Automata-based Reasoning in Fuzzy Description Logics

Rafael Peñaloza

Theoretical Computer Science, TU Dresden, Germany
Center for Advancing Electronics Dresden
penaloza@tcs.inf.tu-dresden.de

Description logics (DLs) are a family of well-studied knowledge representation formalisms designed to express and reason with the conceptual knowledge of application domains in a clear and well-understood manner. They have been successfully applied for representing large application domains, most prominently from the biological and medical fields. In their classical form, DLs are not adequate for handling vague or imprecise knowledge, which is a common staple in bio-medical knowlege. To alleviate this problem, fuzzy extensions of DLs have been introduced. As a prototypical example, we consider here the smallest propositionally closed fuzzy DL, which we call $\otimes$-$\mathcal{ALC}$.[1]

The fuzzy DL $\otimes$-$\mathcal{ALC}$ is based on *concepts* and *roles*, which are interpreted as (fuzzy) unary and binary relations, respectively. Given the disjoint sets $\mathsf{N_R}$, and $\mathsf{N_C}$ of *role*, and *concept names*, respectively, $\otimes$-$\mathcal{ALC}$ *concepts* are built through the grammar rule

$$C ::= A \mid \bot \mid C \sqcap C \mid C \to C \mid \exists r.C \mid \forall r.C,$$

where $A \in \mathsf{N_C}$ and $r \in \mathsf{N_R}$. The concept $\top$ is often used as an abbreviation of $\bot \to \bot$. The terminological knowledge of a domain is represented through a *TBox*: a finite set of *general concept inclusions* (GCIs) of the form $\langle C \sqsubseteq D \geq q\rangle$, where C, D are $\otimes$-$\mathcal{ALC}$-concepts, and $q \in [0, 1]$.

The semantics of this logic is given through *interpretations*, which are pairs $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}}$ is a non-empty set called the *domain*, and $\cdot^{\mathcal{I}}$ is a function that maps every $A \in \mathsf{N_C}$ to a function $A^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \to [0, 1]$, and every $r \in \mathsf{N_R}$ to a function $r^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \to [0, 1]$. Intuitively, for every domain element $x \in \Delta^{\mathcal{I}}$ the value $A^{\mathcal{I}}(x)$ represents the degree to which x is a member of A. The interpretation function is extended to arbitrary concepts using the continuous t-norm $\otimes$ and its (unique) residuum $\Rightarrow$. In the case of G-$\mathcal{ALC}$, where the semantics is based on the Gödel t-norm, complex concepts are interpreted as shown in Table 1.

The interpretation $\mathcal{I}$ is a *model* of the TBox $\mathcal{T}$ if for every GCI of the form $\langle C \sqsubseteq D \geq q\rangle \in \mathcal{T}$ and every $x \in \Delta^{\mathcal{I}}$, $C^{\mathcal{I}}(x) \Rightarrow D^{\mathcal{I}}(x) \geq q$ holds. Reasoning tasks in fuzzy DLs are based on the class of models of a TBox. However, it is customary to further restrict this class allow only so-called *witnessed* models, where the suprema and infima stated by the semantics of the existential and value restrictions, respectively, are in fact maxima and minima. We keep this

[1] Unfortunately, there is no agreed naming for fuzzy DLs. We use this name to emphasize the relationship with $\mathcal{ALC}$, the smallest propositionally closed classical DL.

Table 1: Semantics of $\mathsf{G}\text{-}\mathcal{ALC}$

constructor	**syntax**	**semantics**
bottom concept	$\bot$	0
conjunction	$C \sqcap D$	$\min(C^{\mathcal{I}}(x), D^{\mathcal{I}}(x))$
implication	$C \to D$	$C^{\mathcal{I}}(x) \Rightarrow D^{\mathcal{I}}(x)$
existential restriction	$\exists r.C$	$\sup_{y\in\Delta^{\mathcal{I}}} \min(r^{\mathcal{I}}(x,y), C^{\mathcal{I}}(y))$
value restriction	$\forall r.C$	$\inf_{y\in\Delta^{\mathcal{I}}} r^{\mathcal{I}}(x,y) \Rightarrow C^{\mathcal{I}}(y)$

restriction, and for the rest of this paper call *witnessed models* simply *models* for brevity.

Most reasoning tasks in fuzzy DLs can be reduced to deciding the existence of a model that satisfies an additional set of restrictions, or *restricted consistency*. A *restriction* is an expression of the form $\langle C \triangleright q\rangle$, where C is a concept, $q \in [0,1]$, and $\triangleright \in \{\leq, \geq\}$. A finite set of restrictions $\mathcal{R}$ is *consistent* w.r.t. the TBox $\mathcal{T}$ if there is a model $\mathcal{I}$ of $\mathcal{T}$ and an element $x \in \Delta^{\mathcal{I}}$ such that $C^{\mathcal{I}}(x) \triangleright q$ holds for every restriction $\langle C \triangleright q\rangle \in \mathcal{R}$.

