Maximum Entropy Reasoning via Model Counting in (Description) Logics that Count (Extended Abstract)

Franz Baader[1,2], Anton Claußnitzer[1,2]

[1] TU Dresden, Institute of Theoretical Computer Science, Dresden, Germany

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

Abstract

This extended abstract reports on work that was published in the proceedings of FLAIRS-38. In previous work it was shown that the logic $\mathcal{ALC}^{\mathsf{ME}}$, which extends the description logic (DL) $\mathcal{ALC}$ with probabilistic conditionals, has domain-lifted inference. In the FLAIRS-38 paper, we extend this result from the base logic $\mathcal{ALC}$ to two logics that can count, the two-variable fragment C^2 of first-order logic (FOL) with counting quantifiers, and the DL $\mathcal{ALCSCC}$, which can formulate expressive counting constraints on role successors and is not a fragment of FOL. As an auxiliary result, we prove that model counting in $\mathcal{ALCSCC}$ can be realized in a domain-liftable way.

Keywords

Probabilistic Conditionals, Model Counting, Domain Liftability, Counting Quantifiers

1. Introduction

Description logics (DLs) [1, 2] are a well-investigated family of logic-based knowledge representation formalisms, which can be used to formalize the terminological knowledge of an application domain in a machine-processable way. For instance, large medical ontologies such as SNOMED CT[1] and Galen[2] have been developed using an appropriate DL. While classical DLs are often sufficient for formalizing certain knowledge like the definition of medical terminology, they cannot adequately express uncertain knowledge, which may, e.g., be needed for medical diagnosis. Using a non-medical example, the concept of a father can be formalized by the concept inclusion (CI) $\mathit{Father} \sqsubseteq \mathit{Human} \sqcap \mathit{Male} \sqcap \exists \mathit{child}.\mathit{Human}$, which says that fathers are male humans that have a human child. However, a statement like "Rich persons usually have rich children" should not be expressed with a CI since it does not hold for all rich persons. It is more appropriate to use a probabilistic conditional (PC) of the form $(\forall \mathit{child}.\mathit{Rich} \mid \mathit{Person} \sqcap \mathit{Rich})[p]$, where the probability p may be based on statistical knowledge or express the degree of a subjective belief. The CI and PC of our example can be phrased in the probabilistic DL $\mathcal{ALC}^{\mathsf{ME}}$ [3, 4, 5], which extends the well-known DL $\mathcal{ALC}$ with probabilistic conditionals that are interpreted

DL 2025: 38th International Workshop on Description Logics, September 2–4, 2025, Opole, Poland

franz.baader@tu-dresden.de (F. Baader); anton.claussnitzer1@tu-dresden.de (A. Claußnitzer)

https://lat.inf.tu-dresden.de/~baader (F. Baader); https://lat.inf.tu-dresden.de/claussnitzer (A. Claußnitzer)

0000-0002-4049-221X (F. Baader); 0009-0004-5103-8452 (A. Claußnitzer)

CEUR Workshop Proceedings (CEUR-WS.org)

[1]https://www.snomed.org/

[2]https://bioportal.bioontology.org/ontologies/GALEN

based on the aggregating semantics and the maximum entropy principle. Compared to other probabilistic extensions of DL (such as [6, 7, 8]), $\mathcal{ALC}^{\mathsf{ME}}$ has the advantage that the aggregating semantics smoothly combines the statistical and the subjective view on probabilities and that the maximum entropy approach fulfills a number of reasonable commonsense principles [9, 10, 11]. Like other approaches for probabilistic reasoning in a first-order setting, the aggregating semantics assumes that interpretations are built over a fixed finite domain Δ. To be able to deal with large domain sizes, one needs reasoning to be domain-lifted [12], which means that inferences can be drawn in time polynomial in the size of Δ. The main results of [4, 3, 5] are that $\mathcal{ALC}^{\mathsf{ME}}$ allows for domain-lifted inference.

