Beyond $\mathcal{ALC}_{\mathsf{reg}}$: Exploring Non-Regular Extensions of PDL with Description Logics Features

Bartosz Bednarczyk[1,2]

[1] Computational Logic Group, Technische Universität Dresden, Germany
[2] Institute of Computer Science, University of Wrocław, Poland
bartosz.bednarczyk@cs.uni.wroc.pl

Abstract. We investigate the impact of non-regular path expressions on the decidability of satisfiability checking and querying in description logics. Our primary object of interest is $\mathcal{ALC}_{\mathsf{vpl}}$, an extension of $\mathcal{ALC}$ with path expressions using visibly-pushdown languages, which was shown to be decidable by Löding et al. in 2007. We prove that decidability of $\mathcal{ALC}_{\mathsf{vpl}}$ is preserved when enriching the logic with functionality, but decidability is lost upon adding the seemingly innocent Self operator. We also consider the simplest non-regular (visibly-pushdown) language $r^{\#}s^{\#} := \{r^n s^n \mid n \in \mathbb{N}\}$. We establish undecidability of the concept satisfiability problem for $\mathcal{ALC}_{\mathsf{reg}}$ extended with nominals and $r^{\#}s^{\#}$, as well as of the query entailment problem for $\mathcal{ALC}$-TBoxes, where such non-regular atoms are present in queries.

1 Introduction

Formal ontologies play a crucial role in artificial intelligence, serving as the backbone of various applications such as the Semantic Web, ontology-based information integration, and peer-to-peer data management. In reasoning about graph-structured data, a significant role is played by *description logics* (DLs) [2], a robust family of logical formalisms serving as the logical foundation of contemporary standardized ontology languages, including OWL 2 by the W3C [16,23]. Among many features present in extensions of the basic description logic $\mathcal{ALC}$, an especially useful one is $\cdot_{\mathsf{reg}}$, supported by popular $\mathcal{Z}$-family of description logics [10]. With $\cdot_{\mathsf{reg}}$ one can specify regular path constraints, allowing the user to navigate graph-structured data. In recent years many extensions of $\mathcal{ALC}_{\mathsf{reg}}$ for ontology-engineering were proposed, see *e.g.* [6,11,29], and the complexity landscape of their reasoning problems is now mostly well-understood [10,5,4]. By a mere coincidence, the logic $\mathcal{ALC}_{\mathsf{reg}}$ was already studied in 1979 by the formal-verification community [13], under the name of Propositional Dynamic Logic (PDL). Consult [12] for a discussion on relationship between (extensions of) PDL and $\mathcal{ALC}_{\mathsf{reg}}$.

Due to wideness of the spectrum of recognizable word languages, the question of whether regularity constraints in path expressions of $\mathcal{ALC}_{\mathsf{reg}}$ can be lifted to more expressive classes of languages received a lot of attention from researchers. We call such extensions *non-regular*. After the first undecidability proof of satisfiability of $\mathcal{ALC}_{\mathsf{reg}}$ with context-free languages [20], several decidable cases were identified. For instance, Koren and Pnueli [25] proved that $\mathcal{ALC}_{\mathsf{reg}}$ extended with the simplest non-regular language $r^{\#}s^{\#} := \{r^n s^n \mid n \in \mathbb{N}\}$ for *fixed* roles r, s is decidable; while combining it with $s^{\#}r^{\#}$ leads to undecidability [19]. This surprises at first glance, but as it was shown later [28], PDL extended with a broad class of input-driven context-free languages, called *visibly pushdown languages* [1], remain decidable. This generalizes all previously known decidability results, and partially explains the reason behind known failures (*e.g.* the languages $r^{\#}s^{\#}$ and $s^{\#}r^{\#}$ cannot be both visibly-pushdown under the same partition of the alphabet).

