, for example, how hard it is to change from a 32bit IP address to a 128bit address. For these reasons we need to be careful what we design as the foundational formalisms of this Semantic Web. We need to make sure that we promote knowledge, deepen and broaden it. But alongside knowledge goes doubt. Every careful representation of knowledge needs to be aware of the fallibility of each statement. The Semantic Web offers a great opportunity, but it might also just offer the chance to augment existing junk by a whole lot of meta-junk. We want to improve, not worsen the situation that T.S Eliot depicts: “Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?” (T. S. Eliot)

This paper attempts to name fundamental issues with the use of non-classical representations on the Semantic Web. It does not intend to solve them. Much research has been done in single fields, which we are going to address and correlate with the problems that the semantic web is facing.

The remainder of the paper is organized as follows. First, we give two more detailed examples of real-world situations in which we need a representation of uncertainty. In section 3, we will briefly introduce several kinds of logics that exceed the bivalent paradigm. In section 4 we will outline the envisioned framework of logics for the semantic web. Section 5 is concerned with issues of soundness, completeness, consistency and computability. Section 6 sketches the shortcomings of OWL as a basis for the framework and section 7 introduces related work from the fields of Artificial Intelligence, Computer Science, Philosophy and Psychology. Section 8 finally concludes the sketch of this framework and addresses open questions. The discussion of the logical formalisms is mostly held on an informal level. Essentially we are presenting a vision rather than a theory or its implementation.

2. Some motivating examples

2.1. Glycomics

In a project funded by the National Center for Research Resources [39], the LSDIS lab and the Complex Carbohydrate Research Center at UGA are building an ontology to model the interaction of complex carbohydrates with genes, proteins, enzymes and cells.

Human cells have complex carbohydrates on their surfaces, and various experiments have shown that the surface carbohydrates of some invasive cancer cells have unusual structures. Therefore, biochemists are very interested in the relationship between these atypical cell surface carbohydrates and the process of metastasis (invasion by cancer cells).

The complex carbohydrates are synthesized by enzymes that are encoded by genes, and the amount of each enzyme depends on the level to which the gene is activated in the cell. Although it is relatively straightforward to determine whether a gene is "turned on" in a tissue (i.e., whether it is actively making the RNA message that leads to production of a specific enzyme), characterization of complex carbohydrates at the cell surface is technically challenging.

Nevertheless, one can find correlations between the presence of certain complex carbohydrates on the cell surface and the activation of specific genes (such as glycosyl transferases). There is a general but well-defined sequence of events starting from gene activation to carbohydrate synthesis, but the players (specific genes and enzymes) involved in the production of a particular carbohydrate are often not known with certainty. Correlative evidence (e.g., gene homologies and relationships between the amount of carbohydrate and the level of gene activation) is often available, but not sufficient to say with certainty that "activation of gene X results in the production of enzyme Y, which synthesizes carbohydrate Z." Similarly, the hypothesis that the carbohydrate is involved in causing a cancer cell to become invasive is based on correlative evidence. Hence, observation of a measurable quantity - the level of gene activation - only gives us a likelihood of the presence of the abnormal glycoprotein and that only provides the likelihood that the cancer cell is invasive. Additional evidence for these relationships must be accumulated piecemeal. Every bit of evidence modifies the likelihood of the initial hypothesis, which may eventually have to be rejected if the probability of its accuracy becomes too low.

That in cases like this the representation of likelihoods is absolutely necessary is not disputed. But since BioInformatics data is usually obtained and used by experts within the BioInformatics domain, it is not necessary to have a representation of these uncertainties at the lowest possible level, viz. the level of the semantic web layer cake. BioInformatics could agree on an upper ontology that is used for all representations of uncertainties within this field. On the other hand, the next example shows a more common scenario that will arise when we are using advanced Semantic Web technologies for information retrieval. If situations like the following shows the need for a representation of uncertainty at a lower level, we will have a reason to argue for its implementation in the layer cake.

2.2. Information Retrieval

Imagine an advanced information retrieval situation. A question we could ask the question answering system would be “what is the population of New York City?” The system could either return one value or it could ask several other question answering systems and return all answers ranked by some sort of confidence value. It could also return the one that was given with the highest confidence or an average value. In any case, it can be given directly to the user. Hence no formal representation is necessary.

