Querying Circumscribed Description Logic Knowledge Bases (Extended Abstract)

Carsten Lutz[1], Quentin Manière[1,2] and Robin Nolte[3]

[1]Department of Computer Science, Leipzig University, Germany

[2]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany

[3]University of Bremen, Digital Media Lab, Germany

Abstract

We summarize our recent work on evaluating (unions of) conjunctive queries on circumscribed versions of description logic ranging from $\mathcal{ALCHIO}$ via $\mathcal{EL}$ to various versions of DL-Lite [1].

Keywords

Description Logics, Circumscription, Query answering, Complexity of reasoning

1. Introduction

While standard description logics (DLs), such as those underlying the OWL 2 ontology language, do not include non-monotonic features, it is generally acknowledged that extending DLs with such features is very useful. Examples of applications include ontological modeling in the biomedical domain [2, 3] and the formulation of access control policies [4]. Circumscription is one of the traditional AI approaches to non-monotonicity, and it provides an important way to define non-monotonic DLs. In contrast to other approaches, such as default rules, it does not require the adoption of strong syntactic restrictions to preserve decidability. DLs with circumscription are closely related to several other approaches to non-monotonic DLs, in particular to DLs with defeasible inclusions and typicality operators [5, 6, 7, 8, 9].

The main feature of circumscription is that selected predicate symbols can be minimized, that is, the extension of these predicates must be minimal regarding set inclusion. Other predicates may vary freely or be declared fixed. In addition, a preference order can be declared on the minimized predicates. The traditional AI use of circumscription is to introduce and minimize abnormality predicates, which makes it possible to formulate defeasible implications.

Circumscription is also closely related to the closure of predicates symbols as studied, for instance, in [10, 11, 12]. While DLs usually assume open-world semantics and represent incomplete knowledge, such closed predicates are interpreted under a closed-world assumption, reflecting that complete knowledge is available regarding those predicates. Circumscription may

DL 2023: 36th International Workshop on Description Logics, September 2–4, 2023, Rhodes, Greece

clu@informatik.uni-leipzig.de (C. Lutz); quentin.maniere@uni-leipzig.de (Q. Manière); nolte@uni-bremen.de (R. Nolte)

home.uni-leipzig.de/clu (C. Lutz); home.uni-leipzig.de/qmaniere (Q. Manière); www.uni-bremen.de/dmlab/team/robin-nolte (R. Nolte)

0000-0002-8791-6702 (C. Lutz); 0000-0001-9618-8359 (Q. Manière); 0009-0004-2975-6378 (R. Nolte)

	$\mathcal{EL}$, $\mathcal{ALCHIO}$	DL-Lite$^{\mathcal{H}}_{\mathsf{core}}$, DL-Lite$^{\mathcal{H}}_{\mathsf{bool}}$	DL-Lite$_{\mathsf{bool}}$	DL-Lite$_{\mathsf{core}}$, DL-Lite$_{\mathsf{horn}}$
Combined	2Exp-c.	2Exp-c.$^{(\dagger)}$	coNExp-c.	coNExp-c.$^{(\dagger)}$
Data	Π^p_2-c.	coNP-c.	coNP-c.	coNP-c.

Table 1
Complexity of (U)CQ evaluation on circumscribed KBs. $\cdot^{(\dagger)}$ indicates the lower bound relies on UCQs.

be viewed as a soft form of closing concept names: there are no other instances of a minimized concept name except the explicitly asserted ones unless we are forced to introduce (a minimal set of) additional instances to avoid inconsistency.

A primary application of DLs is ontology-mediated querying, where an ontology is used to enrich data with domain knowledge. Surprisingly, relatively little is known about ontology-mediated querying with DLs that support circumscription. The most popular choice of queries are conjunctive queries (CQs) and unions thereof (UCQs), and to the best of our knowledge, in this case not even decidability is known. In [1], we aim to close this gap and study the decidability and precise complexity of ontology-mediated querying for DLs with circumscription, both w.r.t. combined complexity and data complexity. We consider the expressive DL $\mathcal{ALCHIO}$, the tractable (without circumscription) DL $\mathcal{EL}$, and several DL-Lite family members tailored specifically towards ontology-mediated querying.

