Uncertainty and RuleML Rulebases: A Preliminary Report

Giorgos Stoilos[1], Giorgos Stamou[1], Vassilis Tzouvaras[1], and Jeff Z. Pan[2]

[1] Department of Electrical and Computer Engineering, National Technical University of Athens, Zographou 15780, Greece
[2] School of Computer Science, The University of Manchester
Manchester, M13 9PL, UK

Abstract. Uncertainty, like imprecision and vagueness, has gained considerable attention the last decade. To this extend we present a preliminary report on extending the Rule Markup Language (RuleML) with fuzzy set theory, in order to be able to represent and handle vague knowledge. We also provide semantics for the case of fuzzy FOL RuleML.

1 Introduction

According to widely known proposals for a Semantic Web architecture, ontologies will play a key role in the Semantic Web This has led to considerable efforts to developing a suitable ontology language, culminating in the design of the OWL Web Ontology Language [1], which is now a W3C recommendation. Although OWL adds considerable expressive power with respect to languages such as RDF, it does have expressive limitations, particularly with respect to what can be said about properties, as well as the total absence of rules, which are valuable in many real life applications. To this end the extension of the current semantic web architecture with some form of rules language has lead to several proposals, like SWRL [2], FOL RuleML [3], and many more.

Even though the combination of OWL and rules results in the creation of a highly expressive language, there are still many occasions where this language fails to accurately represent knowledge of our world. In particular these languages fail at representing vague and imprecise knowledge and information. Uncertainty, is both a characteristic of information itself, like the concepts of a “tall” preson, a “happy” person, a “fast” car, a “hot” place, a “faulty” part, and many more. Experience with these domains has shown that in many cases dealing with such type of information yields more realistic applications, which derive better results. The need for covering uncertainty in the Semantic Web context has been stressed out in literature many times the last years; some examples are [4–6].

In order to capture imprecision in rules, we propose a fuzzy extension of the RuleML framework, called f-RuleML. In f-RuleML, facts about the world can include a specification of the “degree” (a truth value between 0 and 1) of confidence with which one can assert that a tuple of individuals is an instance of a given relation.

A. Adi et al. (Eds.): RuleML 2005, LNCS 3791, pp. 199–203, 2005.

2 Preliminaries

Fuzzy set theory and fuzzy logic constitute a widely used framework for the representation and management of various forms of uncertainty, like vagueness and imprecision, introduced in real-life applications. They are based on the notion of fuzzy sets, introduced in [7]. While in terms of classical set theory any element belongs or not to a set, in fuzzy set theory this is a matter of degree. More formally, let X be a collection of elements with cardinality m, i.e $X = \{x_1, x_2, \ldots, x_m\}$. A crisp subset S of X is any collection of elements of X that can be defined with the aid of its *characteristic function* $\chi_A(x)$ that assigns any $x \in X$ to a value 1 or 0 if this element belongs to X or not, respectively. On the other hand, a fuzzy subset A of X, is defined by a membership function $\mu_A(x)$, or simply $A(x)$, $x \in X$. This membership function assigns any $x \in X$ to a value between 0 and 1 that represents the degree in which this element belongs to X.

The classical set theoretic operation of complement, union, intersection and the logical operation of implication are also extended to this new framework and are performed by special mathematical functions over the unit interval called fuzzy complement (c), fuzzy intersection or t-norm (t or $*$), fuzzy union or t-conorm (u) and fuzzy implication ($\Rightarrow$), respectively [8]. Among these operations fuzzy implications play an important role as they determine the properties of the resulting fuzzy logic [9]. In our context we will consider only the class of R-implications [8], cause of their interesting properties. These implications are given by the equation: $x \Rightarrow y = \sup\{z \in [0,1] \mid t(x,z) \leq y\}$, and they have the nice property that $x \Rightarrow y = 1$ if and only if $x \leq y$.

3 Dealing with Uncertainty in RuleML

In the current section we will use a motivating use case to present the syntactic changes that need to be applied to the RuleML framework.

Consider a casting company, which has a knowledge base that consists of models. Advertisement companies are using this knowledge base to look for models to be used in tv commercials. Each entry in the knowledge base contains a photo of the model, personal information and some body and face characteristics. The casting company has created a user interface for inserting the information of the models as instances of a predefined ontology. It also provides a query engine to search for models with specific characteristics, which in the case of advertisement companies usually are complex characteristics, like the hair quality, color, body fitness, skin quality, etc, in order to determine if they qualify for a certain commercial. Obviously, such a knowledge base contains a wealth of vague concepts, like long hair, brown eyes, athletic body, and many more. In such a case one might want to specify the membership degree of an individual, say "SUSAN" to a fuzzy concept like $brown\_eyes$, as $brown\_eyes(SUSAN) \geq 0.7$, to indicate the least degree that "SUSAN" participates to the fuzzy concept $brown\_eyes$. We can make these assertions explicit to the system by encoding them as RuleML *fuzzy facts*. In that case we can write,

```
<Atom>
  <degree><Data>0.8</Data></degree>
  <_opr><Rel>brown_eyes</Rel></_opr>
  <Ind>SUSAN</Ind>
</Atom>
```

