Extending PSL with Fuzzy Quantifiers

Golnoosh Farnadi[1,2], Stephen H. Bach[3], Marie-Francine Moens[2], Lise Getoor[4], Martine De Cock[1,5]
[1]Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium,
[2]Department of Computer Science, Katholieke Universiteit Leuven, Belgium,
[3]Statistical Relational Learning Group, University of Maryland, USA,
[4]University of California, Santa Cruz, USA,
[5]Center for Data Science, University of Washington Tacoma, USA

Introduction

Many interesting tasks in artificial intelligence require the ability to work with *imperfect relational* data. Examples include social network applications in which partial information about users and about their connections is available. For instance, we may have *attribute* information about users; we may know the age, gender, personality, likes and dislikes etc. of a user, but this knowledge is usually incomplete in the sense that we do not have all attribute values for all users. The knowledge can also be uncertain, for instance when the attribute values are not given directly by the user but instead they are inferred from user generated content.

The *relational* aspect of social network data stems from the connections between the users, e.g., Facebook users are connected with their friends; Twitter users can follow one another; in Amazon, users can bookmark other interesting users; in Epinions users can trust other users or include them in their block list (i.e. distrust) etc. Again, the available knowledge about the connections is typically incomplete and uncertain. An interesting task is to perform most probable explanation (MPE) inference over a given social network, i.e. based on (1) given attribute values about users and their connections and (2) some background knowledge about the domain, expressed as probabilistic rules in first-order logic, infer missing attribute values such that the given and the inferred values combined adhere to the background knowledge as well as possible. MPE inference is an important task in *statistical relational learning* (SRL).

Probabilistic Soft Logic (PSL)

A particularly interesting framework for SRL is *probabilistic soft logic* (PSL). Unlike in other forms of SRL, in which predicates are Boolean, in PSL they can take soft truth values in the interval $[0, 1]$. Conjunction and disjunction in PSL rules are respectively modelled by the Łukasiewicz t-norm and t-conorm because their particular piecewise linear form allows to cast MPE inference for a PSL program as a convex optimization problem. Negation is the standard negation from fuzzy logic, i.e. for x and y in $[0, 1]$ (the ˜ indicates the relaxation over Boolean values):

$$x \tilde{\wedge} y = \max(0, x + y - 1), x \tilde{\vee} y = \min(x + y, 1)$$

$$\tilde{\neg} x = 1 - x$$

In a PSL program, relationships and attributes are modelled by predicates, and first-order rules model constraints on these predicates. MPE inference is concerned with finding the most probable assignment of truth values. The probability of truth value assignments is defined by a weighting function, namely, distance to satisfaction (d_r). Generally, the distance to satisfaction of an unsatisfied rule is the difference between the truth values of the body and the head. A rule r is satisfied when the truth value of its head is at least as high as the truth value of its body. Therefore, the rule's distance to satisfaction under an interpretation I is defined in Equation (1). An interpretation I is a mapping that associates a truth value to each element, i.e., $I(x) \in [0, 1]$.

$$d_r(I) = \max\{0, I(r_{body}) - I(r_{head})\} \quad (1)$$

A PSL program, i.e. a set of PSL rules, induces a distribution over interpretations I. The probability density function is (Brocheler, Mihalkova, and Getoor 2012):

$$f(I) = \frac{1}{Z} \exp[-\sum_{r \in R} \lambda_r (d_r(I))^p] \quad (2)$$

where λ_r is the weight of the rule r, Z is a normalization constant (see Equation (3)), and $p \in \{1, 2\}$ provides a choice of two different loss functions. $p = 1$ favors the satisfaction of one rule, and $p = 2$ favors the satisfaction of all rules to some degree. These probabilistic models are instances of hinge-loss Markov random fields (Bach et al. 2013).

$$Z = \int_I \exp[-\sum_{r \in R} \lambda_r (d_r(I))^p] \quad (3)$$

For example *Trusts*$(A, B) \rightarrow$ *Knows*(A, B) models that "if A trusts B then A knows B" where A and B are variables referring to arbitrary objects (Huang et al. 2012). An example of a grounded version of this PSL rule is *Trusts(Alice,Bob)* $\rightarrow$ *Knows(Alice,Bob)*. If *Trusts(Alice,Bob)* and *Knows(Alice,Bob)* are true to respectively degree 0.7 and 0.5, then the rule is satisfied to degree $1 - (0.7 + 0.5) = 0.8$. The truth degree of some predicates is given in advance, and, roughly speaking, the goal of MPE inference is to find a truth assignment for the other predicates such that the combined degree of satisfaction of all grounded rules is as high as possible.

Variables in PSL rules are implicitly universally quantified. A PSL rule *Trusts*$(A, X) \wedge$ *Trusts*$(X, B) \rightarrow$ *Trusts*(A, B), models that "A trusts B" is true to the degree to which there is a trusted third party X. In standard PSL there is no direct way to express that A can trust B if *most* friends of A trust B. The notion *most* in this statement represents a fuzzy quantifier. Atoms whose truth values are functions of other atoms' truth values are defined using *aggregates*. PSL supports linear aggregates, e.g., the average truth value of a set of atoms. However, non-linear aggregates such as the fuzzy quantifiers as we propose in the next section require additional capabilities.

