Expressive Description Logics with Rich Yet Affordable Numeric Constraints (Extended Abstract)

Federica Di Stefano, Sanja Lukumbuzya, Magdalena Ortiz and Mantas Šimkus

Institute of Logic and Computation, TU Wien

Abstract

We summarize our recent work on extending Description Logics with numeric constraints presented in [1]. We add numeric features to the expressive DL $\mathcal{ALCHOIQ}$ with closed predicates, and study reasoning problems that go beyond satisfiability, such as finding models that optimize user-defined objectives.

Keywords

Expressive DLs with Closed Predicate, Numeric Constraints, Optimal Models, Complexity of Reasoning

1. Introduction

Description Logics (DLs) have proven to be excellent formalisms for describing different domains and reasoning about them, but their very limited support for numeric constraints and quantitative reasoning is a significant weakness. In many applications, one would like to be able to refer to integers or real numbers and basic comparisons between them. In expressive DLs like $\mathcal{ALCHIOQ}$, we can describe complex requirements in the form of concept inclusions, such as $\mathsf{Project} \sqcap \neg\{\mathsf{proj1}\} \sqsubseteq\, \geq 3\, \mathsf{assignedTo}^{-}.\mathsf{Employee}$, but we cannot reason, e.g., about the number of hours that employees work for a project, or verify that the sum of the monthly salaries of all employees assigned to a project does not exceed the corresponding budget. Standard DLs cannot express this type of quantitative constraints, and adding them is far from easy, and even simple numeric reasoning easily becomes computationally costly, or even undecidable.

Overcoming this limitation is a long-standing challenge that has received significant attention since the early days of DL research. The most prominent line of work here is *concrete domains* [2]. In addition to concepts and roles, we use concrete *features* such as age, salary, size, etc., to relate objects to values from a domain, like the reals or the integers, for example. A fixed set of predicates over these domains, such as addition and comparisons, can then be used to incorporate numeric constraints into concept descriptions. DLs with concrete domains are very powerful, but this comes at a high computational cost. Strong restrictions must be imposed on the concrete domains to preserve decidability (see [3, 4] and their references). Most works require the so-called ω-admissibility, which, in a nutshell, guarantees that the infinite systems of constraints that may arise when reasoning about infinite models can be effectively decided. Tight complexity bounds for standard reasoning tasks like concept satisfiability and instance checking have been obtained for expressive DLs with ω-admissible concrete domains [3, 4]—and even a few isolated results for non-ω-admissible ones [5, 6]—but those works focus on worst-case bounds and generic restrictions that preserve decidability in many domains, rather than on practical usability. In contrast, efforts to support powerful and useful domains by restricting the interaction with the logics, and practicable algorithms for them, have been limited [7, 8].

When DLs are extended with numeric features, many novel and natural reasoning problems arise. For example, if a feature corresponds to a cost, it is natural to minimize it: a company may not only be interested in staying within a given budget when fulfilling certain requirements, but may naturally want to minimize their cost. Given the open domains of DLs, the cost associated with a specific feature may not always be finite, so it is natural to ask whether the KB can be satisfied while guaranteeing that certain cost remains bounded, or below a certain value. Popular reasoning problems such as concept

DL 2025: 38th International Workshop on Description Logics, September 3–6, 2025, Opole, Poland

federica.stefano@tuwien.ac.at (F. Di Stefano); sanja.lukumbuzya@tuwien.ac.at (S. Lukumbuzya); magdalena.ortiz@tuwien.ac.at (M. Ortiz); mantas.simkus@tuwien.ac.at (M. Šimkus)

satisfiability, instance checking, and query answering can all be defined in terms of optimal models. Additionally, one can ask what values certain features might take when other features are optimized. Despite their evident potential, to our knowledge, this kind of numeric reasoning service has not yet been developed for DLs.

