

# $\mathcal{EL}$ description logic modeling querying web and learning imperfect user preferences

Peter Vojtáš

Dpt. Software Eng., Fac. Math. Phys., Charles University Prague  
Inst. Computer Sci., Czech Acad. Sci, Prague

**Abstract.** In this position paper we share ideas on modeling querying web resources by (imperfect) combination of particular user preferences based on description logic. Our basic assumption is, that web resources are modeled crisp. Imperfection (uncertainty, vagueness,...) comes from user context and preferences. We offer a model based on connection between three  $\mathcal{EL}$ -description logic systems: classical, annotated(fuzzy) and a new variant of Bayesian description logic. The Bayesian part enables learning each single user's combination function and concepts.

In [2] R. Fagin et al. propose heuristics for a middleware for finding best (top-k) answer when the data that we wish to access and combine may reside in a variety of web repositories. We can look to the Fagin's model as working with fuzzy RDF data (ordered by user preference of particular attribute) and using a combination function providing global score for ordering of resources. Our main goal is to provide description logic counterpart for Fagin's model.

Typical example is a user looking for a hotel for *vacation* which is *cheap* and *close* to a beach. Particular user preferences (wrt a fixed ordinal preference scale  $\mathcal{T}$ ) need to be combined to get an overall ordering of results. Our model should be also able, for another user, to deliver e.g. hotels with medium price and distance.

In [9] we have proposed f- $\mathcal{EL}^{\text{®}}$  description logic allowing existential restrictions, crisp roles,  $\mathcal{T}$ -fuzzy concepts and  $\mathcal{T}$ -fuzzy aggregation functions. Our system is less expressive than fuzzy description logic of U. Straccia [7] but more effective and embeddable into RDF (and OWL) without reification.

The user's aggregation function is learned from his/her global preferences by an Inductive logic programming system ([3]), combining  $\mathcal{T}$ -degree of being *cheap* and *close* to a global  $\mathcal{T}$ -degree of being good for *vacation*). In practical situations a problem can occur: when user evaluates a sample of resources by a global preference, few inconsistencies in evaluation can lead to a constant aggregation function as an only possible, strictly logical, solution. To overcome this, we need to extend our model with a Bayesian part.

For this, we can transform our previous work on imperfection in logic programming [8] to description logic framework. In [8] we have shown that generalized annotated programs (GAP) of M. Kifer and V.S. Subrahmanian [5] are

---

<sup>1</sup> Email: peter.vojtas@mff.cuni.cz, supported by Czech IT project 1ET100300517, 1ET100300419 and MSM-0021620838

equivalent with an extension of classical logic programs with monotonicity axioms added, and these are (weakly) embeddable into Bayesian logic program of K. Kersting and L. De Raedt [4]. Experiments were provided with Bayesian networks BN (*cheap*, *close* and *vacation* are random variables over  $\mathcal{T}$ ).

In this position paper we share ideas on using connections between classical  $\mathcal{EL}$  description logic,  $f\text{-}\mathcal{EL}^{\circledast}$  description logic and a new variant of Bayesian description logic B- $\mathcal{EL}$  for modeling user preference querying web resources.

Several models of Bayesian description logic were described already in [1]. In [6] Bayesian description logic programs are described. All mentioned systems use a random distribution on the domain  $\Delta^{\mathcal{I}}$ , or a measure on the power set of  $\Delta^{\mathcal{I}}$  and provide a challenge of inducing probability distribution over web resources, properties and values.

Our system does not model randomness over web data. Uncertainty and vagueness comes with user and his/her context, interpretation of web data and preferences. Resource data (roles *price*, *distance*) remain crisp, randomness is touching only user query concepts (like a hotel for *vacation*) and user preference concepts (like *cheap*, *close*) which are random variable over the ordinal preference scale  $\mathcal{T}$ . A cpd can describe user preference depending on distribution of  $\exists price.cheap$  and  $\exists distance.close$ .

We did not provide any experiments with this model. Our experiments in the logic programming framework give a well supported expectation that this could be a technically sound candidate for a simple, sufficiently expressive model for description of user preference querying of web resources. This leads to a layered model - web resources are crisp and each single user specific part with imperfection and learning ability.

## References

1. F. Baader, R. Kuesters, F. Wolter. Extensions to Description Logic, Handbook DL, 2003, 219-261
2. R. Fagin, A. Lotem and M. Naor. Optimal aggregation algorithms for middleware, J. Computer and System Sciences 66 (2003), pp. 614–656
3. T. Horváth, P. Vojtáš. Ordinal Classification with Monotonicity Constraints, In Proc. ICDM2006, LNAI 4065, Springer-Verlag 2006, 217 - 225
4. K. Kersting, L. De Raedt. Bayesian Logic Programs, Technical Report 151, University of Freiburg, 52 pages
5. M. Kifer, V. S. Subrahmanian. "Theory of generalized annotated logic programming and its applications", J. Logic Programming, 12 (1992) pp 335–367
6. L. Predoiu, Information Integration with Bayesian Description Logic Programs, WWW2006, May 2326, 2006, Edinburgh, UK
7. U. Straccia. A fuzzy description logic for the Semantic web, In FLSW, Elsevier (2006) 73–90
8. P. Vojtáš, M. Vomlelová. "On models of comparison of multiple monotone classifications", In Proc. IPMU'2006, Éditions EDK, Paris, pp. 1236-1243
9. P. Vojtáš. A description logic with combination of user preference concepts, crisp roles and existential restrictions, In EJC 2006, pp. 176-183, 2006