Reasoning in OWL 2 EL with Hierarchical Concrete Domains

Extended Abstract*

Francesco Kriegel^{1,2}

¹Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany ²Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden and Leipzig, Germany

Concrete Domains. Concrete domains can be integrated in description logics (DLs) in order to refer to concrete knowledge expressed by numbers, strings, and other concrete datatypes [3]. They have mainly been investigated with DLs that are not Horn, such as ALC and its extensions, regarding decidability and complexity [4–9], reasoning procedures [6, 9–13], an algebraic characterization [14, 15], and their expressive power [16, 17].

For computationally tractable description logics, other conditions on the concrete domains than above must be imposed. Suitable for the \mathcal{EL} family are p-admissible concrete domains [18]: they are convex (i.e. every finite disjunction of constraints and negated constraints is already equivalent to one disjunct) and they guarantee that reasoning in the concrete domain is tractable. Due to convexity, it is impossible to introduce disjunction into the ontological domain so that the DL part retains its Horn character. \mathcal{EL} underpins the profile OWL 2 EL of the Web Ontology Language [19], and we here use " \mathcal{EL} " and "OWL 2 EL" as synonyms despite some minor technical differences. Concrete domains have also been integrated with DL-Lite [20].

State of the Art in OWL 2 EL. Existing p-admissible concrete domains for \mathcal{EL} provide only limited utility. Using the concrete domain $\mathcal{D}_{\mathbb{Q},\text{diff}}$ [18], we could express with the concept inclusions (sys = 140) \sqsubseteq Hypertension, (sys > 140) \sqsubseteq Hypertension, (dia = 90) \sqsubseteq Hypertension, and (dia > 90) \sqsubseteq Hypertension that a systolic blood pressure of 140 or higher indicates hypertension, as does a diastolic blood pressure of at least 90, and for example specific values of a patient Bob can be expressed by a concept assertion bob : (sys = 114) \sqcap (dia = 69). However, neither non-elevated blood pressure (dia. below 120 and sys. below 70) nor elevated blood pressure (dia. between 120 and 140, and sys. between 70 and 90) are expressible since the other relations \geq , \leq , < are unavailable in order to avoid introducing disjunctions "through the backdoor." Otherwise the TBox { $\top \sqsubseteq (f > 0), (f = 3) \sqsubseteq C, (f > 3) \sqsubseteq C, (f < 3) \sqsubseteq A, C \sqcap A \sqsubseteq \bot$ } could

^{*}This is an extended abstract of an article accepted at the 15th International Symposium on Frontiers of Combining Systems (FroCoS 2025) [1] and of its extended version [2].

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

[☆] francesco.kriegel@tu-dresden.de (Francesco Kriegel)

https://tu-dresden.de/inf/lat/francesco-kriegel (Francesco Kriegel)

D 0000-0003-0219-0330 (Francesco Kriegel)

^{© 0 2025} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

enforce that the atomic concept A is the complement of the concept C, enabling emulation of the expressivity of ALC (which has exponential-time reasoning complexity).

Mixed inequalities \geq , >, \leq , < may be used under certain limitations which of them may occur in left-hand sides and, respectively, in right-hand sides of concept inclusions [21]. While this ensures convexity, reasoning is rather impaired since the usual completion procedure is only complete for consistency and classification, but not for subsumption.

An algebraic characterization of p-admissible concrete domains has put forth a further concrete domain $\mathcal{D}_{\mathbb{Q},\text{lin}}$, which supports linear combinations of numerical features [22, 23]. For instance, the concept inclusion $\top \sqsubseteq (\text{sys} - \text{dia} - \text{pp} = 0)$, where - is the difference operation in real arithmetic, expresses that the pulse pressure is the difference between the systolic and the diastolic blood pressure. In the medical domain, the combined expressivity of $\mathcal{D}_{\mathbb{Q},\text{diff}}$ and $\mathcal{D}_{\mathbb{Q},\text{lin}}$ would be useful since then with the concept inclusion ICUPatient \sqcap (pp > 50) \sqsubseteq NeedsAttention it could be expressed that intensive-care patients with a pulse pressure exceeding 50 need attention – but this combination is not convex anymore [24]. For instance, the TBox { $\top \sqsubseteq (f + g = 0), (f = 0) \sqsubseteq C, (f > 0) \sqsubseteq C, (g > 0) \sqsubseteq A, C \sqcap A \sqsubseteq \bot$ } declares A as the complement of C. Apart from these p-admissible concrete domains involving numbers, there is another involving strings [18] but it is also too restricted to be of practical use.