This extended abstract summarizes our recent work [1]. We enhance very expressive DLs with simple yet useful numeric reasoning capabilities, and propose a toolbox of reasoning services that seamlessly integrate ontological and numeric reasoning. We use features to assign numeric values to domain objects, and the concept descriptions in our knowledge bases may comprise constraints on the sums of the feature values in the neighborhood of objects. As usual in DLs, domains may be arbitrarily large, but we require a finite bound on the range of possible feature values of each object. In this way, decidability is not compromised, even in very expressive DLs like $\mathcal{ALCHOIQ}$ and its extension with closed predicates [9]. Reasoning in our numeric extension can be seamlessly achieved with the standard reasoning algorithm for $\mathcal{ALCHOIQ}$ with closed predicates by reduction to a system of linear inequalities. We identify some conditions on the numeric ranges that guarantee that the worst-case complexity of reasoning is not higher than for plain $\mathcal{ALCHOIQ}$. A highlight of our approach is its natural support for many novel reasoning services, thus revealing a new tool for flexible quantitative inference in the presence of rich ontological knowledge.

2. The Formalism

We refer to [10] for standard preliminaries on DLs. We extend DLs with closed predicates [11] by allowing to assign numeric values to concept names and control how some values for the objects compare to aggregated values of its neighbors. For example, assuming an ontology describing a company setting, we want to be able to assign possible salaries to employees within described ranges, and ensure that the sum of yearly salaries of the employees does not exceed a given yearly budget. We introduce *features*, which assign numeric values to elements, and a new form of concept expressions that locally restricts feature values. Let N_F be a countably infinite set of *feature*, disjoint from the sets of concept names N_C, role names N_R and individual names N_I. Let $N_C^+ = N_C \cup \{\top, \bot\} \cup \{\{a\} | \, a \in N_I\}$ be the set of basic concepts. Given a KB $\mathcal{K}$, we denote with $N_F(\mathcal{K})$ and $N_C^+(\mathcal{K})$ the sets of features and basic concepts (resp.) occurring in $\mathcal{K}$.

Definition 1. *Let $\mathcal{L}$ be a DL with closed predicates. A* feature annotation α *is a partial function from $N_C^+ \times N_F$ to finite discrete subsets of the non-negative rationals $\mathbb{Q}^+ \cup \{0\}$. We call the set $D = \alpha(B, f)$ the* domain *of feature f for B. A* neighborhood restriction *takes the form*

$$w_0 + w_1 \sum f_1[r_1.C_1] + \cdots + w_n \sum f_n[r_n.C_n] \circ wf, \tag{$*$}$$

where $\circ \in \{<, \leq, \geq, >\}$, $r_1, \ldots, r_n$ are roles, $C_1, \ldots, C_n$ are concepts, $f, f_1, \ldots, f_n$ are features, and $w, w_1, \ldots, w_n$ are non-negative rational numbers. Concepts in $\mathcal{L}^{\mathrm{NF}}$ are defined as for $\mathcal{L}$, but also allowing neighborhood restrictions. A KB is defined as a 4-tuple $\mathcal{K} = (\mathcal{T}, \Sigma, \alpha, \mathcal{A})$, where $(\mathcal{T}, \Sigma, \mathcal{A})$ is an $\mathcal{L}^{\mathrm{NF}}$ KB with closed predicates and α is a feature annotation function.

The feature annotation function α determines the set D of values that objects participating in a basic concept B can take as the value of feature f. For instance, $\alpha(\mathsf{ExecEmpl}, \mathit{salary}) = [3.5, 5.5, +0.1]$ indicates that executive employees can have their salary-value between 3.5k and 5.5k, with increases of 100€. In our case, we restrict our attention to sets D that consist of finitely many, non-negative rational values. Note that α is a partial function, as for instance we may not want to prescribe values for $\alpha(\mathsf{Office}, \mathit{salary})$, since offices have no salary.