From the above examples we can see that the syntactic changes that need to take place are minimal and only involve the syntax of fuzzy facts. So the additional change that needs to take place in the XML Schema definition of the element `Atom` [10] is the following:

```
<xs:group name="Atom.content">
...................
  <xs:choice>
    <xs:sequence>
      <xs:element name="degree" type="degree.type" minOccurs="0"
maxOccurs="1"/>
    <xs:choice>
      <xs:element ref="opr"/>
....................
</xs:group>
```

The XML Schema definition of the new tag `degree` could be given by the following schema definition:

```
<xs:attributeGroup name="degree.attlist"/>
  <xs:group name="degree.content">
    <xs:sequence>
      <xs:element ref="Data"/>
    </xs:sequence>
  </xs:group>
  <xs:complexType name="degree.type">
    <xs:group ref="degree.content"/>
    <xs:attributeGroup ref="degree.attlist"/>
  </xs:complexType>
<xs:element name="degree" type="degree.type"/>
```

Subsequently, one can use such fuzzy facts to encode knowledge in the form of fuzzy rules without any further syntax modification.

4 Fuzzy FOL RuleML

In the current section we will present the syntax and semantics of Fuzzy FOL RuleML (f-FOL RuleML). Our presentation follows the one in [11].

Definition 1. *[11] A* predicate language *consists of a non-empty set of* predicates*, each together witha positive natural number (the arity), and a (possibly empty) set of* object constants*. Predicates are mostly denoted by* $P, Q, R, ...$, *constants by* $c, d, ...$. Logical Symbols *are* object variables $x, y, ...$, connectives $\&, \rightarrow$, truth constants $\bar{r}$ *for each rational* $r \in [0, 1]$ *and* quantifiers $\exists, \forall$. *Other connectives are defined as follows:*

$$\begin{array}{ll} \phi \wedge \psi = \psi \& (\phi \rightarrow \psi), & \phi \equiv \psi = (\phi \rightarrow \psi) \& (\psi \rightarrow \phi) \\ \neg \phi = \phi \rightarrow \bar{0}, & \phi \vee \psi = ((\phi \rightarrow \psi) \rightarrow \psi) \wedge ((\psi \rightarrow \phi) \rightarrow \phi) \end{array}$$

Atomic formulas *have the form* $P(t_1, ..., t_n)$*, where* P *is a predicate of arity* n *and* $t_1, ..., t_n$ *are terms. If* ϕ, ψ *are formulas and* x *is an object variable then* $\phi \rightarrow \psi, \phi \& \psi, (\exists x)\phi, (\forall x)\phi, \bar{r}$ *are formulas.*

Definition 2. *[11] Let* $\mathcal{J}$ *be a predicate language and let* $\boldsymbol{L}$ *be a linearly ordered BL-algebra. An* $\boldsymbol{L}$-structure $\boldsymbol{M} = \langle M, (r_P)_P, (m_c)_c \rangle$ *for* $\mathcal{J}$ *has a non-empty* domain M*, for each* n*-ary predicate* P *a* $\boldsymbol{L}$*-fuzzy* n*-ary relation* $r_P : M^n \rightarrow \boldsymbol{L}$ *on* M*, associating to each* n*-tuple of elements of* M *the degree* $r_P(m_1, ..., m_n) \in \boldsymbol{L}$ *of the membership of* $(m_1, ..., m_n)$ *to the fuzzy relation, and for each object constant* c*,* m_c *is an element of* M*.*

Definition 3. *[11] Let* $\mathcal{J}$ *be a predicate language and* $\boldsymbol{M}$ *an* $\boldsymbol{L}$*-structure for* $\mathcal{J}$*. An* $\boldsymbol{M}$-evaluation *of object variables is a mapping* u *assigning to each object variable* x *an elements* $u(x) \in M$*. Let* u, u' *be two evaluations.* $u \equiv_x u'$ *means that* $u(y) = u'(y)$ *for each variable* y *distinct from* x*. The value of a term given by* $\boldsymbol{M}$*,* u *is defined as follows:* $\|x\|_{\boldsymbol{M},u} = u(x)$*,* $\|c\|_{\boldsymbol{M},u} = m_c$*. We define the* truth value $\|\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u}$ *of a formula. as follows:*

$$\begin{array}{rlrl} \|P(t_1, ..., t_n)\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= r_P(\|t_1\|_{\boldsymbol{M},u}, ..., \|t_n\|_{\boldsymbol{M},u}), & \|\bar{r}\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= r \\ \|\phi \& \psi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= \|\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} * \|\psi\|^{\boldsymbol{L}}_{\boldsymbol{M},u}, & \|\phi \rightarrow \psi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= \|\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} \Rightarrow \|\psi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} \\ \|(\exists x)\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= \sup\{\|\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u'} | u \equiv_x u'\} & \|(\forall x)\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u} &= \inf\{\|\phi\|^{\boldsymbol{L}}_{\boldsymbol{M},u'} | u \equiv_x u'\} \end{array}$$