Novel Contributions. We introduce a novel form of concrete domains based on semi-lattices. A *semi-lattice* $\mathbf{L} := (L, \leq, \wedge)$ consists of a set L, a partial order \leq , and a binary meet operation \wedge . The elements of L are taken as concrete values, and \leq is understood as an "information order," i.e. $p \leq q$ means that p is equal to or more specific than q, like a subsumption order between concepts. The meet operation \wedge is used to combine two values p and q to their meet value $p \wedge q$, which is the most general value that is equal to or more specific than both p and q.

The hierarchical concrete domain $\mathcal{D}_{\mathbf{L}}$ has values in $\text{Dom}(\mathcal{D}_{\mathbf{L}}) \coloneqq L$ and supports only constraints of the form $f \leq p$ involving a feature f and a value p. Like atomic concepts, these constraints $f \leq p$ can be used within compound concepts, i.e. the concepts' syntax is $C \coloneqq \perp | \top | \{i\} | A | f \leq p | C \sqcap C | \exists r. C$. Their semantics is $(f \leq p)^{\mathcal{I}} = \{x | f^{\mathcal{I}}(x) \leq p\}$ where $f^{\mathcal{I}}$ is a partial function from the domain of \mathcal{I} to the concrete values. Recall: this means that f's value is p or more specific, not smaller like in the aforementioned examples. For instance, real intervals form a semi-lattice with subset inclusion \subseteq as partial order and intersection \cap as meet operation. With that, the statement NonElevatedBP \equiv (sys $\subseteq [0, 120)$) \sqcap (dia $\subseteq [0, 70)$) defines non-elevated blood pressure, where [0, 120) and [0, 70) are real intervals.

In addition, we introduce *FBoxes* consisting of feature inclusions that describe dependencies between features as well as aggregations of features. A *feature inclusion* $f \leq H(g_1, \ldots, g_n)$ consists of features f, g_1, \ldots, g_n and a computable *n*-ary operation $H: L^n \to L$ that is *monotonic* in the sense that $H(p_1, \ldots, p_n) \leq H(q_1, \ldots, q_n)$ whenever $p_1 \leq q_1, \ldots$, and $p_n \leq q_n$ (i.e. applying *H* to equal or more specific values yields equal or more specific values). For instance, through the feature inclusion $pp \subseteq sys - dia$ we can obtain an interval value of the pulse pressure given intervals of the systolic and the diastolic blood pressure. The operator *H* is the difference operation – in real interval arithmetic, which, when applied to intervals *P*, *Q*, yields the set of all numbers p - q where $p \in P$ and $q \in Q$. It is monotonic w.r.t. subset inclusion \subseteq since, simply put, more numbers in *P* or *Q* yield more numbers in P - Q. For



Figure 1: Three graphs representing chemical compounds

instance, we have $H([p_1, q_1], [p_2, q_2]) \coloneqq [p_1 - q_2, q_1 - p_2]$ and similarly for the other interval types. With the concept inclusion ICUPatient $\sqcap (pp \subseteq (50, \infty)) \sqsubseteq$ NeedsAttention we can now express that intensive-care patients having a pulse pressure above 50 need attention and, unlike in the combination of $\mathcal{D}_{\mathbb{Q},\text{diff}}$ and $\mathcal{D}_{\mathbb{Q},\text{lin}}$, computationally reason with that in polynomial time.