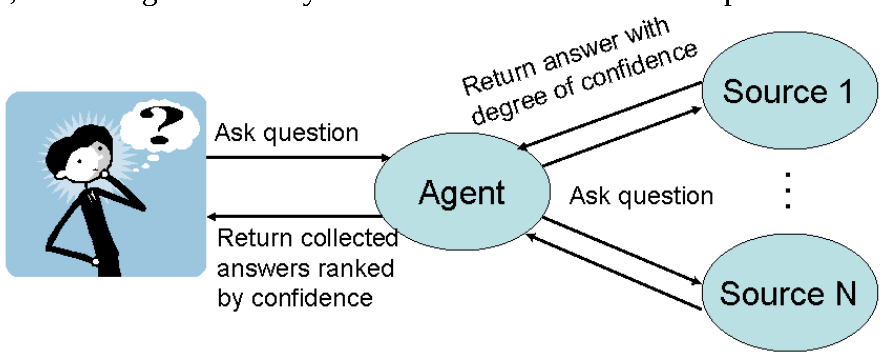

Figure 1: Simple scenario, one step between user and question answering systems

A more complicated case arises when we ask an agent to deduce an answer. We could, for example, ask the agent whether the largest city in the USA has more inhabitants than the largest city in Mexico. Now the system has to deal with several levels of uncertainty. First of all, it has to retrieve the name of the largest cities of both countries. Depending on the systems it consults, this could potentially return different answers with varying degrees of certainty. Then it has to take those results and ask again for the population of the cities. In the deductive step, the agent takes either all or the most confident results of the queries and performs the comparison. Depending on the settings, the user sees either the answer that was deduced with the highest confidence or all answers ranked by their degrees of confidence.

In this case, the representation of uncertainty needs to be on a fairly low level, preferably directly within the Semantic Web layer cake. We need to assume that the different agents that are consulted by the deducing agent do not agree on the same domain ontology, but use all the same knowledge exchange format. Hence uncertainty does not only arise due to an initial uncertainty returned by the QA systems, but also due to different ontological representations that need to be merged.

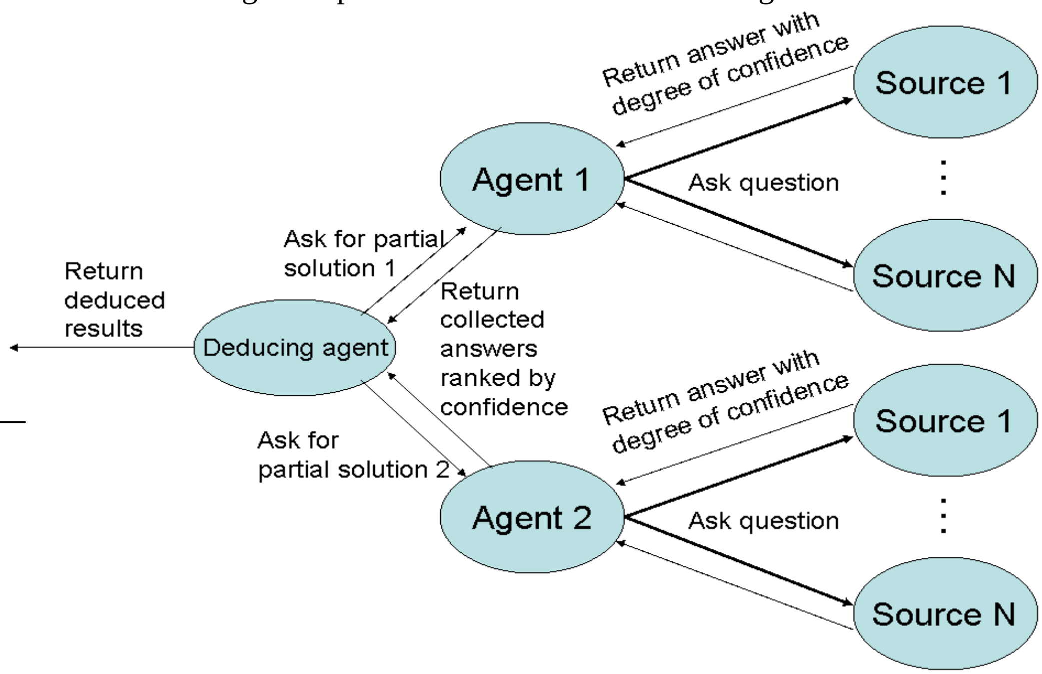

Figure 2: Complex scenario, involving deductive steps between user and QA systems

3. *Challenges for multi-valued logics*

3.1. The fuzzy case

Fuzzy logic has a special place amongst those logics that allow the assignment of continuous values. In its simplest case, it exhibits monotonicity. This means that the scalability argument cannot be used against an application of a fuzzy logic. Fuzzy logic is based on the notion of fuzzy sets. In classical set theory a set has crisp boundaries, i.e. an object is either a member of a set A or it is member of the set ¬A (the complement of A). In fuzzy set theory sets have fuzzy boundaries. This notion is closely associated with prototype theory. A set has prototypical members that best describe the set. Then there are things that closely resemble the prototype, but are not considered to meet the properties of the prototype completely. A 7 feet tall person is prototypical for a tall person, while a 5 foot 8 inches tall person is less prototypical, but can still be considered as somewhat tall. In fuzzy logic, how much or little prototypical an object is, is identified by the set-membership function, which maps an attribute value to the range [0..1], which is the degree of membership. These set membership function can potentially take arbitrary forms, for example linear or Gaussian. In the special case of a step function the fuzzy set takes the form of a classical set.

Umberto Straccia [34] has developed a fuzzy description logic that can easily be expressed in the form of OWL statements. SWRL [18] can then be used to map values to degrees of membership in fuzzy sets.