2. Contributions

One of our main results is that UCQ evaluation is decidable in all these DLs when circumscription is added, a summary of the complexities can be found in Table 1. It is 2Exp-complete in $\mathcal{ALCHIO}$ w.r.t. combined complexity, and thus not harder than query evaluation without circumscription. W.r.t. data complexity, there is a significant increase from coNP- to Π^p_2-completeness. For $\mathcal{EL}$, both combined and data complexity turns out to be identical to that of $\mathcal{ALCHIO}$, which improves lower bounds from [5]. All these lower bounds already hold for CQs. Remarkably, the Π^p_2 lower bound for data complexity already holds when there is only a single minimized concept name and without fixed predicates. The complexities for DL-Lite are lower, though still high. Evaluation is 'only' coNP-complete w.r.t. data complexity. The combined complexity remains at 2Exp with role inclusions and drops to coNExp without them. The lower bounds already apply to very basic positive versions of DL-Lite that do not provide concept disjointness constraints, and the upper bounds to expressive versions that include all Boolean operators.

We also study the evaluation of the basic yet important atomic queries (AQs), conjunctive queries of the form $A(x)$ with A a concept name. Also here, we obtain a rather complete picture of the complexity landscape. It is known from [13] that AQ evaluation in $\mathcal{ALCHIO}$ is coNExp$^{\mathrm{NP}}$-complete w.r.t. combined complexity. We show that the lower bound holds already for $\mathcal{EL}$. Moreover, our Π^p_2-lower bound for the data complexity of (U)CQ-evaluation in $\mathcal{EL}$ mentioned above only requires an AQ, and thus AQ evaluation in both $\mathcal{ALCHIO}$ and $\mathcal{EL}$ are Π^p_2-complete w.r.t. data complexity. For DL-Lite, the data complexity drops to PTime in all considered versions, and the combined complexity ranges from coNExp- to Π^p_2-complete, depending on which Boolean operators are admitted. A summary can be found in Table 2.

	$\mathcal{EL}$, $\mathcal{ALCHIO}$	DL-Lite$_{\mathsf{bool}}$, DL-Lite$^{\mathcal{H}}_{\mathsf{bool}}$	DL-Lite$_{\mathsf{core}}$, DL-Lite$^{\mathcal{H}}_{\mathsf{horn}}$
Combined	coNExp$^{\text{NP}}$-c.$^{(\dagger)}$	coNExp-c.	Π^{p}_2-c.$^{(\ddagger)}$
Data	Π^{p}_2-c.	in PTime	in PTime

Table 2
Complexity of AQ evaluation on circumscribed KBs. $^{(\dagger)}$: completeness already known for $\mathcal{ALC}(\mathcal{IO})$. $^{(\ddagger)}$: hardness already known.

3. Manipulating models of circumscribed KBs

We highlight a key-ingredient that underlies the obtained decidability results and upper complexity bounds. Recall that an $\mathcal{ALCIO}$ *concept* C is built according to the rule $C, D ::= \top \mid A \mid \{a\} \mid \neg C \mid C \sqcap D \mid \exists r.D$ where A ranges over *concept names*, a over *individual names*, and r over (possibly inverse) roles. An $\mathcal{ALCHIO}$ *knowledge base (KB)* takes the form $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ with $\mathcal{T}$ and $\mathcal{A}$ being an $\mathcal{ALCHIO}$ *TBox* and *ABox*, defined in the standard way. The set of individual names used in $\mathcal{A}$ is denoted $\mathsf{ind}(\mathcal{A})$. The semantics is defined as usual in terms of *interpretations*, we refer to [14] for full details.