Our new hierarchical concrete domains are convex by design. This is because models can assign to features any elements of the semi-lattice, and thus a general value of a feature does not imply the disjunction of all more specific feature values. For example with real intervals, a model of the constraint sys \subseteq [110, 120) can assign the interval [110, 120) to the feature sys, and thus this constraint does not imply the disjunction of, say, sys \subseteq [110, 115) and sys \subseteq [115, 120). In a nutshell, the semi-lattice semantics effectively expels disjunction. Atomic feature values are supported nonetheless when these are available as atoms in the semi-lattice (e.g. singleton intervals [p, p] represent specific numerical values p). In general, $\mathcal{D}_{\mathbf{L}}$ is convex w.r.t. every FBox if the underlying semi-lattice \mathbf{L} is complete (i.e. every subset $P \subseteq L$ has a meet $\bigwedge P \in L$). Furthermore, for each semi-lattice \mathbf{L} that is computable (i.e. L and \leq are decidable and \land is computable) and bounded (i.e. it has a greatest element \top such that $p \leq \top$ for every $p \in L$), $\mathcal{D}_{\mathbf{L}}$ is convex and decidable w.r.t. an FBox \mathcal{F} if \mathbf{L} is well-founded or \mathcal{F} is acyclic.

New Concrete Domains. Besides real intervals already mentioned above, we provide further hierarchical concrete domains based on 2D-polygons, regular languages, and graphs.

With a finite automaton \mathfrak{A} such that $L(\mathfrak{A}) = \Sigma^* \circ \{\text{description logic}\} \circ \Sigma^*$, the concept inclusion ScientificArticle \sqcap (hasTitle $\preceq \mathfrak{A}$) \sqsubseteq DLPaper expresses that all scientific articles with a title containing "description logic" as substring are DL papers.

Structural formulas of molecules can be represented as labeled graphs. Each node is labeled with the atom it represents, and the edges are labeled with the binding type (e.g. single bond, double bond, etc.). The partial order \leq is defined by $\mathcal{G} \leq \mathcal{H}$ if there is a homomorphism from \mathcal{H} to \mathcal{G} , and the meet of two graphs is their disjoint union. Figure 1 shows three exemplary graphs.¹ Graph (c) represents L-leucine, and we can integrate it into a knowledge base with the statement L-Leucine \equiv (hasMolecularStructure $\leq \mathcal{G}_{L-leucine}$). Moreover, the statement AminoAcid \equiv (hasMolecularStructure $\leq \mathcal{G}_{carboxylic acid group}$) \sqcap (hasMolecularStructure $\leq \mathcal{G}_{amino group}$) expresses that amino acids are organic compounds that contain both amino and carboxylic acid functional

¹Graphs (a) and (b) are molecule parts whereas Graph (c) is a complete molecule, which cannot be a part of another molecule. The lower left node in (a) and all outer nodes in (b) can match any element in a larger molecule, be it partial or complete. In Graph (c) the skeletal formula is shown, where labels are optional for carbon atoms (C) and the hydrogen atoms (H) attached to them.

groups. If \mathcal{K} is the knowledge base consisting of the aforementioned statements, then $\mathcal{K} \models$ L-Leucine \sqsubseteq AminoAcid since $\mathcal{G}_{L-leucine} \leq \mathcal{G}_{carboxylic acid group} \land \mathcal{G}_{amino group}$.

Reasoning. Reasoning in \mathcal{EL} can be done by means of a rule-based calculus [18, 25–27], and a hierarchical concrete domain $\mathcal{D}_{\mathbf{L}}$ can be seamlessly integrated into this calculus. Compared to the primal calculus [18, 25], it is only necessary to take the feature inclusions into account (which can now be contained in knowledge bases). For integration into the improved calculus [26, 27] we only need to add the following two rules responsible for interaction between concrete and logical reasoning (where \mathcal{F} consists of all feature inclusions in the knowledge base).