To define the semantics of our formalism, we extended standard interpretation functions to features by requiring that an interpretation function $\cdot^{\mathcal{I}}$ assigns to every feature $f \in N_F$ a partial function $f^{\mathcal{I}}$ from $\Delta^{\mathcal{I}}$ to non-negative rational values. The semantics of a neighborhood restriction $(*)$ is now straightforward: at a domain element, we compare the feature value of f multiplied by the weight w

with the weighted sum of the feature values of the relevant neighbors, i.e.

$$\begin{aligned}&(w_0 + w_1 \sum f_1[r_1.C_1] + \cdots + w_n \sum f_n[r_n.C_n] \circ wf)^{\mathcal{I}}\\ &= \{o \in \Delta^{\mathcal{I}} : \big(w_0 + \sum_{\substack{(o,o')\in r_i^{\mathcal{I}}, o' \in C_i^{\mathcal{I}},\\ f_i^{\mathcal{I}}(o') \text{ def.}, 1\leq i\leq n}} w_i f_i^{\mathcal{I}}(o')\big) \circ w f^{\mathcal{I}}(o)\}\end{aligned}$$

Satisfaction of axioms in $\mathcal{I}$ involving the new concept constructor is defined as usual.

Definition 2. *Given a KB $\mathcal{K} = (\mathcal{T}, \Sigma, \alpha, \mathcal{A})$, an interpretation $\mathcal{I}$ is $\mathcal{K}$-*suitable*, if for each $e \in \Delta^{\mathcal{I}}$, each feature $f \in N_F(\mathcal{K})$, and each basic concept $A \in N_C^+(\mathcal{K})$ with $e \in A^{\mathcal{I}}$ and $\alpha(A, f)$ defined, we have that $f^{\mathcal{I}}(e) \in \alpha(A, f)$. We say that $\mathcal{I}$ *satisfies* $\mathcal{K}$, and call $\mathcal{I}$ a *model* of $\mathcal{K}$, if (i) $\mathcal{I}$ is $\mathcal{K}$-suitable and (ii) $\mathcal{I}$ satisfies $(\mathcal{T}, \Sigma, \mathcal{A})$.*

3. Reasoning Services and their Complexity

Standard reasoning tasks can be easily defined in our setting. By adapting the well-known *mosaic technique* [12, 13] for $\mathcal{ALCHOIQ}$ with closed predicates [9], we provide tight complexity results. Roughly speaking, the core idea of the mosaic technique is to succinctly represent small fragments of models that serve as variables of a system of inequalities. A solution to the system consists of a set of fragments that can be successfully plugged together into a model. Given a KB, to accommodate our numeric features and constraints, we extend the model fragments of [9] with vectors called *feature types*, where the i-th position corresponds to the value of the feature f_i or is undefined. Via the feature types, we ensure that when constructing the model, we respect the feature annotations of $\mathcal{K}$ and satisfy the axioms with neighborhood constraints.

Theorem 1. *Satisfiability of $\mathcal{ALCHOIQ}^{\mathrm{NF}}$ KBs with closed predicates is NExpTime-complete.*

Our formalism supports other novel reasoning problems that arise naturally by restricting the models of KBs to those that are *optimal* w.r.t. some given objective. Given a KB $\mathcal{K}$, a *cost function for $\mathcal{K}$* is an expression of the form $F = w_0 + \sum_{i=1}^{n} w_i \cdot f_i[B_i]$, where $w_0, w_i \in \mathbb{Q}^+$ are *weights*, $f_i \in N_F(\mathcal{K})$, and $B_i \in N_C^+(\mathcal{K})$, for all $1 \leq i \leq n$. Given a concrete interpretation $\mathcal{I}$ and a cost function F, the *value of $\mathcal{I}$ w.r.t. F*, denoted $v_F(\mathcal{I})$ is defined as

$$v_F(\mathcal{I}) = w_0 + \sum_{i=1}^{n} \sum_{d\in B_i^{\mathcal{I}} \text{ s.t. } f_i^{\mathcal{I}}(d) \text{ def.}} w_i f_i^{\mathcal{I}}(d).$$