3.2. Probabilities and Possibilities

Probabilistic reasoning gives us a way of finding out how likely something is the case. If 75% of the students in a class passed the exam, then we have a 75% chance to randomly pick one that passed. However, it doesn't make a claim about every particular student. Each single one either passed or failed. There is a fundamental difference in the semantics of fuzzy logic and probabilistic logic. In fuzzy logic, a statement can be true to a certain extent or an entity belongs to a class to a certain degree. This degree is assumed to be known with certainty. In probabilistic reasoning, there is a probability that a statement is true or false, but the statement itself is either true or false, but neither both nor something in between. Hence fuzzy logic sees the world as continuous instead of binary, while probabilistic logics make a claim about the randomness of the world or the observer's state of certainty.

For the task of knowledge acquisition, the statistical analysis of data allows the exploration of relationships that are not explicitly stated. Statistical techniques give us great insight into a corpus of documents or a large collection of data in general, when a program exists that can actually "pose the right questions to the data", i.e. analyze the data according to our needs. All derived relationships are statistical in nature and we only have an idea or a likelihood of their validity.

In the traditional bivalent-logic based formalisms we, that is the users or the systems, have to make a decision. Once two contradictory statements are identified, one has to be chosen as the right one. While this is possible in domains that are axiomatized, fully explored or in which statements are true by definition, it is not possible for most scientific domains. In the life sciences for instance, hypotheses have to be evaluated, contradicting statements have promoting data, etc. Decisions have to be deferred until enough data is available that either verifies or falsifies the hypothesis. Nevertheless, it is desirable to express these hypotheses formally to have means to computationally evaluate them on the one hand and to exchange them between different systems on the other.

3.3. Belief and trust

As the number of ontologies available grows and maybe in the future also their interoperability, we will be facing the problem of inference across ontologies. We will have to take into account that ontologies evolve and that facts change. In the paradigm of FOL and FOL-based formalisms, this can completely throw off the inference engine. It is very likely that a set of related ontologies is inconsistent. Assume a good-case scenario in which the interoperability issue is solved insofar as a common thesaurus has been used to name concepts in a set of ontologies and the merging of the ontologies has been accomplished already. Imagine further a set of news related ontologies O. And for simplicity, assume that all ontologies follow a bivalent paradigm and that each ontology by itself is consistent. It is very likely that some subset $O' \subset O$ entails a fact a, while some other subset $O'' \subset O$, $O'' \cap O' = \emptyset$, $O'' \cup O' = O$, entails $\neg$a. Each entailment is logically valid and sound. But the merged set is inconsistent and will lead to false inferences. A simple modal solution is to put a big epistemological **B** in front of each ontology o_i to show that “it is believed by the creator(s) of o_i, that the facts in o_i are true”.

With such a framework for distributed belief, the problem of inconsistency in merged ontologies is somewhat solved, but on the other hand, it does not make us any wiser. There exists no calculus that, based on these beliefs, tells us, what is more likely to be a true belief.

The first possibility is to assign a belief value to every belief. Each knowledge bearing agent, i.e. each ontology and its creator(s), can assign confidence values to their beliefs [5]. A calculus allows the evaluation of the overall distributed belief. This, however, requires epistemic honesty on the side of each agent and, in a distributed environment with unreliable sources, may result in wrong conclusions, since “The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people so full of doubts” (Bertrand Russell).

Even though it might not be epistemologically correct, we are making the simplified assumption that the complement to belief on the side of the knowledge providing agent is trust on the side of the knowledge consuming agent. Trust, as much as belief, is a state of mind and hence subjective and biased. However, the consumer is in charge of it. As opposed to relying on the opposite side’s honesty, the consumer can assign a trust value to sets of concepts. Each of these sets might encompass a complete ontology, a subset of the ontology or single classes or facts. The consumer can now assign trust values of 0.8 to the New York times and the Economist, 0.6 to CNN and BBC, 0.01 to the Enquirer[2]. The same calculus as used for belief in [5] can now be applied for trust. Combining both is also possible and is in the hands of the consumer.

The statement that trust is purely subjective can be relaxed at this point. The statements of other sources can be compared to a trusted source of choice and the trust value can be a function of the jointly held beliefs vs. contradicting beliefs.

4. A framework for logical extensions on the Semantic Web

This section will outline the parts of a possible framework for several multi-valued logics. We will present one example of a combined fuzzy and probabilistic logic and a partial translation into RDF as an extension to OWL.

4.1. A fuzzy uncertain logic

Cao and Rossiter [9] describe a fuzzy and probabilistic language for database systems, implemented in the language FRIL++ [3]. In this language, a property is of the form **φ[l,u]**, where φ is a fuzzy atom and **[l,u]** describes the probability range that the proposition of the atom is the case. A class *Dinner* could have a fuzzy probabilistic attribute *time_of_day(evening)[0.8, 0.95]*, indicating, that there is an 80% to 95% chance that dinners are held in the evening. The label *evening* of course identifies a fuzzy set, indicating the times of day roughly between late afternoon and nightfall.

[2] One might actually imagine a negative assignment such as in “I trust with certainty 0.5 that everything what the government says is false”. Consider as well, that is it not the same as claiming that the opposite is true or believed.

For practical purposes, it is assumed, that a class totally subsumes its subclasses, even though fuzzy logic would allow an overlap. Hence the knowledge engineer needs to be careful when designing a taxonomy of fuzzy concepts.