$$\mathsf{R}_{\mathcal{D}}: \quad \frac{C \sqsubseteq (f_1 \le p_1) \cdots C \sqsubseteq (f_m \le p_m)}{C \sqsubseteq (g \le q)}: \mathcal{D}_{\mathbf{L}}, \mathcal{F} \models \prod_{i=1}^m (f_i \le p_i) \sqsubseteq (g \le q)$$
$$\mathsf{R}_{\mathcal{D},\perp}: \quad \frac{C \sqsubseteq (f_1 \le p_1) \cdots C \sqsubseteq (f_m \le p_m)}{C \sqsubseteq \bot}: \quad \prod_{i=1}^m (f_i \le p_i) \text{ unsatisfiable in } \mathcal{D}_{\mathbf{L}}, \mathcal{F}$$

W.r.t. p-admissible hierarchical concrete domains $\mathcal{D}_{\mathbf{L}}$ (e.g. the interval domain, or the convexpolygon domain), the following reasoning tasks can be done in polynomial time: consistency, classification, subsumption checking, instance checking, and concept satisfiability. If concrete reasoning in $\mathcal{D}_{\mathbf{L}}$ is not tractable, then ontological reasoning in the pure \mathcal{EL} part of the knowledge base is not affected and still requires only polynomial time. However, the combined complexities of the aforementioned reasoning tasks are then dominated by the complexity of concrete reasoning (e.g. non-deterministic polynomial time with the graph domain, and exponential time with the regular-language domain or the polygon domain).

Future Prospects. An interesting question for future research is whether *non-local* feature inclusions $f \leq H(R_1 \circ g_1, \ldots, R_n \circ g_n)$ would lead to undecidability or could be reasoned with, where the R_i are role chains. The operator must then be defined for lists of values, like in the non-local feature inclusion combinedWealth $\subseteq \sum$ (hasAccount \circ balance) + \sum (holdsAsset \circ value) over the interval domain, which computes the aggregated wealth of a person or company. At first sight, it seems that the undecidability proof for $\mathcal{EL}(\mathcal{D}_{\mathbb{Q}^2, aff})$ [22] cannot be adapted to this setting. (Mind the braces: (\mathcal{D}) instead of [\mathcal{D}] allows for role chains in front of features.) The computation of canonical valuations must then take into account the graph structure induced by the role assertions entailed by the knowledge base.

In general, it is unclear whether a hierarchical concrete domain is convex and decidable w.r.t. cyclic FBoxes. According to our results for intervals and regular languages, convexity and decidability can be ensured by approaches to solving systems of equations or inequations involving elements of the underlying semi-lattice. This is still open for polygons and graphs.

Since hierarchical concrete domains are convex by design, they are also appropriate for other Horn logics [28] such as \mathcal{ELI} [18], Horn- \mathcal{ALC} [29], Horn- \mathcal{SROIQ} [30], and existential rules [31] — extending the chase procedure with support for them would be practically relevant. Interesting would further be an empirical evaluation, at best with a clear separation of logical and concrete reasoning — especially when tractable logics are equipped with intractable concrete domains. More hierarchical concrete domains of practical relevance should be explored.

Acknowledgments

This work has been supported by Deutsche Forschungsgemeinschaft (DFG) in Project 389792660 (TRR 248: Foundations of Perspicuous Software Systems) and in Project 558917076 (Construction and Repair of Description-logic Knowledge Bases) as well as by the Saxon State Ministry for Science, Culture, and Tourism (SMWK) by funding the Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI).