Restricted consistency and other associated reasoning tasks have been recently shown to be hard (even undecidable) for non-idempotent t-norms; i.e., any continuous t-norm that is not Gödel [5]. One culprit for this hardness is the fact that, for those t-norms, $\otimes\text{-}\mathcal{ALC}$ does not have the finite model property nor the *finitely-valued model property*. That is, there exist consistent restrictions that are only satisfied by infinite models that use infinitely many different membership degrees [3]. This fact is used to prove that the existence of such a model cannot be decided in finite time. Given the simplicity of the operators associated to the Gödel t-norm, it was generally believed that $\mathsf{G}\text{-}\mathcal{ALC}$ has the finite model property. Moreover, it is often claimed that all reasoning tasks in this logic can be restricted to only a finite set of truth degrees, which can be computed *a priori*, depending only on the values explicitly provided in the input. This belief seems to arise from the results in [6] which, however, depend on different semantics.

Consider the set of restrictions $\mathcal{R} = \{\langle A \leq 0.6\rangle\}$ and the TBox

$$\mathcal{T} = \{\langle \forall r.A \sqsubseteq A \geq 1\rangle\,,\ \langle \exists r.\top \sqsubseteq A \geq 1\rangle\}.$$

It is easy to see that $\mathcal{R}$ is consistent w.r.t. $\mathcal{T}$. For any model $\mathcal{I}$ of $\mathcal{T}$ that satisfies $\mathcal{R}$ there must exist an element $x_1 \in \Delta^{\mathcal{I}}$ such that $A^{\mathcal{I}}(x_1) < 0.6$. As $\mathcal{I}$ is witnessed, there exists a $x_2 \in \Delta^{\mathcal{I}}$ with $(\forall r.A)^{\mathcal{I}}(x_1) = r^{\mathcal{I}}(x_1,x_2) \Rightarrow A^{\mathcal{I}}(x_2)$. The first axiom of $\mathcal{T}$ entails $r^{\mathcal{I}}(x_1,x_2) \Rightarrow A^{\mathcal{I}}(x_2) \leq A^{\mathcal{I}}(x_1) < 1$, and in particular $r^{\mathcal{I}}(x_1,x_2) > A^{\mathcal{I}}(x_2)$. The second axiom of the TBox $\mathcal{T}$ implies that

$$r^{\mathcal{I}}(x_1,x_2) = \min(r^{\mathcal{I}}(x_1,x_2),1) \leq (\exists r.\top)^{\mathcal{I}}(x_1) \leq A^{\mathcal{I}}(x_1),$$

and thus $A^{\mathcal{I}}(x_1) > A^{\mathcal{I}}(x_2)$. Repeating the same argument, there must exist elements $x_3, x_4, \ldots \in \Delta^{\mathcal{I}}$ such that $A^{\mathcal{I}}(x_i) > A^{\mathcal{I}}(x_{i+1})$ for all $i \geq 1$. This means

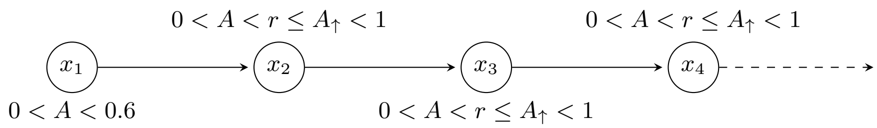

Fig. 1: An abstract description of models of $\mathcal{T}$ satisfying $\mathcal{R}$

that any model of $\mathcal{T}$ satisfying the restriction $\mathcal{R}$ must have infinitely many elements that belong to the concept A to a different degree.

While it is not possible to explicitly construct a model that uses infinitely many membership degrees in finite time, we can still decide its existence by considering the local ordering relations between the membership degrees of all relevant concepts, at every node and its parents. As seen in the example above, it is possible to provide an abstract description of the models of interest through a preorder over all subconcepts and membership degrees explicitly appearing in the input TBox and set of restrictions. Figure 1 provides an abstract representation of all models of $\mathcal{T}$ that satisfy the restriction $\mathcal{R}$. In the figure, $A_{\uparrow}$ represents the membership degree of the parent node to A. As it can be seen, although the models of this TBox satisfying the restriction can be arbitrarily complex, they can all be represented using a very simple recurrent structure. In general, the existence of a model satisfying a set of restrictions can be characterised through *Hintikka trees*.