4.2. OWL-FP: Combining OWL and a fuzzy-probabilistic logic

From a theoretical point of view it is easy to conceive of a language just like OWL that has a meaningful continuous representation for all sorts of relationships. The beauty of it is, that it can handle standard bivalent facts as a special case in which the class membership value *1.0* and the probability range is *[1, 1]*. Hence, a reasoner for a fuzzy and probabilistic logic can translate an OWL property into a fuzzy-probabilistic-OWL property. We will exemplify this on material implication:

$$\frac{\mathbf{a} \supset \mathbf{b} \qquad \mathbf{a}}{\mathbf{b}} \quad \equiv \quad \frac{\mathbf{Pr(b \mid a) = 1\ ;\ \ Pr(b \mid \neg\, a) = [0, 1]} \qquad \mathbf{Pr(a) = 1}}{\mathbf{Pr(b) = Pr(b \mid a)Pr(a) + Pr(b \mid \neg\, a)Pr(\neg\, a) = 1}}$$

The reverse translation is much harder, as it is always more difficult to map a description in a more powerful language to a less powerful one. If the ontology is meant to be translated, the burden lies on the ontology designer who has to be careful to design the ontology in such a way that a less powerful inference engine than the one the ontology was designed for can still use the contained information. For example, when an ontology that has been designed for a reasoner that processes a fuzzy and probabilistic logic is given to a traditional OWL reasoner, it is desirable that the OWL part of the ontology still be consistent. There are different ways to solve this problem; a simple solution is to just disregard the non-OWL constructs, similar to an HTML parser that just skips non-HTML tags. If a fuzzy assignment made the instance a fuzzy member of two classes that are according to the OWL representation supposed to be disjoint, then it can be made a member of the lowest common super class in the class hierarchy. An ontology designer can make sure that this happens by having this class be the primary class of the instance. For example

```
<X:instance_1>
   <fuzzyOWL:type rdf:resource= “X.1” membership_value = “0.4”/>
   <fuzzyOWL:type rdf:resource= “X.2” membership_value = “0.6”/>
</X:instance_1>
```

In this case, *instance_1* is still an instance of X. The assignment of *instance_1* with a membership degree to the subsets can either be discarded or treated as a simple named relationship within OWL.

Given the power that a language like Fril++ offers for fuzzy and probabilistic reasoning, it is easy to conceive of an extension to the syntax of OWL that allows the representation of fuzzy and probabilistic constraints:

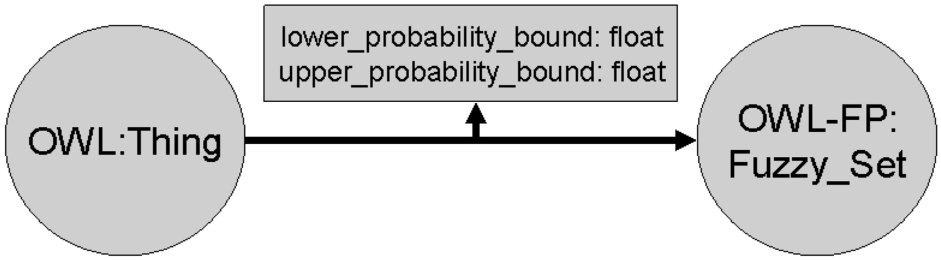

Figure 3: Example of a fuzzy and probabilistic attribute

This is an example of a fuzzy and probabilistic attribute. It describes the classes OWL-FP:Property, OWL-FP:Fuzzy_Set as well as the slots OWL-FP:fuzzy_value, OWL-FP:upper_probability_bound and OWL-FP:lower_probability_bound, which, together with the domain of Property, describe **φ, l** and **u.** Since the second order attributes of the relationship are limited, it is still computable. The RDF syntax for it is:

```
<rdfs:Class rdf:ID="Fuzzy-Set">
    <rdfs:label>Datatype</rdfs:label>
    <rdfs:comment>
        The class of all fuzzy sets
    </rdfs:comment>
    <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdfs:Class>

<rdf:Property rdf:ID="upper_probability_bound">
    <rdfs:label>upper probability bound</rdfs:label>
    <rdfs:comment>
        upper probability bound, u in [l, u]
    </rdfs:comment>
    <rdfs:domain rdf:resource="#Restriction"/>
    <rdfs:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#float"/>
</rdf:Property>

<rdf:Property rdf:ID="lower_probability_bound">
    <rdfs:label>lower probability bound</rdfs:label>
    <rdfs:comment>
        lower probability bound, l in [l, u]
    </rdfs:comment>
    <rdfs:domain rdf:resource="#Restriction"/>
    <rdfs:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#float"/>
</rdf:Property>

<rdf:Property rdf:ID="has_fuzzy_value">
    <rdfs:label>has fuzzy value</rdfs:label>
    <rdfs:comment>
        for a fuzzy set F and (a, b) in Property P, b is in F
    </rdfs:comment>
```

```
    <rdfs:domain rdf:resource="#Restriction"/>
    <rdfs:range rdf:resource="#Fuzzy-Set"/>
</rdf:Property>
```

Because of the ugliness of all XML, we forego a further description of the OWL-FP syntax. This example is just meant to show, that it is possible to easily define a language more powerful than OWL, if the underlying semantics are already well defined. Definition is one step. In order to have any sort of semantics, the notion of the fuzzy set and the probability bounds have to be computationally used, of course.