Fuzzy quantifiers

The idea of fuzzy quantifiers (Delgado et al. 2013) was first introduced by Zadeh (Zadeh 1983). Quantifiers represent notions such as "most" and "a few". The necessity of introducing a new notation to define quantifiers based on a linguistic perspective has also been studied in (Barwise and Cooper 1981; Keenan and Westerstahl 2011) in the context of generalization quantifiers.

Intuitively, quantifiers relate to the concept of cardinality of sets. Recall that a fuzzy set A in a universe X is a $X - [0, 1]$ mapping. The cardinality of A is defined as the sum of the individual membership degrees, i.e.

$$|A| = \sum_{x \in X} A(x) \tag{4}$$

A fuzzy quantifier Q is a $[0, 1] - [0, 1]$ mapping. Q is called coherent if Q is non-decreasing and if it satisfies the boundary conditions $Q(0) = 0$ and $Q(1) = 1$. Figures 1 and 2 depict two coherent fuzzy quantifiers.

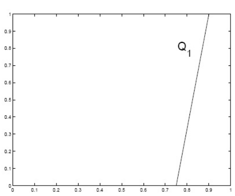

Figure 1: Fuzzy quantifier Q_1

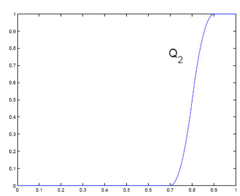

Figure 2: Fuzzy quantifier Q_1

Zadeh (Zadeh 1983) suggested to calculate the truth value of "Q A's are Bs", with A and B fuzzy sets in X, as:

$$Q\left(\frac{|A \cap B|}{|A|}\right) \tag{5}$$

Other approaches for fuzzy quantifiers and aggregation operators have been proposed, most notably Yager's OWA-operators (Yager 1988).

Extending the semantics of PSL to allow the use of these quantifiersis non-trivial. The main challenge is finding a way in which MPE inference over PSL programs can be cast as a convex optimization problem.

Extension approach

A PSL model consists of a collection of weighted rules. Given the data, rules are grounded out with substitution of logical terms, and form hinge-loss potentials for Equation (2). Hinge-loss MRFs have log concave density functions and finding a MPE is a convex optimization problem, which is solvable in polynomial time. Although, the numerator of the fuzzy quantifier (i.e., Equation (5)) is grounded out using the Łukasiewicz t-norm, grounding the function to some threshold is not linear and cannot be represented with linear constraints.

However, one could initialize the aggregate values by grounding the quantifier operator. Each fuzzy aggregate can be defined as a new predicate, e.g., most of the trusted friends can be defined by $MostFriendsTrust(A, B)$ which calculates the cardinality of the trusted friends of A who trust B over all trusted friends of A. By using this new predicate, we are able to define rules similar to the standard PSL rules using conjunction and disjunction, e.g., the following PSL rule indicates that if most friends of A trust B, A trusts B: $MostFriendsTrust(A, B) \rightarrow Trusts(A, B)$

The value of each grounded predicate will be updated to satisfy all grounded rules during the MPE inference or learning phase which maximizes the likelihood of each variable conditioned on all other variables. Thus, to update the value of the fuzzy aggregates, we propose to ground the aggregate function iteratively during the inference and weight learning phase until the aggregate value has converged.

Conclusion

PSL is a probabilistic modeling framework which uses first-order logic and soft truth values in the interval $[0, 1]$ for reasoning in relational domains. PSL uses the Łukasiewicz t-norm and t-conorm from fuzzy logic to model respectively conjunction and disjunction. In the current version of PSL, rules representing notions such as "most" and "a few" have not been addressed. We propose an extension to PSL with fuzzy quantifiers to overcome this limitation. We believe fuzzy quantifiers expand the ability of PSL to model social network applications which work with *imperfect relational data*, e.g., trust propagation, link prediction, node labelling.

Acknowledgements: This project is part of the SBO Program of the IWT-SBO-Nr. 110067 and NSF CCF0937094.

References

Bach, S. H.; Huang, B.; London, B.; and Getoor, L. 2013. Hinge-loss Markov random fields: Convex inference for structured prediction. In *Uncertainty in Artificial Intelligence (UAI).*

Barwise, J., and Cooper, R. 1981. Generalized quantifiers and natural language. *Linguistics and philosophy* 4(2):159–219.

Brocheler, M.; Mihalkova, L.; and Getoor, L. 2012. Probabilistic similarity logic. *arXiv preprint arXiv:1203.3469*.

Delgado, M.; Ruiz, M. D.; Sánchez, D.; and Vila, M. A. 2013. Fuzzy quantification: a state of the art. *Fuzzy Sets and Systems* 242:1–30.

Huang, B.; Kimmig, A.; Getoor, L.; and Golbeck, J. 2012. Probabilistic soft logic for trust analysis in social networks. In *International Workshop on Statistical Relational AI*, 1–8.

Keenan, D., and Westerstahl, D. 2011. Generalized quantifiers in linguistics and logic. *van Benthem and A. ter Meulen (eds.), Handbook of Logic and Language* 837–893.

Yager, R. R. 1988. On ordered weighted averaging aggregation operators in multicriteria decision making. *Systems, Man and Cybernetics, IEEE Transactions on* 18(1):183–190.

Zadeh, L. A. 1983. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with Applications* 9(1):149–184.