In the FLAIRS-38 paper [13], we extend these results from the base logic $\mathcal{ALC}$ to logics that can count. Number restrictions [14, 15] are DL concept constructors that can express simple numerical constraints on the number of role successors of an individual, such as that it has three children that are rich, and only two that are not rich, whereas cardinality restrictions on concepts [16, 17] can constrain the overall number of elements of a concept, e.g., expressing that there are more than 500,000 rich people living in Florida. Description logics offering such counting features are contained in C^2, the two-variable fragment of FOL with counting quantifiers, and are thus decidable [18, 19]. In [20, 21] it was recently shown that (extended versions of) model counting in C^2 can be realized in a domain-liftable way. We use this result in [13] to prove that ${\mathsf{C}^2}^{\mathsf{ME}}$ allows for domain-lifted inference. The DL $\mathcal{ALCSCC}$ [22] offers more expressive counting constraints on role successors, which in general cannot be expressed in C^2 or even full FOL [23]. For example, in $\mathcal{ALCSCC}$ we can describe persons that have more rich than non-rich children without specifying how many children of each type the person actually has, and in $\mathcal{ALCSCC}^{\mathsf{ME}}$ we can say that, with a high probability (say .8), rich persons have more rich than non-rich children. We show in [13] that (an extended version of) model counting in $\mathcal{ALCSCC}$ can be realized in a domain-liftable way, and use this result to prove that $\mathcal{ALCSCC}^{\mathsf{ME}}$ allows for domain-lifted inference.

In the following, we briefly sketch the main results obtained in [13]. More details can be found in the full paper [13].

2. Concept-Constrained Model Counting

Model counting usually asks how many models over a given finite domain Δ a given sentence has. In [13], we consider a slightly extended version of this task, called *concept-constrained model counting*, where the underlying logic is either C^2 or $\mathcal{ALCSCC}$. Concepts C of $\mathcal{ALCSCC}$ and their extensions C^I as well as $\mathcal{ALCSCC}$ TBoxes and their models are defined in [22]. For the two-variable fragment C^2 of FOL with counting quantifiers, concepts C are formulas with one free variable x, and their extension C^I consists of those elements of I that make the formula true when substituted for x. A C^2 TBox is a sentence (i.e., formula without free variables) of C^2.

Let $\mathcal{T}$ be a TBox, $C_1, \dots, C_n$ concepts, $c_1, \dots, c_\ell$ non-negative integers, and Δ a finite set. Then

$$\mathrm{ccmc}(\mathcal{T}, C_1, \dots, C_\ell, c_1, \dots, c_\ell, \Delta)$$

is defined to be the number of models I of $\mathcal{T}$ with domain Δ that satisfy $|C_i^I| = c_i$ $(1 \leq i \leq \ell)$. We say that concept-constrained model counting is *domain-liftable* if this number can be computed in polynomial time in the size of the input Δ (i.e., where the other inputs of the function ccmc are assumed to be of constant size).

Theorem 1 ([13]). *Concept-constrained model counting in C^2 and in $\mathcal{ALCSCC}$ is domain-liftable.*

For C^2, this is an easy consequence of the results on model counting in C^2 in [20] (Proposition 4 together with Theorem 4). For $\mathcal{ALCSCC}$, this is explicitly proved in [13], and constitutes one of the main results of this paper.

3. The Logics $\mathcal{ALCSCC}^{\mathsf{ME}}$ and $\mathsf{C}^{2^{\mathsf{ME}}}$

In the following, let $\mathcal{L}$ be either $\mathcal{ALCSCC}$ or C^2. In the logic $\mathcal{L}^{\mathsf{ME}}$, we consider *probabilistic conditionals (PCs)* of the form $(D \mid C)[p]$, where C, D are $\mathcal{L}$ concepts and p is a rational number. An $\mathcal{L}$ *knowledge base* $\mathcal{K} = (\mathcal{T}, \mathcal{C})$ consists of an $\mathcal{L}$ TBox $\mathcal{T}$ together with a finite set $\mathcal{C}$ of PCs. To define the semantics of such a knowledge base $\mathcal{K}$, we follow [3, 4] and consider interpretations over a fixed, finite domain Δ of the signature of $\mathcal{K}$. We denote the (finite) set of all these interpretations with $\mathcal{I}^\Delta$ and the set of probability distributions $P\colon \mathcal{I}^\Delta \to [0, 1]$ over $\mathcal{I}^\Delta$ with $\mathfrak{P}^\Delta$. The distribution $P \in \mathfrak{P}^\Delta$ is a *model* of $\mathcal{K} = (\mathcal{T}, \mathcal{C})$ if all interpretations I that are not models of $\mathcal{T}$ satisfy $P(I) = 0$ and the following holds for all PCs $(F_i \mid E_i)[p_i]$ in $\mathcal{C}$: $\sum_{I\in\mathcal{I}^\Delta} P(I) \cdot |E_i^I| > 0$ and

$$\sum_{I\in\mathcal{I}^\Delta} P(I) \cdot |E_i^I \cap F_i^I| = p_i \cdot \sum_{I\in\mathcal{I}^\Delta} P(I) \cdot |E_i^I|. \tag{1}$$

This semantics for PCs is called *aggregating semantics* [10]. A knowledge base with at least one model is *consistent*. Using the fact that concept-constrained model counting for $\mathcal{L}$ is domain-liftable (see Theorem 1), consistency checking for $\mathcal{L}^{\mathsf{ME}}$ is shown to be also domain-liftable in [13].