5. Soundness and Completeness

Soundness and completeness are seen as crucial properties of a logical inference. Soundness refers to the fact that given true premises, the inferred conclusions are always true within the system. Completeness means that every conclusion that is true in the system can be derived. For many applications however, the property of completeness of the inference mechanism will be secondary to its soundness. Hence the first step will have to be the proof of the soundness of different combinations of logics.

The relaxation of the requirement of completeness also gives rise to the design of more efficient algorithms. Domain independent and domain specific heuristics can be applied to the inference task. Some domain independent heuristics could be "follow the sequences with the higher likelihoods first", "consider the most prototypical members of fuzzy sets first" or the converse rules.

The language proposed by Cao and Rossiter [9] was proven to be sound but not complete. Since for now, we have based our formalism on this language, reasoning with OWL-FP will also be sound, but not complete. Soundness always means soundness with respect to the semantics of the language, not with respect to the "real world". We have seen above that non-monotonic logics need to do global inference in order to give complete results. This, in our case, is equivalent to finding the tightest probability bounds. However, the derived probability bounds are not inconsistent with the tightest ones. This allows us to restrict the reasoning to local inference instead of global. The danger in not finding the tightest intervals is, though, that we might find many meaningless statements, such as $P(a) = [0, 1]$, which means that the probability of a is somewhere between zero and one This is apparently true, but does not convey much information. It would be an interesting exercise to see whether, if the knowledge base grows indefinitely, all inference would result in meaningless statements.

Relaxing the completeness requirement also raises the question of relaxing the notion of soundness. While it is absolutely necessary to have a sound inference leveling a deductive proof, many tasks cannot be sound by definition. Any sort of inductive or abductive reasoning faces the danger of producing incorrect results. Here, the focus is merely to have a statistical bound on the percentage of errors. Any tool that is meant to assist with the derivation of new knowledge needs human evaluation. It can merely suggest the results of its reasoning and add a confidence measure. In essence, it again depends on the area of application. The difference to traditional models of inductive and abductive reasoning is, however, that the Semantic Web technologies are used for the exchange of data amongst agents and the reasoning across local domains or knowledge bases. A general purpose reasoner need not consider the abductive and inductive cases.

6. Fundamental limitations of the current standard

While it is one problem to clearly define these different modal and uncertain properties, the challenge is to define their actual function by defining operations on these properties. With the latest versions of SWRL [18], this is possible in the context of the Semantic Web layers themselves. And even though this is not the most efficient way of computing, the ontologies for uncertain reasoning can carry their own rules and functions. In the case of OWL-FP, the use of FRIL++ improves performance.

While for every local use, any sort of extension to OWL can be used to formalize and a proprietary inference mechanism can be used to meaningfully interpret the data, in a global scenario these extensions become meaningless. With "Contextualizing Ontologies" and the C-OWL formalism, [6] point in the right direction. The C-OWL language allows assigning certain parts of remote ontologies as consistent or equivalent, other parts as inconsistent. This allows us to use parts of remote ontologies and to just leave out the inconsistent statements. However, a lot of information that could have been gained from the remote ontologies is lost, because the approach is too rigid. One can picture overlapping concepts A and B, such that $|A \cap B| > 0$, $|A \cup B| < |A| + |B|$, that in this case, since neither $A \supseteq B$, nor $A \subseteq B$, nor $A \equiv B$, the concepts would be inconsistent. For knowledge discovery it is desirable to know about the concept overlap. $A \approx B$ to a degree λ.

In a recent Weblog [41], basically the same question was raised that we raise here. The author's conclusion as to why we need a representation of uncertainty is also very much along our lines: "*Fundamentally because the semantic web is about finding connections.*" And most likely we want to find connections across ontologies that have not been found before and declared to be consistent. OWL and bivalent DLs in general lack the basic capabilities of expressing degrees of consistency in a meaningful way.

7. Related Work

Even though, with the exception of a probabilistic extension to OWL [10], no XML-style serialization of a more powerful formalism than that of OWL has been published, AI, KR and IR research has brought forward many formalisms that extend FOL in one or the other direction. The extremely powerful Knowledge Interchange Format (KIF) language [43] is an example for a higher order language that is actually used in ontology creation, but due to its complexity on the one hand and its requirement of locality of the knowledge on the other hand not considered for the semantic web stack. It is known that increasing the expressive power of a KR language causes problems relating to computability and decidability. This has been the main reason for limiting the expressive power of KR languages. The real power behind human reasoning however is the ability to do so in the face of imprecision, uncertainty, inconsistencies, partial truth, and approximation. There have been attempts made in the past at building KR languages that allow such expressive power.

As previously outlined, major approaches to reasoning with imprecision are 1) probabilistic reasoning, 2) possibilistic reasoning [11] and 3) fuzzy reasoning. In [38], Lotfi Zadeh proposed a formalism that combines fuzzy logic with probabilistic reasoning to exploit the merits of both approaches. Zadeh stresses the different semantics of fuzzy sets. For background on this, see also Dubois and Prade [12]. With the *Fuzzy Logic for the Internet (FLINT)* initiative, the BISC group presents fuzzy logic as a part of a larger framework including the notions of precisiated natural language (PNL), fuzzy knowledge bases for background world knowledge, etc. as a basis for future question answering systems. [27] present a survey of applications of fuzzy logic in contemporary search engines and challenges for future applications of fuzzy logic in perception based querying and question answering.