Consider the set $\mathcal{U} := \mathsf{V}_{\mathcal{T},\mathcal{R}} \cup \mathsf{sub}(\mathcal{T},\mathcal{R}) \cup \mathsf{sub}_{\uparrow}(\mathcal{T},\mathcal{R}) \cup \{\lambda\}$, where $\mathsf{V}_{\mathcal{T},\mathcal{R}}$ represents the set of all constants appearing in the input extended with $0, 1$, $\mathsf{sub}(\mathcal{T},\mathcal{R})$ is the set of all subconcepts from $\mathcal{T},\mathcal{R}$, and λ is an arbitrary new symbol. A *Hintikka order* is a total preorder $\lesssim$ over $\mathcal{U}$ that preserves the standard ordering of real numbers over $\mathsf{V}_{\mathcal{T},\mathcal{R}}$ and is consistent with the semantics of the propositional constructors. For example, if $X, C \sqcap D \in \mathcal{U}$ and $X \lesssim C \sqcap D$, then it must also hold that $X \lesssim C$ and $X \lesssim D$. All other cases can be treated similarly. Intuitively, a Hintikka ordering represents the relation between the membership degrees at a specific element of the domain of an interpretation. To ensure that it is a model, this ordering must also be *compatible* with the GCIs in the TBox; that is, for every $\langle C \sqsubseteq D \geq q \rangle \in \mathcal{T}$, either $C \lesssim D$ or $q \lesssim D$.

Existential and value restrictions are verified producing a sequence of successors that witness them. For each existential restriction $E = \exists r.C$ in the input, every node in the Hintikka tree has a distinguished successor $\phi(E)$. The Hintikka ordering associated with this node is required to satisfy $(\exists r.C)_{\uparrow} \equiv \min(\lambda, C)$, thus serving as a witness for the concept at the parent node. Moreover, for all other successors associated to a concept quantified over the same role r, the ordering must satisfy $\min(\lambda, C) \lesssim (\exists r.C)_{\uparrow}$. These conditions ensure that the semantics of existential restrictions are satisfied. Similar conditions guarantee the satisfaction of value restrictions $\forall r.C$.

A *Hintikka tree* for $\mathcal{T},\mathcal{R}$ is an infinite tree of constant arity where every node is labelled with a Hintikka ordering compatible with the TBox $\mathcal{T}$, the successors satisfy the transition conditions, and the root node satisfies the restrictions in $\mathcal{R}$.

It can be shown that $\mathcal{R}$ is consistent w.r.t. $\mathcal{T}$ if and only if there is a Hintikka tree for $\mathcal{T}, \mathcal{R}$. Notice moreover that there are only finitely many partial orderings over the set $\mathcal{U}$, and hence also finitely many Hintikka orderings. In fact, the number of Hintikka orderings is bounded exponentially by the size of the input.

To decide the existence of a Hintikka tree, we construct a simple looping automaton on (unlabeled) infinite trees. The set of Hintikka orderings defines the states of the automaton; the transition relation is determined by the transition conditions for quantified concepts; and the initial states are those that satisfy the input restrictions. Essentially, the successful runs of this automaton correspond to the Hintikka trees sought. Thus, the automaton has a successful run iff a Hintikka tree for $\mathcal{T}, \mathcal{R}$ exists iff $\mathcal{R}$ is consistent w.r.t. $\mathcal{T}$. For further details see [4].

This automata-based decision procedure not only provides a tight complexity bound for reasoning in the fuzzy DL G-$\mathcal{ALC}$. It also opens the door to the application of other automata-based techniques, originally developed for classical DLs (e.g. [1, 2]), to this and other fuzzy DLs based on the Gödel t-norm.

Acknowledgements

The ideas presented in this paper have arisen from joint work and lengthy discussions with Franz Baader, Stefan Borgwardt, and Felix Distel, from TU Dresden, Germany.

This work is partially supported by the German Research Foundation (DFG) under grant BA 1122/17-1 (FuzzyDL) and within the Cluster of Excellence 'Center for Advancing Electronics Dresden' (cfAED).

References

1. F. Baader, J. Hladik, and R. Peñaloza. Automata can show PSPACE results for description logics. *Information and Computation*, 206(9-10):1045–1056, 2008.
2. F. Baader and R. Peñaloza. Automata-based axiom pinpointing. *Journal of Automated Reasoning*, 45(2):91–129, August 2010. Special Issue: Selected Papers from IJCAR 2008.
3. F. Bobillo, F. Bou, and U. Straccia. On the failure of the finite model property in some fuzzy description logics. *Fuzzy Sets and Systems*, 172(1):1–12, 2011.
4. S. Borgwardt, F. Distel, and R. Peñaloza. Gödel description logics: Decidability in the absence of the finitely-valued model property. LTCS-Report 13-09, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany, 2013. See `http://lat.inf.tu-dresden.de/research/reports.html`.
5. S. Borgwardt and R. Peñaloza. Undecidability of fuzzy description logics. In G. Brewka, T. Eiter, and S. A. McIlraith, editors, *Proc. of the 13th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'12)*, pages 232–242. AAAI Press, 2012.
6. U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.