Our motivation and contribution. Despite the presence of a plethora of various results concerning non-regular extensions of PDL [25,18,22,21,8], no one considered their extensions with popular features supported by W3C ontology languages. Such extensions are, *e.g. nominals* (constants), *inverse roles* (inverse programs), *functionality* (deterministic programs), and Self operator (self-loops). The honourable exception is the unpublished undecidability result for $\mathcal{ALC}_{\mathsf{reg}}$ extended with the language $r^{\#}s(r^{-})^{\#}$ (with r^{-} denoting the converse of r) from Göller's thesis [15]. The lack of results on entailment of non-regular queries over ontologies is also intriguing, taking into account positive results for conjunctive visibly-pushdown queries in the setting of relational-databases [27].
In this paper we contribute to a further understanding of the aforementioned questions. Our results are mostly negative. For the first part of the paper, we investigate $\mathcal{ALC}_{\mathsf{reg}}$ with $r^{\#}s^{\#}$. In Section 3 we prove that its extension with nominals has an undecidable satisfiability problem. In Section 4 we show that, already for $\mathcal{ALC}$, the query entailment problem of queries involving $r^{\#}s^{\#}$, is also undecidable. For the second part of the paper, we study $\mathcal{ALC}_{\mathsf{vpl}}$, the extension of $\mathcal{ALC}_{\mathsf{reg}}$ with visibly pushdown languages (that generalize $r^{\#}s^{\#}$). By a simple translation, Section 5 establishes decidability of the logic extended with functionality. As opposed to this, adding the seemingly innocent Self renders the logic undecidable.

Because of lack of space, the journal version of this paper contains all missing proofs, extra pictures and expanded definitions.

2 Preliminaries

We assume familiarity with basics on description logic $\mathcal{ALC}$ [2, Sec. 2.1–2.3], regular and context-free languages, Turing machines and computability [33, Sec. 1–5]. As usual, $\mathbb{N}$ denotes non-negative integers, and $\mathbb{Z}_n$ denotes the set $\{0, 1, \ldots, n-1\}$.

Basics on $\mathcal{ALC}$. We fix countably infinite pairwise disjoint sets of *individual names* $\mathbf{N_I}$, *concept names* $\mathbf{N_C}$, and *role names* $\mathbf{N_R}$ and introduce the description logic $\mathcal{ALC}$. Starting from $\mathbf{N_C}$ and $\mathbf{N_R}$, the set $\mathbf{C}_{\mathcal{ALC}}$ of *$\mathcal{ALC}$-concepts* is built using the following concept constructors: *negation* ($\neg$C), *conjunction* (C $\sqcap$ D), *existential restriction* ($\exists r$.C), and the *top concept* $\top$ with the grammar:

$$\mathrm{C}, \mathrm{D} ::= \top \mid \mathrm{A} \mid \neg\mathrm{C} \mid \mathrm{C} \sqcap \mathrm{D} \mid \exists r.\mathrm{C},$$

where $\mathrm{C}, \mathrm{D} \in \mathbf{C}_{\mathcal{ALC}}$, $\mathrm{A} \in \mathbf{N_C}$ and $r \in \mathbf{N_R}$. We employ the following abbreviations: $\mathrm{C} \sqcup \mathrm{D} := \neg(\neg\mathrm{C} \sqcap \neg\mathrm{D})$, $\forall r.\mathrm{C} := \neg\exists r.\neg\mathrm{C}$, $\bot := \neg\top$, and $\mathrm{C} \to \mathrm{D} := \neg\mathrm{C} \sqcup \mathrm{D}$.

The semantics of $\mathcal{ALC}$ is defined via *interpretations* $\mathcal{I} \coloneqq (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ composed of a non-empty set $\Delta^{\mathcal{I}}$ called the *domain of* $\mathcal{I}$ and an *interpretation function* $\cdot^{\mathcal{I}}$ mapping individual names to elements of $\Delta^{\mathcal{I}}$, concept names to subsets of $\Delta^{\mathcal{I}}$, and role names to subsets of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. This mapping is then extended to concepts.

