Conjunctive Query Answering with Finitely Many Truth Degrees*

Stefan Borgwardt[1], Theofilos Mailis[2],
Rafael Peñaloza[3]**, and Anni-Yasmin Turhan[4]

[1] Chair for Automata Theory, Theoretical Computer Science, TU Dresden, Germany
[2] Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece
[3] KRDB Research Centre, Free University of Bozen-Bolzano, Italy
[4] Department of Computer Science, University of Oxford, UK

1 Introduction

Fuzzy description logics (FDLs) have arisen as suitable formalisms for representing and reasoning with the vague or imprecise knowledge that is intrinsic to many application domains. They extend classical description logics by allowing additional truth degrees that lie between the classical "true" and "false" values. These truth degrees typically belong to a subset of the interval $[0, 1]$.

For example, in a cloud computing environment, one might be interested in modeling the notion of an *overused* component. This is a typical example of an imprecise concept, since it is impossible to give a precise point where a component starts being overused. Instead, in FDLs, all components are assigned the degree to which they are being overused, where a higher degree implies a more extensive usage. For example, an idle component is overused with degree 0, while a component running at half its capacity might be overused to degree 0.8. The axioms

$$\langle \mathsf{Overused}(\mathsf{cpuA}) \geq 0.8 \rangle,$$
$$\langle \mathsf{Server} \sqcap \exists \mathsf{hasPart}.\mathsf{Overused} \sqsubseteq \mathsf{ServerWithLimitedResources} \geq 0.9 \rangle$$

express that object cpuA is overused to a degree of at least 0.8, and that every server that has an overused part is a server with limited resources with a degree of at least 0.9, respectively. The different concept constructors are interpreted by a t-norm and its associated operators [13]. One important t-norm is the Łukasiewicz t-norm, which is defined by $x \otimes y := \max\{x + y - 1, 0\}$.

Since dealing with infinitely many truth degrees easily leads to undecidability of reasoning [1, 7, 12], we focus on finitely valued FDLs, where the degrees are

* This work was partially supported by DFG under grant BA 1122/17-1 'FuzzyDL' (S. Borgwardt), the European project Optique (T. Mailis), the Cluster of Excellence 'cfAED' (R. Peñaloza), and EPSRC grant EP/J0083-46/1 'PrOQAW: Probabilistic Ontological Query Answering on the Web' (A.-Y. Turhan).

** The work was developed while the author was still affiliated with TU Dresden and the Center for Advancing Electronics Dresden, Germany.

ordered in a finite chain. In this case, standard reasoning in expressive FDLs has been shown to be decidable, and in the same complexity class, as reasoning in their classical counterparts [9–11]. One proposed method for reasoning in finitely valued FDLs is based on *crispification*. The idea of this method is to transform the fuzzy ontology into a classical ontology that preserves all the information about the truth degrees expressed in the original ontology. This is achieved through new concept and role names like $\mathsf{Fast}_{\geq 0.8}$ that intuitively contain all the elements that belong to Fast to a degree of at least 0.8. In this way, one can reduce reasoning in fuzzy DLs to reasoning in classical DLs, for which highly optimized reasoners exist. However, this approach only works for DLs that include at least the expressivity of $\mathcal{ALCH}$.

2 Types of Fuzzy Queries

A reasoning problem extensively studied for DLs over the last years is (conjunctive) query answering, together with the associated query entailment problem. Briefly, a conjunctive query q is a finite set of concept and role atoms, which intuitively are ABox assertions that might contain variables in place of individuals. An ontology $\mathcal{O}$ *entails* the query q if every model $\mathcal{I}$ of $\mathcal{O}$ has a match for q; that is, if all the variables in q can be mapped to elements of the domain of $\mathcal{I}$ in a way that all the atoms are satisfied. In FDLs, the matches of an atom need not be absolute, but might also hold with a truth degree between 0 and 1.

The existence of intermediate truth degrees gives rise to two different notions of conjunctive queries that can be entailed by a fuzzy ontology. The first one, called *threshold conjunctive query*, extends the notion of an atom to express additionally the least degree to which the atom must be satisfied in each model. Thus, for example, we can ask whether $\mathsf{server1}$ is fast (to degree 0.8) and has an overused (to degree 0.6) component through the threshold query

$$\{\mathsf{Fast}(\mathsf{server1}) \geq 0.8, \quad \mathsf{hasPart}(\mathsf{server1}, x) \geq 1, \quad \mathsf{Overused}(x) \geq 0.6\}. \tag{1}$$

Such a query is entailed by $\mathcal{O}$ if every model of $\mathcal{O}$ has a match with at least the given degrees. Notice that the result of a threshold query entailment check is either "yes" (if the query is entailed by $\mathcal{O}$) or "no." There are no intermediate degrees associated with these answers.

The second type of query, called *fuzzy conjunctive query*, asks for the *best entailment degree*; i.e., the largest possible degree d such that every model of the ontology has a match to degree at least d. For example, using the fuzzy conjunctive query

$$\{\mathsf{Fast}(\mathsf{server1}), \quad \mathsf{hasPart}(\mathsf{server1}, x), \quad \mathsf{Overused}(x)\}, \tag{2}$$

we can find the best degree to which $\mathsf{server1}$ is fast and has an overused component, where the conjunction between the atoms is interpreted using a t-norm. In the case of fuzzy conjunctive queries, there is only one degree that is global for the whole match of the query. Thus, it is possible that the threshold query above is not entailed (i.e., answers "no") while this fuzzy conjunctive query returns a positive degree.