References

- [1] Francesco Kriegel. Reasoning in OWL 2 EL with Hierarchical Concrete Domains. In: Proceedings of the 15th International Symposium on Frontiers of Combining Systems (FroCoS 2025). To appear. 2025.
- Francesco Kriegel. Reasoning in OWL 2 EL with Hierarchical Concrete Domains (Extended Version). LTCS-Report 25-04. Technische Universität Dresden, 2025. DOI: 10.25368/2025. 127.
- [3] Franz Baader, Philipp Hanschke. A Scheme for Integrating Concrete Domains into Concept Languages. In: Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI). 1991, pp. 452–457. URL: http://ijcai.org/Proceedings/91-1/Papers/070.pdf.
- [4] Franz Baader, Ulrike Sattler. Description logics with aggregates and concrete domains. In: *Information Systems* 28.8 (2003), pp. 979–1004. DOI: 10.1016/S0306-4379(03)00003-6.
- [5] Carsten Lutz. NEXPTIME-complete description logics with concrete domains. In: *ACM Transactions on Computational Logic* 5.4 (2004), pp. 669–705. DOI: 10.1145/1024922.1024925.
- [6] Carsten Lutz. The complexity of description logics with concrete domains. Doctoral Thesis. RWTH Aachen University, Germany, 2002. URL: http://nbn-resolving.org/urn:nbn:de:hbz: 82-opus-3032.
- [7] Claudia Carapelle, Anni-Yasmin Turhan. Description Logics Reasoning w.r.t. General TBoxes Is Decidable for Concrete Domains with the EHD-Property. In: *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI)*. 2016, pp. 1440–1448. DOI: 10.3233/978-1-61499-672-9-1440.
- [8] Stefan Borgwardt, Filippo De Bortoli, Patrick Koopmann. The Precise Complexity of Reasoning in ALC with ω-Admissible Concrete Domains. In: Proceedings of the 37th International Workshop on Description Logics (DL). 2024. URL: https://ceur-ws.org/Vol-3739/paper-1.pdf.
- [9] Carsten Lutz, Carlos Areces, Ian Horrocks, Ulrike Sattler. Keys, Nominals, and Concrete Domains. In: *Journal of Artificial Intelligence Research* 23 (2005), pp. 667–726. DOI: 10.1613/ JAIR.1542.
- [10] Jeff Z. Pan, Ian Horrocks. Reasoning in the SHOQ(D_n) Description Logic. In: Proceedings of the 15th International Workshop on Description Logics (DL). 2002. URL: https://ceurws.org/Vol-53/Pan-Horrocks-shoqdn-2002.ps.

- [11] Carsten Lutz, Maja Miličić. A Tableau Algorithm for Description Logics with Concrete Domains and General TBoxes. In: *Journal of Automated Reasoning* 38.1-3 (2007), pp. 227– 259. DOI: 10.1007/S10817-006-9049-7.
- [12] Volker Haarslev, Ralf Möller, Michael Wessel. The Description Logic \mathcal{ALCNH}_{R^+} Extended with Concrete Domains: A Practically Motivated Approach. In: *Proceedings of the 1st International Joint Conference on Automated Reasoning (IJCAR).* 2001, pp. 29–44. DOI: 10.1007/3-540-45744-5_4.
- [13] Volker Haarslev, Carsten Lutz, Ralf Möller. A Description Logic with Concrete Domains and a Role-forming Predicate Operator. In: *Journal of Logic and Computation* 9.3 (1999), pp. 351–384. DOI: 10.1093/LOGCOM/9.3.351.
- [14] Franz Baader, Jakub Rydval. Description Logics with Concrete Domains and General Concept Inclusions Revisited. In: *Proceedings of the 10th International Joint Conference on Automated Reasoning (IJCAR)*. 2020, pp. 413–431. DOI: 10.1007/978-3-030-51074-9_24.
- [15] Jakub Rydval. Using Model Theory to Find Decidable and Tractable Description Logics with Concrete Domains. Doctoral Thesis. Dresden University of Technology, Germany, 2022. URL: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-799074.
- [16] Franz Baader, Filippo De Bortoli. Logics with Concrete Domains: First-Order Properties, Abstract Expressive Power, and (Un)Decidability. In: SIGAPP Applied Computing Review 24.3 (Oct. 2024), pp. 5–17. DOI: 10.1145/3699839.3699840.
- [17] Franz Baader, Filippo De Bortoli. The Abstract Expressive Power of First-Order and Description Logics with Concrete Domains. In: Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (SAC). 2024, pp. 754–761. DOI: 10.1145/3605098.3635984.
- [18] Franz Baader, Sebastian Brandt, Carsten Lutz. Pushing the *EL* Envelope. In: *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*. 2005, pp. 364–369. URL: http://ijcai.org/Proceedings/05/Papers/0372.pdf.
- [19] Markus Krötzsch. OWL 2 Profiles: An Introduction to Lightweight Ontology Languages. In: Reasoning Web. Semantic Technologies for Advanced Query Answering - 8th International Summer School 2012. Proceedings. 2012, pp. 112–183. DOI: 10.1007/978-3-642-33158-9_4.
- [20] Franz Baader, Stefan Borgwardt, Marcel Lippmann. Query Rewriting for DL-Lite with n-ary Concrete Domains. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI). 2017, pp. 786–792. DOI: 10.24963/IJCAI.2017/109.
- [21] Despoina Magka, Yevgeny Kazakov, Ian Horrocks. Tractable Extensions of the Description Logic *EL* with Numerical Datatypes. In: *Journal of Automated Reasoning* 47.4 (2011), pp. 427–450. DOI: 10.1007/S10817-011-9235-0.
- [22] Franz Baader, Jakub Rydval. Using Model Theory to Find Decidable and Tractable Description Logics with Concrete Domains. In: *Journal of Automated Reasoning* 66.3 (2022), pp. 357–407. DOI: 10.1007/S10817-022-09626-2.
- [23] Franz Baader, Jakub Rydval. An Algebraic View on p-Admissible Concrete Domains for Lightweight Description Logics. In: *Proceedings of the 17th European Conference on Logics in Artificial Intelligence (JELIA).* 2021, pp. 194–209. DOI: 10.1007/978-3-030-75775-5_14.