Name	Syntax	Semantics
top concept	$\top$	$\Delta^{\mathcal{I}}$
concept negation	$\neg\mathrm{C}$	$\Delta^{\mathcal{I}} \setminus \mathrm{C}^{\mathcal{I}}$
concept intersection	$\mathrm{C} \sqcap \mathrm{D}$	$\mathrm{C}^{\mathcal{I}} \cap \mathrm{D}^{\mathcal{I}}$
existential restriction	$\exists r.\mathrm{C}$	$\{\mathrm{d} \mid \exists \mathrm{e} \in \mathrm{C}^{\mathcal{I}}\ (\mathrm{d},\mathrm{e}) \in r^{\mathcal{I}}\}$

An interpretation $\mathcal{I}$ *satisfies* a concept C (or $\mathcal{I}$ is a *model* of C, written: $\mathcal{I} \models \mathrm{C}$) if $\mathrm{C}^{\mathcal{I}} \neq \emptyset$. A concept is *satisfiable* if it has a model. In the *satisfiability problem* we ask, whether an input concept has a model. We consider three popular description-logics features: *nominals* ($\mathcal{O}$), *functionality* ($\mathcal{F}$), and the Self *operator* ($\cdot^{\mathsf{Self}}$). Their semantics is recalled in the table below, assuming that $r, s \in \mathbf{N_R}$, and $\mathsf{a} \in \mathbf{N_I}$.

Name	Syntax	Semantics
functionality	$\mathsf{func}(r)$	$\mathcal{I} \models \mathsf{func}(r)$ if $\forall \mathrm{d} \forall \mathrm{e}_1 \forall \mathrm{e}_2 \big((\mathrm{d},\mathrm{e}_1) \in r^{\mathcal{I}} \wedge (\mathrm{d},\mathrm{e}_2) \in r^{\mathcal{I}} \Rightarrow \mathrm{e}_1 = \mathrm{e}_2\big)$
nominal	$\{\mathsf{a}\}$	$\{\mathsf{a}^{\mathcal{I}}\}$
self-operator	$\exists r.\mathsf{Self}$	$\{\mathrm{d} \mid (\mathrm{d},\mathrm{d}) \in r^{\mathcal{I}}\}$

A path ρ in an interpretation $\mathcal{I}$ is a finite word in $(\Delta^{\mathcal{I}})^*$. We usually enumerate its components with $\rho_1, \ldots, \rho_{|\rho|}$, where the number $|\rho|-1$ is called the *length* of ρ. We say that ρ *starts from* (resp. *ends in*) d if $\rho_1 = \mathrm{d}$ holds (resp. $\rho_{|\rho|} = \mathrm{d}$). If $\mathsf{N} \subseteq \mathbf{N_I}$ is given, we call an element $\mathrm{d} \in \Delta^{\mathcal{I}}$ N-*named* if $\mathrm{d} = \mathsf{a}^{\mathcal{I}}$ holds for some $\mathsf{a} \in \mathsf{N}$.