A *circumscription pattern* is a tuple $\mathsf{CP} = (\prec, \mathsf{M}, \mathsf{F}, \mathsf{V})$, where $\prec$ is a strict partial order on M called the *preference relation*, and M, F and V are a partition of the set of all concept names. The elements of M, F and V are the *minimized*, *fixed* and *varying* concept names. Role names always vary to avoid undecidability [13]. The preference relation $\prec$ on M induces a preference relation $<_{\mathsf{CP}}$ on interpretations by setting $\mathcal{J} <_{\mathsf{CP}} \mathcal{I}$ if the following conditions hold:

1. $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}}$,
2. for all $A \in \mathsf{F}$, $A^{\mathcal{J}} = A^{\mathcal{I}}$,
3. for all $A \in \mathsf{M}$ with $A^{\mathcal{J}} \not\subseteq A^{\mathcal{I}}$, there is a $B \in \mathsf{M}$, $B \prec A$, such that $B^{\mathcal{J}} \subsetneq B^{\mathcal{I}}$,
4. there exists an $A \in \mathsf{M}$ such that $A^{\mathcal{J}} \subsetneq A^{\mathcal{I}}$ and for all $B \in \mathsf{M}$, $B \prec A$ implies $B^{\mathcal{J}} = B^{\mathcal{I}}$.

A *circumscribed KB* takes the form $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ where $\mathcal{K}$ is a KB and CP a circumscription pattern. A model $\mathcal{I}$ of $\mathcal{K}$ is a *model* of $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ if no $\mathcal{J} <_{\mathsf{CP}} \mathcal{I}$ is a model of $\mathcal{K}$. This minimality condition poses a challenge for algorithms that try to find a *countermodel* $\mathcal{I}$ for a query $q(\bar{x})$ and tuple $\bar{a}$, that, is a model $\mathcal{I}$ of $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ with $\mathcal{I} \not\models q(\bar{a})$. Indeed, such algorithms rely on first establishing a 'regular model property' which states that if a countermodel exists, then there is a 'regular' one, typically tree-shaped or forest-shaped. This is proved by starting with any countermodel and then manipulating it, e.g. by unraveling. In the presence of circumscription, such manipulations are more challenging as they must preserve minimality w.r.t. the circumscription pattern. In what follows, we present a condition that is sufficient for the preservation of minimality and underlies several of our constructions. It is stated as Lemma 1 below.

We first observe that a nominal may be viewed as a (strictly) closed concept name with a single instance. From this, we exhibit a reduction from UCQ evaluation on circumscribed $\mathcal{ALCHIO}$ KBs to UCQ evaluation on circumscribed $\mathcal{ALCHI}$ KBs. We are thus left with $\mathcal{ALCHI}$ TBoxes, which we generally assume to be in *normal form*, meaning that every concept inclusion has one of the following shapes:

$$\top \sqsubseteq A \qquad A \sqsubseteq \exists r.B \qquad \exists r.B \sqsubseteq A \qquad A_1 \sqcap A_2 \sqsubseteq A \qquad A \sqsubseteq \neg B \qquad \neg B \sqsubseteq A$$

where A, A_1, A_2, B range over concept names and r ranges over (possibly inverse) roles.

Let $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ be a circumscribed KB with $\mathcal{K} = (\mathcal{T}, \mathcal{A})$. The set of concept names in $\mathcal{T}$ is denoted $\mathsf{N_C}(\mathcal{T})$. We define a *type* to be a set of concept names $t \subseteq \mathsf{N_C}(\mathcal{T})$. For an interpretation $\mathcal{I}$ and $d \in \Delta^\mathcal{I}$, let $\mathsf{tp}_\mathcal{I}(d) := \{A \in \mathsf{N_C}(\mathcal{T}) \mid d \in A^\mathcal{I}\}$, and for a subset $\Delta \subseteq \Delta^\mathcal{I}$, let $\mathsf{tp}_\mathcal{I}(\Delta) = \{\mathsf{tp}_\mathcal{I}(d) \mid d \in \Delta\}$. We further write $\mathsf{TP}(\mathcal{I})$ for $\mathsf{tp}_\mathcal{I}(\Delta^\mathcal{I})$. Finally, we set

$$\mathsf{TP}(\mathcal{T}) := \bigcup_{\mathcal{I} \text{ model of } \mathcal{T}} \mathsf{TP}(\mathcal{I}).$$

We next show how to identify a 'core' part of a model $\mathcal{I}$ of $\mathcal{K}$. These core parts play an important role in dealing with circumscription in our upper bound proofs.