3 Results

We propose a query answering procedure based on the crispification approach. In addition to crispifying the ontology, we also translate a threshold query into a classical conjunctive query that preserves the semantics w.r.t. the crispified ontology. For example, the threshold query (1) is crispified into the classical conjunctive query

$$\{\mathsf{Fast}_{\geq 0.8}(\mathsf{server1}), \quad \mathsf{hasPart}_{\geq 1}(\mathsf{server1}, x), \quad \mathsf{Overused}_{\geq 0.6}(x)\}.$$

Recall that $\mathsf{Fast}_{\geq 0.8}$ is a classical concept name from the crispified ontology. Thus, deciding entailment of this conjunctive query suffices for deciding entailment of the original threshold query.

A similar translation is used for fuzzy conjunctive queries, except that the result is a union of conjunctive queries, where each conjunctive query considers a certain combination of individual degrees whose combination leads to the entailment of the fuzzy query. For example, to decide whether the fuzzy conjunctive query (2) is entailed to a degree of at least 0.8, say in the presence of the degree set $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$, one has to consider a union of several CQs such as

$$\{\mathsf{Fast}_{\geq 0.8}(\mathsf{server1}), \quad \mathsf{hasPart}_{\geq 1}(\mathsf{server1}, x), \quad \mathsf{Overused}_{\geq 1}(x)\} \text{ and}$$
$$\{\mathsf{Fast}_{\geq 1}(\mathsf{server1}), \quad \mathsf{hasPart}_{\geq 1}(\mathsf{server1}, x), \quad \mathsf{Overused}_{\geq 0.8}(x)\},$$

where the t-norm of the individual degrees is equal to 0.8. After the translation, one can use any existing query answering system for classical DLs.

While studying the crispification approach for expressive FDLs, we encountered two issues. First, we noticed that some of the previous crispification approaches, such as those in [4, 5], do not treat number restrictions correctly when the Łukasiewicz t-norm is used. Second, the previously known crispifications (see also [2, 6]) produce an exponential blow-up, which makes almost any instance of the problem infeasible. We solved the second issue by introducing a linear normalization step that ensures a polynomial bound on the size of the crispification. Essentially, the normalization process introduces abbreviations that avoid copying complex concepts during the crispification step.

Using this normalization step, we are able to prove tight complexity bounds for answering threshold conjunctive queries; the complexity is always the same as for classical conjunctive query answering (in DLs more expressive than $\mathcal{ALCH}$ that do not have number restrictions). Unfortunately, the translation of fuzzy conjunctive queries causes an exponential blow-up, which is avoided when the simple Gödel t-norm $x \otimes y := \min\{x, y\}$ is used. Moreover, the *data complexity* of classical query answering in DLs is not affected when considering finitely valued semantics, as the reduction of the ABox (the data) is linear. Finally, our method for fuzzy query answering can be applied to any crispification approach, i.e. also in the cases that correctly handle number restrictions [3].

More details can be found in [8], which has been submitted to a journal. A preliminary version of these results appeared in [14].

References

1. Baader, F., Penaloza, R.: Are fuzzy description logics with general concept inclusion axioms decidable? In: Fuzzy Systems (FUZZ), 2011 IEEE International Conference on. pp. 1735–1742. IEEE (2011)
2. Bobillo, F., Delgado, M., Gómez-Romero, J.: Crisp representations and reasoning for fuzzy ontologies. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 17(4), 501–530 (2009)
3. Bobillo, F., Delgado, M., Gómez-Romero, J., Straccia, U.: Fuzzy Description Logics under Gödel semantics. International Journal of Approximate Reasoning 50(3), 494–514 (2009)
4. Bobillo, F., Delgado, M., Gómez-Romero, J., Straccia, U.: Joining Gödel and Zadeh fuzzy logics in fuzzy description logics. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 20(4), 475–508 (2012)
5. Bobillo, F., Straccia, U.: Reasoning with the finitely many-valued Łukasiewicz fuzzy Description Logic $\mathcal{SROIQ}$. Information Sciences 181(4), 758–778 (2011)
6. Bobillo, F., Straccia, U.: Finite fuzzy Description Logics and crisp representations. In: Uncertainty Reasoning for the Semantic Web II, pp. 99–118. Springer (2013)
7. Borgwardt, S., Distel, F., Peñaloza, R.: The limits of decidability in fuzzy description logics with general concept inclusions. Artificial Intelligence 218, 23–55 (2015)
8. Borgwardt, S., Mailis, T., Peñaloza, R., Turhan, A.Y.: Answering fuzzy conjunctive queries over finitely valued fuzzy ontologies (2015), under submission to a journal.
9. Borgwardt, S., Peñaloza, R.: The complexity of lattice-based fuzzy description logics. Journal on Data Semantics 2(1), 1–19 (2013)
10. Borgwardt, S., Peñaloza, R.: Consistency reasoning in lattice-based fuzzy description logics. International Journal of Approximate Reasoning 55(9), 1917–1938 (2014)
11. Bou, F., Cerami, M., Esteva, F.: Finite-valued Lukasiewicz modal logic is PSPACE-complete. In: Walsh, T. (ed.) Proc. of the 22nd Int. Joint Conf. on Artificial Intelligence (IJCAI'11). pp. 774–779. AAAI Press (2011)
12. Cerami, M., Straccia, U.: On the (un)decidability of fuzzy Description Logics under Łukasiewicz t-norm. Information Sciences 227, 1–21 (2013)
13. Hájek, P.: Metamathematics of Fuzzy Logic (Trends in Logic). Springer-Verlag (2001)
14. Mailis, T., Peñaloza, R., Turhan, A.Y.: Conjunctive query answering in finitely-valued fuzzy description logics. In: Kontchakov, R., Mugnier, M.L. (eds.) Proceedings of the 8th International Conference on Web Reasoning and Rule Systems (RR 2014). vol. 8741, pp. 124–139. Springer (2014)