Definition 3. *Given $\mathcal{K} = (\mathcal{T}, \Sigma, \alpha, \mathcal{A})$ and a cost function F, an interpretation $\mathcal{I}$ is an* optimal model *of $\mathcal{K}$ w.r.t. F if: (i) $\mathcal{I} \models \mathcal{K}$, (ii) $v_F(\mathcal{I})$ is finite, and (iii) there exists no $\mathcal{J}$ such that $\mathcal{J} \models \mathcal{K}$ and $v_F(\mathcal{J}) \leq v_F(\mathcal{I})$.*

Standard reasoning tasks such as KB satisfiability, concept satisfiability, and instance checking can be easily transferred to the setting of *optimal models.* Furthermore, we can define ad hoc reasoning tasks in which we check if a specific requirement over feature values is fulfilled. We call *cost query* an expression q of the form $F \circ w$, where F is a cost function, w is a non-negative rational, and $\circ \in \{<, \leq, =, \geq, >\}$. We say that a cost query q is true in an interpretation $\mathcal{I}$, in symbols $\mathcal{I} \models q$, if $v_F(\mathcal{I}) \circ w$ is true. The reasoning tasks of *brave* and *cautious* query answering can be easily defined for cost queries. Given a cost query q and a KB $\mathcal{K}$, we say that q is bravely entailed by $\mathcal{K}$ if there exists a model $\mathcal{I}$ of $\mathcal{K}$ such that $\mathcal{I} \models q$. While, we say that q is cautiously entailed by $\mathcal{K}$ if $\mathcal{I} \models q$, for all models $\mathcal{I}$ of $\mathcal{K}$.

Theorem 2. *In $\mathcal{ALCHIOQ}^{\mathrm{NF}}$, KB satisfiability w.r.t. optimal models is NExpTime-complete. Concept satisfiability, instance checking, cautious and brave cost query answering w.r.t. optimal models are $\text{ExpTime}^{\text{NP}}$-complete.*

We can analogously define optimal models as those that *maximize* the value of F.

4. Discussion

Our logic has a form of concrete domains [2, 14, 3] combined to rich cardinality constraints, closely related but orthogonal to [15, 16]; the combination of these two types of numeric reasoning, despite being very natural, has not received much attention. Optimal models are close in spirit to *minimal models*, however, our formalism is orthogonal to *circumscribed DLs* [17]. The restrictions on feature annotations and numeric constraints are fundamental for our upper bounds. An interesting challenge is to relax them, e.g., by allowing negative values, but it is not straightforward. We also plan to explore multiobjective optimization and other types of comparison among models, e.g. *Pareto* or *lexicographic* orders. An intriguing direction is using our approach for inconsistency-tolerant reasoning [18], where inconsistencies can be detected with an 'error' cost to be minimized.

Acknowledgments

This work was partially supported by the Austrian Science Fund (FWF) projects PIN8884924, P30873 and 10.55776/COE12.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] F. Di Stefano, S. Lukumbuzya, M. Ortiz, M. Šimkus, Expressive Description Logics with Rich Yet Affordable Numeric Constraints, (To appear in) Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning (KR), 2025.

[2] F. Baader, P. Hanschke, A Scheme for Integrating Concrete Domains into Concept Languages, in: Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI), IJCAI'91, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1991, p. 452–457.

[3] S. Borgwardt, F. De Bortoli, P. Koopmann, The Precise Complexity of Reasoning in $\mathcal{ALC}$ with ω-Admissible Concrete Domains, in: L. Giordano, J. C. Jung, A. Ozaki (Eds.), Proceedings of the 37th International Workshop on Description Logics (DL), volume 3739 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3739/paper-1.pdf.

[4] C. Lutz, The Complexity of Description Logics with Concrete Domains, Ph.D. thesis, RWTH Aachen University, Germany, 2002. URL: http://sylvester.bth.rwth-aachen.de/dissertationen/2002/042/index.htm.

[5] N. Labai, M. Ortiz, M. Šimkus, An ExpTime Upper Bound for $\mathcal{ALC}$ with Integers, in: D. Calvanese, E. Erdem, M. Thielscher (Eds.), Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR), 2020, pp. 614–623. doi:10.24963/KR.2020/61.

