Reasoning and Quantification in Fuzzy Description Logics

Daniel Sánchez[1] and Andrea G.B. Tettamanzi[2]

[1] University of Granada, Department of Computer Science and Artificial Intelligence, Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
daniel@decsai.ugr.es
[2] University of Milan, Department of Information Technologies, Via Bramante 65, I-26013 Crema (CR), Italy
andrea.tettamanzi@unimi.it

Abstract. In this paper we introduce reasoning procedures for $\mathcal{ALCQ}_F^+$, a fuzzy description logic with extended qualified quantification. The language allows for the definition of fuzzy quantifiers of the absolute and relative kind by means of piecewise linear functions on $\mathbb{N}$ and $\mathbb{Q} \cap [0, 1]$ respectively. In order to reason about instances, the semantics of quantified expressions is defined based on recently developed measures of the cardinality of fuzzy sets. A procedure is described to calculate the fuzzy satisfiability of a fuzzy assertion, which is a very important reasoning task. The procedure considers several different cases and provides direct solutions for the most frequent types of fuzzy assertions.

1 Introduction

Description logics (DL) [1] are a family of logic-based knowledge-representation formalisms, which stem from the classical AI tradition of semantic networks and frame-based systems. DLs are well-suited for the representation of and reasoning about terminological knowledge, configurations, ontologies, database schemata, etc.

The need of expressing and reasoning with imprecise knowledge and the difficulties arising in classifying individuals with respect to an existing terminology is motivating research on nonclassical DL semantics, suited to these purposes. To cope with this problem, fuzzy description logics have been proposed that allow for imprecise concept description by using fuzzy sets and fuzzy relations. For instance, a fuzzy extension of the description logic $\mathcal{ALC}$ has been introduced in [6], with complete algorithms for solving the entailment problem, the subsumption problem, as well as the best truth-value bound problem.

In [5] we introduced $\mathcal{ALCQ}_F^+$, a fuzzy description logic with extended qualified quantification[1] that allows for the definition of fuzzy quantifiers of the absolute

[1] In keeping with DL naming conventions, the superscript plus is to suggest that, in addition to **q**ualified number restrictions available in the description logic $\mathcal{ALCQ}$ introduced by De Giacomo and Lenzerini [2], we provide also more general fuzzy linguistic **q**uantifiers. The subscript F means that the language deals with infinitely many truth-values, as in the language $\mathcal{ALC}_{F_M}$ of Tresp and Molitor [7].

I. Bloch, A. Petrosino, and A.G.B. Tettamanzi (Eds.): WILF 2005, LNAI 3849, pp. 81–88, 2006.

and relative kind by means of piecewise linear functions on $\mathbb{N}$ and $\mathbb{Q} \cap [0,1]$ respectively. These quantifiers extends the usual (qualified) $\exists$, $\forall$ and number restriction.

Incorporating fuzzy quantification into fuzzy description logics is important by several reasons. On the one hand, number restriction is a kind of quantification that arises very frequently in concept description, so it is necessary to extend it to the fuzzy case. But another important reason is that not only concepts, but also quantifiers are imprecise in many cases (e.g. "around two", "most").

For example, suppose you are the marketing director of a supermarket chain. You are about to launch a new line of low-calorie products. In order to set up your budget, you need to project the sales of this new line of products. This can be done either by means of an expensive market research, or by means of some kind of inference based on your knowledge of customer habits. For instance, you could expect prospective buyers of this new line of products to be essentially faithful customers who mostly buy foods with low energy value. We have here all the ingredients of imprecise knowledge: a "faithful customer" is a fuzzy concept; "low" energy value is a linguistic value, which might be modelled as a fuzzy number; to "mostly" buy a given kind of product is equivalent to a quantified statement of the form "most of the bought products are of this kind", where "most" is an imprecise quantifier.

Zadeh [8] showed that imprecise quantifiers can be defined by using fuzzy sets, and by incorporating them into the language and providing the tools to define their semantics we can provide a very powerful knowledge representation tool, with greater expressive power, and closer to the humans' way of thinking.