Other formalisms have focused on resolving local inconsistencies in knowledge bases; see for instance the works by Blair, Kifer, Lukasiewicz, Subrahmanian and others in annotated logic and paraconsistent logic [22; 4]. In [26], Thomas Lukasiewicz proposes a weak probabilistic logic and addresses the problem of inheritance. In [8], T.H. Cao proposed an annotated fuzzy logic approach that is able to handle inconsistencies and imprecision.

The easiest transformations to a Semantic Web language can be done from a Description Logic. Many formalisms have been proposed that immediately extend description logics with uncertain or fuzzy reasoning support. Umberto Straccia has done extensive work on Fuzzy Description Logics, see for example [34] or the excellent survey of existing approaches in [35]. With P-CLASSIC, Daphne Koller and others presented an early approach to probabilistic description logics implemented in Bayesian Networks [24]. Other probabilistic description logics have been proposed by Heinsohn [16] and Jaeger [19]. Early research on Bayesian-style inference on OWL was done by Zhongli Ding [10]. In her formalism, OWL is augmented to represent prior probabilities. However, the problem of inconsistencies arising through inheritance of probability values [26] is not taken into account. An interesting probabilistic extension to the Description Logic SHOQ(D), which is very close to the OWL-DL standard, is described in Guigno and Lukasiewicz [14]. The Authors can prove soundness and completeness. An assessment of the computational complexity of P-SHOQ(D) is proposed for future research. A different turn was taken by Wang et al [36], who provided a non-monotonic extension to the SHOQ(D) Description Logic to be able to process incomplete knowledge on the Semantic Web.

In our previous discussion in [33], we give examples of the applications of soft computing techniques in classical database, IR and Internet tasks. We claim that only with the processing of different types of semantics, the Semantic Web vision can be accomplished. It is concluded, that: "[…] the focus of effort in pursuit of the Semantic Web vision needs to move towards an approach that encompasses all three types of semantics both in representation of the knowledge and the computational methods that support the creation of such knowledge." The three types of semantics considered are semantics derived using statistical inference (implicit), semantics based on formal bivalent logics (formal) and formal soft representations (powerful).

In a very comprehensive essay, Wheeler and Pereira [37] analyze the epistemological problems that were only glanced at in this survey and relate them to non-monotonic and default-reasoning. While we have grossly simplified the semantics of fuzzy and probabilistic logic in our approach of a fuzzy and probabilistic OWL, the authors investigate different semantics of the expression of uncertainty, with regards to truth and belief. Dubois and Prade [13] also stress the different semantics that such modal operators can have and the resulting differences for their computation.

Other researchers look at approaches that are not based on logic. Hence the criteria of rigid soundness and completeness do not apply. The paradigm is rather that the computer should assist the human investigator to arrive at conclusions by showing results that are relevant to previously defined formulas and constraints. Golbeck and Hendler [15], true to the slogan "A little semantics goes a long way", coined by James Hendler, define a model of trust that is logically less rigid. An agent assigns trust values to other agents. In a network of agents, it can be inferred by a transitivity function, how much the first agent would trust an agent that he does not immediately know, but that is a few hops down the network path..

Anyanwu et al [1, 2] present an algorithm to find and rank complex relationships in RDF graphs. Any sound deduction in a logic paradigm can be seen as a path in a graph. In addition to the valid deduction, there are most likely several paths in the graph that would produce invalid deductions. According to a classical paradigm, this is, because parts of the paths just do not entail the rest of the path. But this can also arise due to inconsistencies between two RDF schemas. A path finder does not have to worry about these inconsistencies and can simply follow all paths between two entities. The presented ranking mechanism then decides which paths are more likely to be informative and which are less. In this order the paths are then presented to the user. This, like many graph search algorithms, does not really take edge semantics into account that go beyond specificity and other purely structural semantics. However, what at first sight seems to be a drawback proves to be extremely useful, because it allows fast access to the relevant part of the knowledge base and hence allows us to narrow down the number of statements we have to consider in a rigid logical proof.

8. Conclusions

As so often when it comes to practical applications of great ideas, we find ourselves in a conundrum between possibility and usability. Many of the great technologies that could make everything so easy and perfect cannot feasibly be applied. On the other hand, not taking at least the ideas behind those technologies into account will cause the Semantic Web formalisms to be too weak to actually do what we envision it to do.

In our work we are looking towards combining faster, but incomplete graph-based algorithms with more rigid logic formalisms. For future work we want to use the former to identify important regions of a graph that potentially represents multiple ontologies to narrow down the number of statements we have to do logical inference on. We believe that combining different kinds of reasoning on different kinds of semantics is crucial to make reasoning on the Semantic Web a successful and computationally feasible endeavor [33].