[6] S. Demri, K. Quaas, First Steps Towards Taming Description Logics with Strings, in: S. A. Gaggl, M. V. Martinez, M. Ortiz (Eds.), Proceedings of the 18th European Conference on Logic in Artificial Intelligence (JELIA), volume 14281 of *Lecture Notes in Computer Science*, Springer, 2023, pp. 322–337. doi:10.1007/978-3-031-43619-2_23.

[7] V. Haarslev, R. Möller, M. Wessel, The Description Logic $\mathcal{ALCNH}_{\mathrm{R}+}$ Extended with Concrete Domains: A Practically Motivated Approach, in: R. Goré, A. Leitsch, T. Nipkow (Eds.), Proceedings of the 1st International Joint Conference on Automated Reasoning, IJCAR, volume 2083 of *Lecture Notes in Computer Science*, Springer, 2001, pp. 29–44. doi:10.1007/3-540-45744-5_4.

[8] J. Z. Pan, I. Horrocks, Reasoning in the $\mathcal{SHOQ}(D_{\mathrm{n}})$ Description Logic, in: I. Horrocks, S. Tessaris (Eds.), Proceedings of the 2002 International Workshop on Description Logics (DL), volume 53 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2002. URL: https://ceur-ws.org/Vol-53/Pan-Horrocks-shioqdn-2002.ps.

[9] S. Lukumbuzya, M. Ortiz, M. Šimkus, Datalog Rewritability and Data Complexity of $\mathcal{ALCHOIQ}$ with Closed Predicates, Artif. Intell. 330 (2024) 104099. doi:10.1016/J.ARTINT.2024.104099.
[10] F. Baader, I. Horrocks, C. Lutz, U. Sattler, An Introduction to Description Logic, Cambridge University Press, 2017.
[11] C. Lutz, I. Seylan, F. Wolter, Ontology-Based Data Access with Closed Predicates is Inherently Intractable(Sometimes), in: Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI), IJCAI/AAAI, 2013, pp. 1024–1030.
[12] I. Pratt-Hartmann, Complexity of the Two-Variable Fragment with Counting Quantifiers, Journal of Logic, Language and Information 14 (2005) 369–395. doi:10.1007/s10849-005-5791-1.
[13] C. Lutz, U. Sattler, L. Tendera, The Complexity of Finite Model Reasoning in Description Logics, Inf. Comput. 199 (2005) 132–171. doi:10.1016/j.ic.2004.11.002.
[14] C. Lutz, M. Milicic, A Tableau Algorithm for Description Logics with Concrete Domains and General TBoxes, J. Autom. Reason. 38 (2007) 227–259. doi:10.1007/S10817-006-9049-7.
[15] F. Baader, F. De Bortoli, Description logics that count, and what they can and cannot count (extended abstract), in: S. Borgwardt, T. Meyer (Eds.), Proceedings of the 33rd International Workshop on Description Logics (DL), co-located with the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR), volume 2663 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020. URL: https://ceur-ws.org/Vol-2663/abstract-2.pdf.
[16] B. Bednarczyk, O. Fiuk, Presburger Büchi Tree Automata with Applications to Logics with Expressive Counting, in: A. Ciabattoni, E. Pimentel, R. J. G. B. de Queiroz (Eds.), Proceedings of the 28th International Workshop on Logic, Language, Information, and Computation (WoLLIC), volume 13468 of *Lecture Notes in Computer Science*, Springer, 2022, pp. 295–308. doi:10.1007/978-3-031-15298-6_19.
[17] P. A. Bonatti, C. Lutz, F. Wolter, The Complexity of Circumscription in DLs, J. Artif. Intell. Res. 35 (2009) 717–773. URL: https://doi.org/10.1613/jair.2763. doi:10.1613/JAIR.2763.
[18] M. Bienvenu, C. Bourgaux, R. Jean, Cost-Based Semantics for Querying Inconsistent Weighted Knowledge Bases, in: P. Marquis, M. Ortiz, M. Pagnucco (Eds.), Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR, 2024. doi:10.24963/KR.2024/16.