2 The Language $\mathcal{ALCQ}_F^+$

The language $\mathcal{ALCQ}_F^+$ has the following syntax:

$\langle$concept_description$\rangle$ ::= $\langle$atomic_concept$\rangle$ |
$\top$ | $\bot$ | $\neg\langle$concept_description$\rangle$ |
$\langle$concept_description$\rangle \sqcap \langle$concept_description$\rangle$ |
$\langle$concept_description$\rangle \sqcup \langle$concept_description$\rangle$ |
$\langle$quantification$\rangle$
$\langle$quantification$\rangle$ ::= $\langle$quantifier$\rangle\langle$atomic_role$\rangle$.$\langle$concept_description$\rangle$
$\langle$quantifier$\rangle$ ::= "(" $\langle$absolute_quantifier$\rangle$ ")" | "(" $\langle$relative_quantifier$\rangle$ ")" |
$\exists$ | $\forall$
$\langle$absolute_quantifier$\rangle$ ::= $\langle$abs_point$\rangle$ | $\langle$abs_point$\rangle$ + $\langle$absolute_quantifier$\rangle$
$\langle$relative_quantifier$\rangle$ ::= $\langle$fuzzy_degree$\rangle$/u | $\langle$fuzzy_degree$\rangle$/u + $\langle$piecewise_fn$\rangle$
$\langle$piecewise_fn$\rangle$ ::= $\langle$rel_point$\rangle$ | $\langle$rel_point$\rangle$ + $\langle$piecewise_fn$\rangle$
$\langle$abs_point$\rangle$ ::= $\langle$val$\rangle$/$\langle$natural_number$\rangle$
$\langle$rel_point$\rangle$ ::= $\langle$val$\rangle$/$\langle$[0,1]-value$\rangle$
$\langle$val$\rangle$ ::= [$\langle$fuzzy_degree$\rangle \lhd$]$\langle$fuzzy_degree$\rangle$[$\rhd \langle$fuzzy_degree$\rangle$]

In this extension, the semantics of quantifiers is defined by means of piecewise-linear membership functions. In the case of absolute quantifiers, the quantifier is obtained by restricting the membership function to the naturals. The semantics of fuzzy assertions in $\mathcal{ALCQ}_F^+$ is given by the standard fuzzy conjunction, disjunction and negation, as well as method GD [3] for quantified expressions.

The piecewise-linear functions are defined by means of a sequence of points. These points are expressed as $\alpha \triangleleft \beta \triangleright \gamma/x$, where x is the cardinality value, β is the membership degree of x, and α and γ are the limit when the membership function goes to x from the left and from the right, respectively. When the function is continuous, this can be summarized as β/x (since $\alpha = \beta = \gamma$), whereas discontinuities on the left ($\alpha \neq \beta = \gamma$) or right ($\alpha = \beta \neq \gamma$) can be summarized as $\alpha \triangleleft \beta/x$ and $\beta \triangleright \gamma/x$, respectively.

2.1 An Example

Let us go back to the example of the marketing director of a supermarket chain about to launch a line of low-calorie products.

The knowledge base describing the business of running a supermarket chain could contain, among others, the following terminological axioms:

$$\begin{aligned} \mathsf{FaithfulCustomer} &\sqsubseteq \mathsf{Customer} \sqsubseteq \top \\ \mathsf{FoodProduct} &\sqsubseteq \mathsf{Product} \sqsubseteq \top \\ \mathsf{LowCalorie} &\sqsubseteq \mathsf{EnergyMeasure} \sqsubseteq \top \\ \mathsf{LowCalorieFood} &\equiv \mathsf{FoodProduct} \sqcap \forall \mathsf{energyValue.LowCalorie} \end{aligned}$$

The ABox describing facts about your supermarket chain might contain TVBs which we might summarize as follows:

– given an individual customer c and a product p, $\mathsf{buys}(c, p)$ might be interpreted as
$$\mathsf{buys}(c, p) = f(\mathrm{weeklyrevenue}(c, p)),$$
where $f : \mathbb{R} \to [0, 1]$ is nondecreasing, and weeklyrevenue(c, p) : $\mathsf{Customer}^{\mathcal{I}} \times \mathsf{Product}^{\mathcal{I}} \to \mathbb{R}$ returns the result of a database query which calculates the average revenue generated by product p on customer c in all the stores operated by the chain;
– given an individual customer c, $\mathsf{FaithfulCustomer}(c)$ might be interpreted as
$$\mathsf{FaithfulCustomer}(c) = g(\mathrm{weeklyrevenue}(c)),$$
where $g : \mathbb{R} \to [0, 1]$ is nondecreasing, and weeklyrevenue(c) : $\mathsf{Customer}^{\mathcal{I}} \to \mathbb{R}$ returns the result of a database query which calculates the average revenue generated by customer c in all the stores operated by the chain;
– finally, $\mathsf{LowCalorie}(x)$, where x is an average energy value per 100 g of product measured in kJ, could be interpreted as
$$\mathsf{LowCalorie}(x) = \begin{cases} 1 & x < 1000, \\ \frac{2000-x}{1000} & 1000 \leq x \leq 2000, \\ 0 & x > 2000. \end{cases}$$