Undoubtedly, nothing can solve all our problems. Not even a powerful formalism for the representation of logic on the Semantic Web. But it is not supposed to. The secret is, as so often, the application of the right tools at the right time and the right place. Whether we need a fuzzy description of our concepts depends on the domain and the degree of accuracy we need in modeling it. Whether we need a probabilistic description or not depends on our certainty and on the randomness or indeterminacy of the domain. As it happens, it all depends on the scope of our application. When two companies want to use ontologies to speed up their legal contracts, it is mandatory that no fuzziness or uncertainties remain. In this case the user wants to be aware of every inconsistency so that s/he can assess the legal implications.

For the task of finding knowledge across different ontologies, it is necessary to allow for the resolution of inconsistencies and partial belief or knowledge. The question whether the representation of this is internal to the knowledge-acquiring agent or external and explicitly available is crucial. If the acquired knowledge is to be used instantaneously and then thrown away, it is up to each agent to deal with it. If it is used to formalize new knowledge that is available to users and artificial agents on the web, it better reflects the degree of uncertainty under which it was gained and the sources from which it was obtained. Otherwise, errors in knowledge representation can propagate through the web. In a situation in which knowledge is still mostly curated by humans, this is a minor problem. In a scenario in which thousands of ontologies are available and used mainly by machines, the restriction to purely bivalent non-monotonic logics will become a major obstacle for knowledge retrieval and inference.

9. Acknowledgements

This work is funded primarily by Bioinformatics of Glycan Expression, a component of the "Integrated Technology Resource for Biomedical Glycomics," funded by the National Center for Research Resources, July 1, 2003 – June 30, 2008. We gratefully acknowledge the collaboration with Prof. William S. York in the BioChemistry domain which is the basis of one of the examples discussed in this chapter.

10. References

[1] Kemafor Anyanwu, Amit Sheth, *The p Operator: Discovering and Ranking Associations on the Semantic Web*. SIGMOD Record 31(4): 42-47 (2002)

[2] Kemafor Anyanwu, Angela Maduko, Amit Sheth, *SemRank: Ranking Complex Relationship Search Results on the Semantic Web*, The 14th International World Wide Web Conference, (WWW2005), Chiba, Japan, May 10-14, 2005

[3] James F. Baldwin, Tru.H. Cao, Trevor P. Martin, Jonathan M. Rossiter, *Towards soft computing object-oriented logic programming*, in Proc. 9th IEEE Internat. Conf. on Fuzzy Systems, San Antonio, USA, 2000, pp. 768–773.

[4] Howard .A. Blair, V.S. Subrahmanian, *Paraconsistent logic programming*, Theoretical Comp. Sci. 68, 1989, 135-154

[5] Luca Boldrin and Alessandro Saffiotti, *A Modal Logic for Merging Partial Belief of Multiple Reasoners*, Journal of Logic and Computation 9(1):81-103, 1999

[6] Paolo Bouquet, Fausto Giunchiglia, Frank van Harmelen, Luciano Serafini, Heiner Stuckenschmidt, *Contextualizing Ontologies*, Journal of Web Semantics 1(4): 325-343 (2004)

[7] Sergey Brin and Lawrence Page, *The anatomy of a large-scale hypertextual web search engine*, In Proceedings of the Seventh International World Wide Web Conference, 1998.

[8] Tru H.Cao, *Annotated Fuzzy Logic Programs*, International Journal on Fuzzy Sets and Systems 113 (2000) 277-298

[9] Tru H. Cao, Jonathan M. Rossiter, *A deductive probabilistic and fuzzy object-oriented database language*, Fuzzy Sets and Systems 140 (2003) 129–150

[10] Zhongli Ding, Yun Peng, *A Probabilistic Extension to Ontology Language OWL*, Proceedings of the Hawaii International Conference on System Sciences, January 5-8, 2004, Big Island, Hawaii.

[11] Didier Dubois, Jérôme Lang, Henri Prade, *Possibilistic Logic*, In Dov M. Gabbay et al. (Eds), Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3, Oxford University Press, Oxford, 1994, pp 439 - 514

[12] Didier Dubois and Henri Prade, *The three semantics of fuzzy sets*, Fuzzy Sets and Systems 90 (1997) 141-150.

[13] Didier Dubois and Henri Prade, *Possibility theory, probability theory and multiple-valued logics: A clarification*, Annals of Mathematics and Artificial Intelligence 32: 35–66, 2001.

[14] Rosalba Giugno, Thomas Lukasiewicz, *P-SHOQ(D): A Probabilistic Extension of SHOQ(D) for Probabilistic Ontologies in the Semantic Web*, Proceedings of the European Conference on Logics in Artificial Intelligence, 2002

[15] Jennifer Golbeck, James Hendler, *Accuracy of Metrics for Inferring Trust and Reputation*, in Proceedings of 14th International Conference on Knowledge Engineering and Knowledge Management, October 5-8, 2004, Northamptonshire, UK

[16] Jochen Heinsohn, *Probabilistic Description Logics*, UAI 1994: 311-318.

[17] Ian Horrocks, Peter F. Patel-Schneider and Frank van Harmelen, *From SHIQ and RDF to OWL: the making of a Web Ontology Language*, Journal of Web Semantics 1(1): 7-26 (2003)

[18] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean, *SWRL: A Semantic Web Rule Language Combining OWL and RuleML*, W3C Member Submission 21 May 2004 http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/

[19] Manfred Jaeger, *Probabilistic Reasoning in Terminological Logics*, KR 1994: 305-316.