Theorem 2 ([13]). *Consistency of an $\mathcal{L}^{\mathsf{ME}}$ knowledge base $\mathcal{K}$ for a finite domain Δ can be checked in time polynomial in $|\Delta|$.*

Instead of reasoning w.r.t. all models of a consistent knowledge base, we use the maximum entropy distribution as preferred model. In fact, as pointed out in [3], according to Paris, this distribution is the most appropriate choice of model in this setting. The entropy of a probability distribution P is $-\sum_{I\in\mathcal{I}^\Delta} P(I) \cdot \log_2 P(I)$, where we use the convention $0 \cdot \infty = 0$. For every consistent knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{C})$, there is exactly one model of $\mathcal{K}$ with maximal entropy, i.e., the optimization problem

$$\begin{aligned}
&-\sum_{I\in\mathcal{I}^\Delta} P(I) \cdot \log_2 P(I) \overset{!}{=} \max \text{ with the conditions}\\
&\sum_{I\in\mathcal{I}^\Delta} P(I) = 1, \quad \sum_{I\in\mathcal{I}^\Delta} P(I)|E^I| > 0 \text{ for } (F|E)[p] \in \mathcal{C},\\
&\sum_{I\in\mathcal{I}^\Delta} P(I)|E^I \cap F^I| = p \sum_{I\in\mathcal{I}^\Delta} P(I)|E^I| \text{ for } (F|E)[p] \in \mathcal{C},
\end{aligned}$$

$$\forall I \in \mathcal{I}^\Delta\colon P(I) \geq 0 \text{ and } P(I) = 0 \text{ if } I \not\models \mathcal{T},$$

has exactly one solution $P^{\mathsf{ME}}_{\mathcal{K}}$ [10].

Instead of solving this optimization problem directly, one usually considers the dual optimization problem, whose solutions represent $P^{\mathsf{ME}}_{\mathcal{K}}$ in a compact way. Assume that $\mathcal{C} = \{(F_i \mid E_i)[p_i] \mid i = 1, \dots, n\}$ and define the functions f_i $(1 \leq i \leq n)$ as $f_i(I) := |E_i^I \cap F_i^I| - p_i|E_i^I|$. An application of the Lagrange multiplier method to the above optimization problem then yields $P^{\mathsf{ME}}_{\mathcal{K}}(I) = 0$ if $I \not\models \mathcal{T}$ and $P^{\mathsf{ME}}_{\mathcal{K}}(I) = \alpha_0 \alpha_1^{f_1(I)} \cdots \alpha_n^{f_n(I)}$ if $I \models \mathcal{T}$, where the values $\alpha_i > 0$ are solutions to the equations $\sum_{I \in \mathcal{I}^\Delta, I \models \mathcal{T}} f_i(I) \alpha_1^{f_1(I)} \cdots \alpha_n^{f_n(I)} = 0$, $i = 1, \dots, n$, and $\alpha_0 = \left(\sum_{I \in \mathcal{I}^\Delta, I \models \mathcal{T}} \alpha_1^{f_1(I)} \cdots \alpha_n^{f_n(I)}\right)^{-1}$ is a normalization constant.

Since the numbers α_i are solutions of a non-linear optimization problem, they can in general only be approximated (e.g., using Newton's method). Following [3], we do not investigate this approximation process here, but assume that a rational approximation $\boldsymbol{\beta} \in \mathbb{Q}^n_{>0}$ of the exact solution $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n_{>0}$ is given. For such an approximation $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)$, the induced probability distribution $P^{\boldsymbol{\beta}}_{\mathcal{K}}$ on $\mathcal{I}^\Delta$ is defined in [3] as

$$P^{\boldsymbol{\beta}}_{\mathcal{K}}(I) = \begin{cases} \beta_0 \beta_1^{f_1(I)} \cdots \beta_n^{f_n(I)} & \text{if } I \models \mathcal{T}, \\ 0 & \text{else,} \end{cases}$$

where the normalization constant β_0 is defined analogously to α_0.

It is shown in [13] that domain-lifted inference w.r.t. $P^{\boldsymbol{\beta}}_{\mathcal{K}}$ is possible. The main step towards achieving this result is the following theorem.