Table 1. The energy value, membership in the LowCalorieFood, and the degree to which customer CARD0400009324198 buys them for a small sample of products

Product	Energy [kJ/hg]	LowCalorieFood(·)	buys(CARD..., ·)
GTIN8001350010239	1680	0.320	0.510
GTIN8007290330987	1475	0.525	0.050
GTIN8076809518581	1975	0.025	0.572
GTIN8000113004003	1523	0.477	0.210
GTIN8002330006969	498	1.000	1.000
GTIN8005410002110	199	1.000	1.000
GTIN017600081636	1967	0.033	0.184

By using the $\mathcal{ALCQ}_F^+$ language, it is now possible to express the notion of a faithful customer who mostly buys food with low energy value as

$$C \equiv \mathsf{FaithfulCustomer} \sqcap (\mathsf{Most})\mathsf{buys.LowCalorieFood},$$

where $(\mathsf{Most}) \equiv (0/u + 0/0.5 + 1/0.75)$.

A useful deduction this new axiom allows you to make is, for instance, calculating the extent to which a given individual customer or, more precisely, a fidelity card, say CARD0400009324198, is a C. For instance, you could know that

$$\mathsf{FaithfulCustomer(CARD0400009324198)} = 0.8,$$

and, by querying the sales database, you might get all the degrees to which that customer buys each product. For sake of example, we give a small subset of those degrees of truth in Table 1, along with the energy values of the relevant products.

According to the semantics of $\mathcal{ALCQ}_F^+$,

$$C(\mathsf{CARD0400009324198}) \approx 0.742$$

i.e., the degree to which most of the items purchased by this customer are low-calorie is around 0.742. This seems to be in accordance with the data in Table 1,

Table 2. Percentage of purchased items that are low-calorie at significant levels

Level	Percentage
1.000	1.000 = 2/2
0.572	0.667 = 2/3
0.510	0.500 = 2/4
0.320	0.750 = 3/4
0.210	0.800 = 4/5
0.184	0.667 = 4/6
0.050	0.714 = 5/7
0.033	0.857 = 6/7

where we can see that four products (those products p in rows 2, 4, 5, and 6) verify

$$\mathsf{buys}(\mathsf{CARD0400009324198}, p) \leq \mathsf{LowCalorieFood}(p)$$

while for the products in rows 1 and 7 the difference between being purchased and being low-calorie food is not so high. Only the item in row 3 seems to be a clear case of item purchased but not low-calorie.

As another justification of why this result appears in agreement with the data, in Table 2 we show the percentage of purchased items that are low-calorie at α-cuts of the same level. At any other level, the percentage obtained is one of those shown in Table 2.

At many levels the percentage is above 0.75, therefore fitting the concept of Most as we have defined it. At level 0.050 the percentage is almost 0.75. The only level that clearly doesn't fit Most is 0.510, but at the next level (0.320) we have again 0.75 and $\mathsf{Most}(0.75) = 1$.

3 Reasoning with $\mathcal{ALCQ}_F^+$

Of course, the purpose of a knowledge representation system goes beyond storing concept definitions and assertions. A knowledge representation system based on fuzzy DLs should be able to perform specific kinds of reasoning. One particularly important reasoning task is to calculate the fuzzy satisfiability of a fuzzy assertion Ψ, i.e., the interval of values $\mathcal{S}(\Psi) = [\beta_\Psi, \tau_\Psi]$ such that for any interpretation $\mathcal{I}$, the degree of truth of Ψ under $\mathcal{I}$, noted $\text{truth}_\mathcal{I}(\Psi)$, verifies $\text{truth}_\mathcal{I}(\Psi) \in [\beta_\Psi, \tau_\Psi]$ (i.e., the maximum interval $[\beta, \tau]$ such that Ψ is $[\beta, \tau]$–satisfiable in the sense of Navara's definition [4]).