- [24] Christian Alrabbaa, Franz Baader, Stefan Borgwardt, Patrick Koopmann, Alisa Kovtunova. Combining Proofs for Description Logic and Concrete Domain Reasoning. In: *Proceedings* of the 7th International Joint Conference on Rules and Reasoning (RuleML+RR). 2023, pp. 54– 69. DOI: 10.1007/978-3-031-45072-3_4.
- [25] Franz Baader, Sebastian Brandt, Carsten Lutz. Pushing the *EL* Envelope Further. In: Proceedings of the 4th OWLED Workshop on OWL: Experiences and Directions. 2008. URL: https://ceur-ws.org/Vol-496/owled2008dc_paper_3.pdf.
- [26] Yevgeny Kazakov, Markus Krötzsch, František Simančík. The Incredible ELK From Polynomial Procedures to Efficient Reasoning with *EL* Ontologies. In: *Journal of Automated Reasoning* 53.1 (2014), pp. 1–61. DOI: 10.1007/S10817-013-9296-3.
- [27] Yevgeny Kazakov, Pavel Klinov. Advancing ELK: Not Only Performance Matters. In: Proceedings of the 28th International Workshop on Description Logics (DL). 2015. URL: https://ceur-ws.org/Vol-1350/paper-27.pdf.
- [28] George F. McNulty. Fragments of First Order Logic, I: Universal Horn Logic. In: *Journal* of Symbolic Logic 42.2 (1977), pp. 221–237. DOI: 10.2307/2272123.
- [29] Markus Krötzsch, Sebastian Rudolph, Pascal Hitzler. Complexities of Horn Description Logics. In: ACM Transactions on Computational Logic 14.1 (2013), 2:1–2:36. DOI: 10.1145/ 2422085.2422087.
- [30] Magdalena Ortiz, Sebastian Rudolph, Mantas Šimkus. Worst-Case Optimal Reasoning for the Horn-DL Fragments of OWL 1 and 2. In: Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR). 2010. URL: http://aaai.org/ocs/index.php/KR/KR2010/paper/view/1296.
- [31] Jean-François Baget, Michel Leclère, Marie-Laure Mugnier, Eric Salvat. Extending Decidable Cases for Rules with Existential Variables. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI). 2009, pp. 677–682. URL: http: //ijcai.org/Proceedings/09/Papers/118.pdf.