Theorem 3. *Let E, F be $\mathcal{L}$ concepts, $\mathcal{K} = (\mathcal{T}, \mathcal{C})$ with $\mathcal{C} = \{(D_i \mid C_i)[p_i] \mid 1 \leq i \leq n\}$ a consistent $\mathcal{L}$ knowledge base where $p_i = s_i/t_i$ for natural numbers s_i, t_i, and let $P^{\boldsymbol{\beta}}_{\mathcal{K}}$ be an approximation of the maximum entropy distribution, as defined above. Then we can compute (in time polynomial in $|\Delta|$) a polynomial $P(X_1, \dots, X_n)$ in n indeterminates and with rational coefficients such that $p := P(\sqrt[t_1]{\beta_1}, \dots, \sqrt[t_n]{\beta_n})$ satisfies $P^{\boldsymbol{\beta}}_{\mathcal{K}} \models (F \mid E)[p]$.*

Employing results from the theory of algebraic field extensions [24, 25], this theorem is used in [13] to show the following domain-liftability result.

Corollary 1. *Let E, F be $\mathcal{L}$ concepts, $q \in [0, 1]$ a rational number, $\mathcal{K} = (\mathcal{T}, \mathcal{C})$ a consistent $\mathcal{L}$ knowledge base, and $P^{\boldsymbol{\beta}}_{\mathcal{K}}$ a rational approximation of the maximum entropy distribution. Then $P^{\boldsymbol{\beta}}_{\mathcal{K}} \models (F \mid E)[q]$ and $P^{\boldsymbol{\beta}}_{\mathcal{K}} \models E \sqsubseteq F$ can be decided in time polynomial in $|\Delta|$.*

As pointed out in [13], it would also be interesting to know, for a given rational number q, whether q is larger or smaller than the probability p for which $P^{\boldsymbol{\beta}}_{\mathcal{K}} \models (F \mid E)[p]$ holds. At the point of submitting the final version of [13], we were able to show that this problem is decidable (see [26]), but it was not clear to us whether deciding the problem can be done in time polynomial in $|\Delta|$. More recently, we were able to show domain-liftability also for this problem.

Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (BMBF, SCADS22B) and the Saxon State Ministry for Science, Culture and Tourism (SMWK) by funding the competence center for Big Data and AI "ScaDS.AI Dresden/Leipzig".

References

[1] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, P. F. Patel-Schneider (Eds.), The Description Logic Handbook: Theory, Implementation, and Applications, Cambridge University Press, 2003.

[2] F. Baader, I. Horrocks, C. Lutz, U. Sattler, An Introduction to Description Logic, Cambridge University Press, 2017. doi:10.1017/9781139025355.

[3] M. Wilhelm, G. Kern-Isberner, A. Ecke, F. Baader, Counting strategies for the probabilistic description logic $\mathcal{ALC}^{\mathsf{ME}}$ under the principle of maximum entropy, in: F. Calimeri, N. Leone, M. Manna (Eds.), Logics in Artificial Intelligence – 16th European Conference, JELIA 2019, Proceedings, volume 11468 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 434–449. doi:10.1007/978-3-030-19570-0_28.

[4] F. Baader, A. Ecke, G. Kern-Isberner, M. Wilhelm, The complexity of the consistency problem in the probabilistic description logic $\mathcal{ALC}^{\mathsf{ME}}$, in: A. Herzig, A. Popescu (Eds.), Frontiers of Combining Systems – 12th International Symposium, FroCoS 2019, Proceedings, volume 11715 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 167–184. doi:10.1007/978-3-030-29007-8_10.

[5] M. Wilhelm, G. Kern-Isberner, Maximum entropy calculations for the probabilistic description logic $\mathcal{ALC}^{\mathsf{ME}}$, in: C. Lutz, U. Sattler, C. Tinelli, A. Turhan, F. Wolter (Eds.), Description Logic, Theory Combination, and All That – Essays Dedicated to Franz Baader on the Occasion of His 60th Birthday, volume 11560 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 588–609. doi:10.1007/978-3-030-22102-7_28.

[6] T. Lukasiewicz, Expressive probabilistic description logics, Artif. Intell. 172 (2008) 852–883. doi:10.1016/J.ARTINT.2007.10.017.

[7] R. Peñaloza, N. Potyka, Towards statistical reasoning in description logics over finite domains, in: Proceedings of the 11th International Conference on Scalable Uncertainty Management (SUM), volume 10564 of *Lecture Notes in Computer Science*, Springer, 2017, pp. 280–294.