In this work we introduce a PSPACE-complete algorithm to calculate the fuzzy satisfiability of a fuzzy assertion. Though infinite interpretations are taken into account in the definition of the semantics of the language $\mathcal{ALCQ}_F^+$, in practice and due to the physical limitations of computers, we are going to deal with a finite number of individuals. The same limitations put a bound on the number of different membership degrees we can deal with. Therefore we shall calculate the fuzzy satisfiability of a fuzzy assertion up to a certain precision degree, given as a number of decimals p.

In addition to the general algorithm, we have obtained some results showing that we can calculate the fuzzy satisfiability of a fuzzy assertion directly in some (the most common) cases. Some of these results are based on a previous result about *independence* of fuzzy assertions. We introduce the following definition:

Definition 1. *Two concepts A and B are* independent *of each other if, for all $\alpha \in \mathcal{S}(A)$ and $\beta \in \mathcal{S}(B)$, there exists an interpretation $\mathcal{I}$ containing an individual $d \in \Delta_\mathcal{I}$ such that*

$$A^\mathcal{I}(d) = \alpha \quad \textit{and} \quad B^\mathcal{I}(d) = \beta.$$

In other words, A and B are independent if the degree of truth of A does not affect the degree of truth of B and *vice versa*.

In order to determine whether two concepts are independent, we provide the following results:

Proposition 1. *Two concepts A and B, not containing quantifiers, which are neither tautologies nor contradictions, are independent if and only if the following four concepts are satisfiable in the crisp sense: $A \sqcap B$; $A \sqcap \neg B$; $\neg A \sqcap B$; $\neg A \sqcap \neg B$.*

Proposition 2. *Let D_1 and D_2 be two independent concepts. Then, no atomic concept and no role appears in the expansion of both.*

Proposition 3. *Let D_1 be a concept that doesn't contain quantifiers in its expansion, and let $D_2 \equiv QR.C$. Then, D_1 and D_2 are independent, regardless of whether D_1 and C are independent or not.*

Proposition 4. *Let $D_1 \equiv Q_1 R_1.C_1$ and let $D_2 \equiv Q_2 R_2.C_2$. If $R_1 \neq R_2$ or C_1 and C_2 are independent, then D_1 and D_2 are independent.*

A general procedure to determine whether two concepts are independent can be obtained from Proposition 1 using the crisp procedure to check unsatisfiability. However, Propositions 2, 3, and 4 can make things easier in some cases.

On this basis, we introduce the following results on the calculation of satisfiability of fuzzy assertions:

– If A is an atomic concept, then $\mathcal{S}(A) = [0, 1]$.
– (Negation) If $\mathcal{S}(D) = [\beta_D, \tau_D]$, then $\mathcal{S}(\neg D) = [1 - \tau_D, 1 - \beta_D]$. This verifies $\mathcal{S}(\neg\neg D) = \mathcal{S}(D)$ and:
  • if A is atomic, $\mathcal{S}(\neg A) = \mathcal{S}(A) = [0, 1]$;
  • if $C = C_1 \sqcap \cdots \sqcap C_r$, then $\mathcal{S}(\neg C) = \mathcal{S}(\neg C_1 \sqcup \cdots \sqcup \neg C_r)$;
  • if $C = C_1 \sqcup \cdots \sqcup C_r$, then $\mathcal{S}(\neg C) = \mathcal{S}(\neg C_1 \sqcap \cdots \sqcap \neg C_r)$;
  • if $D \equiv QR.C$,

$$\mathcal{S}(\neg D) = \mathcal{S}(\neg(QR.C)) = \mathcal{S}((\neg Q)R.C). \tag{1}$$

– Let Q be an absolute quantifier such that $\text{core}(Q) \neq \emptyset$ and $\mathbb{N} \backslash \text{supp}(Q) \neq \emptyset$. If $D \equiv QR.C$ and $\mathcal{S}(C) = [\beta_C, \tau_C]$, $\mathcal{S}(D) = \mathcal{S}(QR.C) = [\beta_D, \tau_D]$, with