At last a rule is interpreted as: $\|\phi \rightarrow \psi\|_{\boldsymbol{M},u} = 1$

Observe that since fuzzy rules are in general not equivalent fuzzy implications, as it is also the case in classical rules, we have interpreted rules as fuzzy implications which are 1-tautologies, i.e. their truth value is 1, in each safe **L**-structure **M** and each **M**-valuation of object variables. Since in our case we use R-implications the above equation further means that $\|\phi\|^{\mathbf{L}}_{\mathbf{M},u} \leq \|\psi\|^{\mathbf{L}}_{\mathbf{M},u}$.

Logics like the one presented above is often referred to as *Rational Pavelka predicate logic* ($RPL\forall$). The reader is referred to [11] for more information on these logics.

5 Discussion

Recently, a lot of interest towards dealing with uncertainty, like imprecision and vagueness (fuzzyness), in the semantic web context has been demonstrated. It is well established in the AI community that dealing with such type of information yields more intelligent and realistic applications. Since rules will also play an important role in the realization of the semantic web and its wider acceptance by the industry community, the need for dealing with uncertainty in rule systems is evident. The idea of adding fuzzyness in logic programs is not new. Several approaches to fuzzy logic programming have been presented [12–15]. Recently the interest has been extended to the semantic web where both fuzzy description logics [16] and fuzzy rules, are integrated providing a fuzzy extension to the SWRL language [17]. In the current paper we have extended the RuleML framework in order to make it capable to represent and handle imprecise and vague information. We showed that the syntactic changes that need to take place are minimal and only regard facts.

References

1. Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D.L., Patel-Schneider, P.F., eds., L.A.S.: OWL Web Ontology Language Reference (2004)
2. Horrocks, I., Patel-Schneider, P.F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A Semantic Web Rule Language — Combining OWL and RuleML. W3C Member Submission, `http://www.w3.org/Submission/SWRL/` (2004)
3. Boley, H., Dean, M., Grosof, B., Sintek, M., Spencer, B., Tabet, S., Wagner, G.: FOL RuleML: The First-Order Logic Web Language. W3C Member Submission, `http://www.w3.org/Submission/FOL-RuleML/` (2005)
4. Bechhofer, S., Goble, C.: Description Logics and Multimedia - Applying Lessons Learnt from the GALEN Project. In: KRIMS 96 Workshop on Knowledge Representation for Interactive Multimedia Systems, ECAI 96, Budapest (1996)
5. Stoutenburg, S., Obrst, L., Nichols, D., Peterson, J., Johnson, A.: Toward a standard rule language for semantic integration in the dod enterprise, W3C Workshop on Rule Languages for Interoperability (2005)
6. Chen, H., Fellah, S., Bishr, Y.: Rules for geospatial semantic web applications, W3C Workshop on Rule Languages for Interoperability (2005)
7. Zadeh, L.A.: Fuzzy sets. Information and Control **8** (1965) 338–353
8. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall (1995)
9. Klement, E., Navara, M.: A survey of different triangular norm-based fuzzy logics. Fuzzy Sets and Systems **101** (1999) 241–251
10. Hirtle, D., Boley, H., Grosof, B., Kifer, M., Sintek, M., Tabet, S., Wagner, G.: Schema Specification of RuleML 0.89. `http://www.ruleml.org/0.89/` (2005)
11. Hajek, P.: Metamathematics of fuzzy logic. Kluwer (1998)
12. Vojtás, P.: Fuzzy logic programming. Fuzzy Sets and Systems **124** (2001) 361–370
13. Ebrahim, R.: Fuzzy logic programming. Fuzzy Sets and Systems **117** (2001) 215–230
14. Cao, T.: Annotated fuzzy logic programs. Fuzzy Sets and Systems **113** (2000) 277–298
15. Damásio, C.V., Pereira, L.M.: Monotonic and residuated logic programs. In: Proceedings of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Springer-Verlag (2001) 748–759
16. Stoilos, G., Stamou, G., Tzouvaras, V., Pan, J., Horrocks, I.: A fuzzy description logic for multimedia knowledge representation, Proc. of the International Workshop on Multimedia and the Semantic Web (2005)
17. Pan, J.Z., Stamou, G., Tzouvaras, V., Horrocks, I.: f-SWRL: A Fuzzy Extension of SWRL. In: Proc. of the International Conference on Artificial Neural Networks, Special section on “Intelligent multimedia and semantics”. (2005) To appear.