[20] Vipul Kashyap, Cartic Ramakrishnan, Christopher Thomas and Amit Sheth, *TaxaMiner: An Experimental Framework for Automated Taxonomy Bootstrapping*, to appear in the International Journal of Web and Grid Services, Special Issue on Semantic Web and Mining Reasoning, September 2005

[21] Vipul Kashyap, Amit Sheth, *Semantic Heterogeneity in Global Information Systems: The Role of Metadata, Context and Ontologies*, Cooperative Information Systems 1996

[22] Michael Kifer and V.S. Subrahmanian, *Theory of generalized annotated logic programming and its applications*, Journal of Logic Programming, 12:335-367, 1992

[23] Won Kim, Injun Choi, Sunit K. Gala, and Mark Scheevel, *On Resolving Schematic Heterogeneity in Multidatabase Systems*, Distributed and Parallel Databases, 1(3):251--279, 1993.

[24] Daphne Koller, Alon Levy, and Avi Pfeffer, *P-CLASSIC: A tractable probabilistic description logic*, Proc. Of the 14th Nat. Conf. on Artificial Intelligence (AAAI-97), 390-397, 1997

[25] Ravi Krishnamurthy, Witold Litwin, and William Kent , *Language features for interoperability of databases with schematic discrepancies*, in J. Clifford and R. King, editors, Proceedings of ACM SIGMOD Conference, pages 40--49. ACM, May1991.

[26] Thomas Lukasiewicz, *Weak Nonmonotonic probabilistic Logics*, Principles of Knowledge Representation and Reasoning (KR) 2004

[27] Masoud Nikravesh, Vincenzo Loia, and Behnam Azvine, *Fuzzy logic and the Internet (FLINT): Internet, World Wide Web, and search engines*, Soft Computing - A Fusion of Foundations, Methodologies and Applications, Volume 6, Issue 5, Aug 2002, Pages 287 - 299

[28] Borys Omelayenko, *Learning of Ontologies for the Web: the Analysis of Existent approaches*, In Proceedings of the International Workshop on Web Dynamics, 2001.

[29] Ganesh Ramakrishnan, Soumen Chakrabarti, Deepa Paranjpe, Pushpak Bhattacharya, *Is question answering and acquired skill?* Proceedings of the 13th international conference on World Wide Web 2004

[30] Stephen Reed and Douglas B. Lenat, *Mapping Ontologies into Cyc*, In AAAI 2002 Conference Workshop on Ontologies For The Semantic Web, Edmonton, Canada, July 2002.

[31] Jonathan Rossiter, *Fusing Partially Inconsistent Expert and Learnt Knowledge in Uncertain Hierarchies*, Proceedings of the Third International Conference on Intelligent Data Engineering and Automated Learning, 2002

[32] Amit Sheth, *Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics*, in Interoperating Geographic Information Systems. M. F. Goodchild, M. J. Egenhofer, R. Fegeas, and C. A. Kottman (eds.), Kluwer, Academic Publishers, 1999, pp. 5-30.

[33] Amit Sheth, Cartic Ramakrishnan, and Christopher Thomas, *Semantics for the Semantic Web: the Implicit, the Formal and the Powerful*, International Journal on Semantic Web & Information Systems, 1 (1), Jan-Mar 2005, pp. 1-18

[34] Umberto Straccia, *A Fuzzy Description Logic*, Proceedings of AAAI-98, 15th Conference of the American Association for Artificial Intelligence

[35] Umberto Straccia, *Uncertainty and Description Logic Programs: A Proposal for Expressing Rules and Uncertainty on Top of Ontologies*, Technical Report 2004-TR-14

[36] Songxin Wang, Shuigeng Zhou, Fei Wang, and Aoying Zhou, *N-SHOQ(D): A Nonmonotonic Extension of Description Logic SHOQ(D)*, in J.X. Yu, X. Lin, H. Lu, and Y. Zhang (Eds.): APWeb 2004, LNCS 3007, pp. 912–915, 2004.

[37] Gregory Wheeler, Luís Moniz Pereira. *Epistemology and Logical Artificial Intelligence*, Journal of Applied Logic, 2(4): 469-493, 2004

[38] Lotfi A. Zadeh, *Toward a perception-based theory of probabilistic reasoning with imprecise probabilities*, In Journal of Statistical Planning and Inference 105 (2002) 233-264.

[39] Bioinformatics for Glycan Expression. Integrated Technology Resource for Biomedical Glycomics: Technological Research and Development Project IV, A Project funded by National Center for Research Resources http://lsdis.cs.uga.edu/projects/Glycomics/

[40] http://platypuswiki.sourceforge.net/, http://en.wikipedia.org/wiki/Semantic_Wikipedia

[41] SWAD-Europe Weblog 05/30/04: The semantic web and uncertainty: http://esw.w3.org/mt/esw/archives/000055.html

[42] http://www.vivisimo.com

[43] http://logic.stanford.edu/kif/dpans.html