$$\beta_D = (1 - \tau_C)Q(0) \tag{2}$$

$$\tau_D = \max\{(\tau_C + (1 - \tau_C)Q(0)), Q(0)\} \tag{3}$$

– Let Q be a relative quantifier such that $\text{core}(Q) \neq \emptyset$ and $[0,1] \cap \mathbb{Q} \backslash \text{supp}(Q) \neq \emptyset$ with $u^Q = Q(x/0)$ (i.e., u^Q is the value returned by the quantifier when the relative cardinality is undefined). If $D \equiv QR.C$ and $\mathcal{S}(C) = [\beta_C, \tau_C]$,
  • If $Q(0) = 0$ and $Q(1) = 1$,

$$\mathcal{S}(D) = [\min\{u^Q, \beta_C\}, \tau_C + (1 - \tau_C)u^Q]$$

In particular, for quantifiers $\exists$ and $\forall$ we have $u^\exists = 0$, $u^\forall = 1$, and

$$\mathcal{S}(\exists R.C) = [0, \tau_C],$$
$$\mathcal{S}(\forall R.C) = [\beta_C, 1].$$

- If $Q(0) = 1$ and $Q(1) = 0$,

$$\mathcal{S}(QR.C) = [1 - (\tau_C + (1 - \tau_C)u^Q), 1 - \min\{u^Q, \beta_C\}].$$

- If $Q(0) = 0$ and $Q(1) = 0$,

$$\mathcal{S}(D) = [0, \max\{u^Q, (1 - \tau_C)u^Q + (\tau_C - \beta_C)\}].$$

- If $Q(0) = 1$ and $Q(1) = 1$,

$$\mathcal{S}(QR.C) = [1 - \max\{u^Q, (1 - \tau_C)u^Q + (\tau_C - \beta_C)\}, 1].$$

The remaining cases are solved by means of an $O(1)$ algorithm with a fixed precision of p decimals.

– If $D \equiv D_1 \sqcup D_2 \sqcup \cdots \sqcup D_s$ with $\mathcal{S}(D_i) = [\beta_{D_i}, \tau_{D_i}]$, then $\mathcal{S}(D) = [\beta_D, \tau_D]$, with

$$\beta_D \geq \max_{i \in \{1,\ldots,s\}} \{\beta_{D_i}\},$$
$$\tau_D = \max_{i \in \{1,\ldots,s\}} \{\tau_{D_i}\}.$$

In particular, if the fuzzy assertions D_i are pairwise independent,

$$\beta_D = \max_{i \in \{1,\ldots,s\}} \{\beta_{D_i}\},$$
$$\tau_D = \max_{i \in \{1,\ldots,s\}} \{\tau_{D_i}\}.$$

The value of β_D with a precision degree of p decimals, when the D_i are not independent, is obtained by means of an algorithm that performs a dichotomic search after guessing values of the atomic concepts and roles that appear in the D_i's.

– The satisfiability of conjunctions can be obtained by using De Morgan's laws and the latter result.

4 Conclusions

$\mathcal{ALCQ}_F^+$ allows for concept description involving fuzzy linguistic quantifiers of the absolute and relative kind, and using qualifiers. We have introduced algorithms to perform two important reasoning tasks with this logic: reasoning about instances, and calculating the fuzzy satisfiability of a fuzzy assertion. In addition, we have defined *independence* of fuzzy assertions and obtained some results that speed up the calculation of fuzzy satisfiability in some (the most common) cases.

References

1. Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook: Theory, implementation and applications*. Cambridge, 2003.

2. Giuseppe De Giacomo and Maurizio Lenzerini. A uniform framework for concept definitions in description logics. *Journal of Artificial Intelligence Research*, 6:87–110, 1997.
3. M. Delgado, D. Sánchez, and M.A. Vila. Fuzzy cardinality based evaluation of quantified sentences. *International Journal of Approximate Reasoning*, 23:23–66, 2000.
4. Mirko Navara. Satisfiability in fuzzy logics. *Neural Networks World*, 10(5):845–858, 2000.
5. Daniel Sánchez and Andrea G. B. Tettamanzi. Generalizing quantification in fuzzy description logics. In *Proceedings 8th Dortmund Fuzzy Days*, Dortmund, Germany, 2004.
6. Umberto Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.
7. Christopher B. Tresp and Ralf Molitor. A description logic for vague knowledge. In *Proceedings of the 13th biennial European Conference on Artificial Intelligence (ECAI'98)*, pages 361–365, Brighton, UK, 1998. J. Wiley and Sons.
8. L. A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computing and Mathematics with Applications*, 9(1):149–184, 1983.