Definition 1. *Let $\mathcal{I}$ be a model of $\mathcal{K}$. We use $\mathsf{TP}_{\mathsf{core}}(\mathcal{I})$ to denote the set of all types $t \in \mathsf{TP}(\mathcal{I})$ such that $|\{d \in \Delta^\mathcal{I} \setminus \mathsf{ind}(\mathcal{A}) \mid \mathsf{tp}_\mathcal{I}(d) = t\}| < |\mathsf{TP}(\mathcal{T})|$. We set $\mathsf{TP}_{\overline{\mathsf{core}}}(\mathcal{I}) = \mathsf{TP}(\mathcal{I}) \setminus \mathsf{TP}_{\mathsf{core}}(\mathcal{I})$ and $\Delta^\mathcal{I}_{\mathsf{core}} = \{d \in \Delta^\mathcal{I} \mid \mathsf{tp}_\mathcal{I}(d) \in \mathsf{TP}_{\mathsf{core}}(\mathcal{I})\}$.*

So the core of $\mathcal{I}$, which is the restriction of $\mathcal{I}$ to domain $\Delta^\mathcal{I}_{\mathsf{core}}$, consists of all elements whose types are realized not too often, except possibly in the ABox. A good way of thinking about cores is that if a model $\mathcal{I}$ of $\mathcal{K}$ is minimal w.r.t. $<_{\mathsf{CP}}$, then all instances of minimized concept names are in the core. This is, however not strictly true since we may have $A \sqsubseteq B$ where A is $\top$ or fixed, and B is minimized.

The following crucial lemma provides a sufficient condition for a model $\mathcal{J}$ of $\mathcal{K}$ to be minimal w.r.t. $<_{\mathsf{CP}}$, relative to a model $\mathcal{I}$ of $\mathcal{K}$ that is known to be minimal w.r.t. $<_{\mathsf{CP}}$.

Lemma 1. *Let $\mathcal{I}$ be a model of $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ and $\mathcal{J}$ a model of $\mathcal{K}$ with $\Delta^\mathcal{I}_{\mathsf{core}} \subseteq \Delta^\mathcal{J}$. If $\mathsf{tp}_\mathcal{I}(d) = \mathsf{tp}_\mathcal{J}(d)$ for all $d \in \Delta^\mathcal{I}_{\mathsf{core}}$ and $\mathsf{tp}_\mathcal{J}(\Delta^\mathcal{J} \setminus \Delta^\mathcal{I}_{\mathsf{core}}) = \mathsf{TP}_{\overline{\mathsf{core}}}(\mathcal{I})$, then $\mathcal{J}$ is a model of $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$.*

We use the above lemma to show that if there exists a countermodel for a CQ $q(\bar{x})$ and tuple $\bar{a} \in \mathsf{ind}(\mathcal{A})^{|\bar{x}|}$ on $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$, then this is witnessed by a countermodel $\mathcal{I}$ that has a regular shape. By regular shape, we mean that there is a 'base part' that contains the ABox, the core of $\mathcal{I}$, as well as some additional elements as representatives for certain types; all other parts of $\mathcal{I}$ are tree-shaped with their root in the base part, and potentially with edges that go back to the core (but not to other parts of the base). We then show that the existence of a countermodel of such a shape can be decided using a mosaic procedure. The base part $\mathcal{I}_{\mathsf{base}}$ of $\mathcal{I}$ is part of all mosaics (intuitively, it is guessed) and the mosaics are used to ensure that $\mathcal{I}_{\mathsf{base}}$ can be extended into a complete countermodel by adding tree-shaped parts. We trace partial query matches through the mosaics to make sure that the query doesn't map into $\mathcal{I}$. Compliance with $\mathsf{Circ}_{\mathsf{CP}}(\mathcal{K})$ can be checked solely on $\mathcal{I}_{\mathsf{base}}$, using Lemma 1. This gives the 2Exp upper bound w.r.t. combined complexity. Proving the Π^P_2 upper bound in data complexity requires extra work. We add a quotient construction that exploits the above regular shape, and preserves the base part so that Lemma 1 may again guarantee minimality.