[8] V. Gutiérrez-Basulto, J. C. Jung, C. Lutz, L. Schröder, Probabilistic description logics for subjective uncertainty, J. Artif. Intell. Res. 58 (2017) 1–66. doi:10.1613/JAIR.5222.

[9] J. B. Paris, Common sense and maximum entropy, Synthese 117 (1999) 75–93.

[10] G. Kern-Isberner, M. Thimm, Novel semantical approaches to relational probabilistic conditionals, in: F. Lin, U. Sattler, M. Truszczynski (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, AAAI Press, 2010.

[11] C. Beierle, M. Finthammer, G. Kern-Isberner, Relational probabilistic conditionals and

their instantiations under maximum entropy semantics for first-order knowledge bases, Entropy 17 (2015) 852–865. doi:10.3390/E17020852.
[12] G. Van den Broeck, N. Taghipour, W. Meert, J. Davis, L. D. Raedt, Lifted probabilistic inference by first-order knowledge compilation, in: T. Walsh (Ed.), IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI/AAAI, 2011, pp. 2178–2185. doi:10.5591/978-1-57735-516-8/IJCAI11-363.
[13] F. Baader, A. Claußnitzer, Maximum entropy reasoning via model counting in (description) logics that count, in: The International FLAIRS Conference, Proceedings, 38(1), 2025. doi:https://doi.org/10.32473/flairs.38.1.138854, an extended version of this paper is available as a technical report [26].
[14] B. Hollunder, W. Nutt, M. Schmidt-Schauß, Subsumption algorithms for concept description languages, in: 9th European Conference on Artificial Intelligence, ECAI 1990, 1990, pp. 348–353.
[15] B. Hollunder, F. Baader, Qualifying number restrictions in concept languages, in: J. F. Allen, R. Fikes, E. Sandewall (Eds.), Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91), Morgan Kaufmann, 1991, pp. 335–346.
[16] F. Baader, M. Buchheit, B. Hollunder, Cardinality restrictions on concepts, Artif. Intell. 88 (1996) 195–213. doi:10.1016/S0004-3702(96)00010-0.
[17] S. Tobies, The complexity of reasoning with cardinality restrictions and nominals in expressive description logics, J. Artif. Intell. Res. 12 (2000) 199–217. doi:10.1613/JAIR.705.
[18] E. Grädel, M. Otto, E. Rosen, Two-variable logic with counting is decidable, in: Proceedings, 12th Annual IEEE Symposium on Logic in Computer Science, IEEE Computer Society, 1997, pp. 306–317. doi:10.1109/LICS.1997.614957.
[19] L. Pacholski, W. Szwast, L. Tendera, Complexity of two-variable logic with counting, in: Proceedings, 12th Annual IEEE Symposium on Logic in Computer Science, IEEE Computer Society, 1997, pp. 318–327. doi:10.1109/LICS.1997.614958.
[20] O. Kuželka, Weighted first-order model counting in the two-variable fragment with counting quantifiers, J. Artif. Intell. Res. 70 (2021) 1281–1307. doi:10.1613/JAIR.1.12320.
[21] J. Tóth, O. Kuželka, Complexity of weighted first-order model counting in the two-variable fragment with counting quantifiers: A bound to beat, in: P. Marquis, M. Ortiz, M. Pagnucco (Eds.), Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024, 2024. doi:10.24963/KR.2024/64.
[22] F. Baader, A new description logic with set constraints and cardinality constraints on role successors, in: C. Dixon, M. Finger (Eds.), Frontiers of Combining Systems – 11th International Symposium, FroCoS 2017, Proceedings, volume 10483 of *Lecture Notes in Computer Science*, Springer, 2017, pp. 43–59. doi:10.1007/978-3-319-66167-4_3.
[23] F. Baader, F. D. Bortoli, On the expressive power of description logics with cardinality constraints on finite and infinite sets, in: A. Herzig, A. Popescu (Eds.), Frontiers of Combining Systems – 12th International Symposium, FroCoS 2019, Proceedings, volume 11715 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 203–219. doi:10.1007/978-3-030-29007-8_12.

[24] N. Jacobson, Basic Algebra I, W. H. Freeman and Company, 1974.
[25] H. Cohen, A Course in Computational Algebraic Number Theory, volume 138 of *Graduate Texts in Mathematics*, Springer, 1993. URL: https://www.worldcat.org/oclc/27810276.
[26] F. Baader, A. Claußnitzer, Maximum Entropy Reasoning via Model Counting in (Description) Logics that Count (Extended Version), LTCS-Report 25-01, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, 2025. URL: https://lat.inf.tu-dresden.de/research/reports.html. doi:10.25368/2025.015.