To establish the improved bounds for the DL-Lite family, we refine Lemma 1. To prove that AQ evaluation on circumscribed DL-Lite$^\mathcal{H}_{\mathsf{bool}}$ KBs is in PTime in data complexity, for example, we improve Lemma 1 by identifying a core part of the countermodel that lives purely within the ABox. This refinement is made possible by the limited expressivity of DL-Lite.

Acknowledgments

The research reported in this paper has been supported by the German Research Foundation DFG, as part of Collaborative Research Center (Sonderforschungsbereich) 1320 Project-ID 329551904 "EASE - Everyday Activity Science and Engineering", University of Bremen (http://www.ease-crc.org/). The research was conducted in subprojects "P02 – Ontologies with Abstraction" and "P05-N – Principles of Metareasoning for Everyday Activities".

The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany and by the Sächsische Staatsministerium für Wissenschaft Kultur und Tourismus in the program Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification number: ScaDS.AI

This work is partly supported by BMBF (Federal Ministry of Education and Research) in DAAD project 57616814 (SECAI, School of Embedded Composite AI) as part of the program Konrad Zuse Schools of Excellence in Artificial Intelligence.

References

[1] C. Lutz, Q. Manière, R. Nolte, Querying Circumscribed Description Logics Knowledge Bases, in: Proc. of KR, 2023, p. (To appear).

[2] A. L. Rector, Defaults, context, and knowledge: Alternatives for OWL-indexed knowledge bases, in: Proc. of the Pac. Symp. in Biocomp., World Scientific, 2004, pp. 226–237.

[3] R. Stevens, M. E. Aranguren, K. Wolstencroft, U. Sattler, N. Drummond, M. Horridge, A. L. Rector, Using OWL to model biological knowledge, Int. J. Hum. Comput. Stud. 65 (2007) 583–594.

[4] P. A. Bonatti, P. Samarati, Logics for authorization and security, in: Logics for Emerging Applications of Databases, Springer, 2003, pp. 277–323.

[5] P. A. Bonatti, M. Faella, L. Sauro, Defeasible inclusions in low-complexity DLs, J. Artif. Intell. Res. 42 (2011) 719–764.

[6] G. Casini, U. Straccia, Defeasible inheritance-based description logics, J. Artif. Intell. Res. 48 (2013) 415–473.

[7] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, A non-monotonic description logic for reasoning about typicality, Artif. Intell. 195 (2013) 165–202.

[8] P. A. Bonatti, M. Faella, I. M. Petrova, L. Sauro, A new semantics for overriding in description logics, Artif. Intell. 222 (2015) 1–48.

[9] M. Pensel, A. Turhan, Reasoning in the defeasible description logic $\mathcal{EL}$ - computing standard inferences under rational and relevant semantics, Int. J. Approx. Reason. 103 (2018) 28–70.

[10] C. Lutz, I. Seylan, F. Wolter, Ontology-based data access with closed predicates is inherently intractable (sometimes), in: Proc. of IJCAI, IJCAI/AAAI, 2013, pp. 1024–1030.

[11] N. Ngo, M. Ortiz, M. Simkus, Closed predicates in description logics: Results on combined complexity, in: Proc. of KR, AAAI Press, 2016, pp. 237–246.

[12] C. Lutz, I. Seylan, F. Wolter, The data complexity of ontology-mediated queries with closed predicates, Log. Methods Comput. Sci. 15 (2019).

[13] P. A. Bonatti, C. Lutz, F. Wolter, The complexity of circumscription in description logic, J. Artif. Intell. Res. 35 (2009) 717–773.
[14] F. Baader, I. Horrocks, C. Lutz, U. Sattler, An Introduction to Description Logic, Cambridge University Press, 2017.