$\mathcal{ALC}$ with path expressions. We treat $\Sigma_{\mathsf{all}} \coloneqq \mathbf{N_R} \cup \{\mathrm{A}? \mid \mathrm{A} \in \mathbf{N_C}\}$ as an infinite alphabet. Let $\mathbb{ALL}$ and $\mathbb{REG}$ denote classes of all recognizable (resp. regular) finite-word languages over finite subsets of Σ_{all}. For a language $\boldsymbol{\mathcal{L}}$ and a path $\rho \coloneqq \rho_1\rho_2\ldots\rho_n\rho_{n+1}$ in an interpretation $\mathcal{I}$, we say that ρ is an $\boldsymbol{\mathcal{L}}$-*path*, if there exists a word $\mathtt{w} \coloneqq \mathtt{w}_1\mathtt{w}_2\ldots\mathtt{w}_n \in \boldsymbol{\mathcal{L}}$ such that for all $i \leq n$ we have either (i) $\mathtt{w}_i \in \mathbf{N_R}$ and $(\rho_i, \rho_{i+1}) \in (\mathtt{w}_i)^{\mathcal{I}}$, or (ii) $\mathtt{w}_i$ has the form A?, $\rho_i = \rho_{i+1}$ and $\rho_i \in \mathrm{A}^{\mathcal{I}}$. Intuitively $\mathtt{w}$ either traverses roles or loops at an element to check the satisfaction of concepts. We say that $\mathrm{e} \in \Delta^{\mathcal{I}}$ is $\boldsymbol{\mathcal{L}}$-*reachable from* $\mathrm{d} \in \Delta^{\mathcal{I}}$ (or that d $\boldsymbol{\mathcal{L}}$-*reaches* e) if there is an $\boldsymbol{\mathcal{L}}$-path ρ that starts from d and ends in e. The logic $\mathcal{ALC}_{\mathsf{all}}$ extends $\mathcal{ALC}$ with concept constructors of the form $\exists\boldsymbol{\mathcal{L}}.\mathrm{C}$, where $\boldsymbol{\mathcal{L}} \in \mathbb{ALL}$ and C is an $\mathcal{ALC}_{\mathsf{all}}$-concept. Their semantics is as follows: $(\exists\boldsymbol{\mathcal{L}}.\mathrm{C})^{\mathcal{I}}$ is the set of all $\mathrm{d} \in \Delta^{\mathcal{I}}$ that can $\boldsymbol{\mathcal{L}}$-reach some $\mathrm{e} \in \mathrm{C}^{\mathcal{I}}$, and $\forall\boldsymbol{\mathcal{L}}.\mathrm{C}$ stands for $\neg\exists\boldsymbol{\mathcal{L}}.\neg\mathrm{C}$. The logic $\mathcal{ALC}_{\mathsf{reg}}$ (a.k.a. PDL [13]) is a restriction of $\mathcal{ALC}_{\mathsf{all}}$ to regular languages.

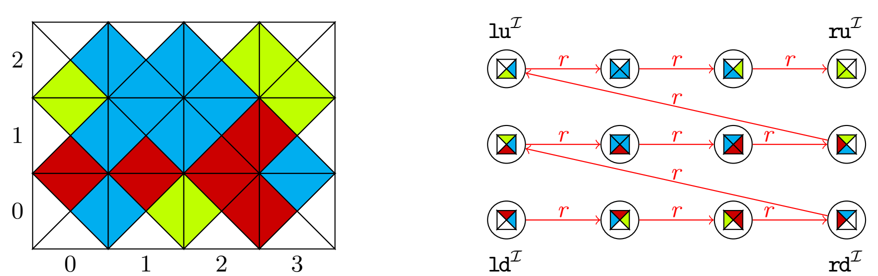
VPLs. The class of *Visibly-pushdown languages* (VPLs) [1] is a well-behaved family of context-free languages, in which the usage of the stack in the underlying pushdown automata model is input-driven. A *pushdown alphabet* Σ is an alphabet equipped with a partition $(\Sigma_c, \Sigma_i, \Sigma_r)$. The elements of Σ_c, Σ_i, and Σ_r are called, respectively, *call* letters, *internal* letters, and *return* letters. A *visibly-pushdown automaton* (VPA) $\boldsymbol{\mathcal{A}}$ over a pushdown alphabet Σ is a deterministic pushdown automaton that can push (resp. pop) a letter from its stack only after reading a call (resp. return) symbol. A *visibly one-counter automaton* [3] (VOCA) is a VPA that can use only a single stack letter. Given a VPA $\mathcal{A}$, we speak about words *accepted* by $\mathcal{A}$, and the language $\mathcal{L}(\mathcal{A})$ of $\mathcal{A}$ defined in the usual way. As an example, suppose that $r \in \Sigma_c$ and $s \in \Sigma_r$. Then the language $r^{\#}s^{\#} := \{r^n s^n \mid n \in \mathbb{N}\}$ is visibly-pushdown, but the language $s^{\#}r^{\#}$ over the same alphabet is not. What is more, every regular language is visibly-pushdown. We present Σ_{all} as a pushdown alphabet $((\mathbf{N_R})_c, (\mathbf{N_R})_i \cup \{A? \mid A \in \mathbf{N_C}\}, (\mathbf{N_R})_r)$. The logic $\mathcal{ALC}_{\mathsf{vpl}}$ is defined as the restriction of $\mathcal{ALC}_{\mathsf{all}}$ to visibly-pushdown languages over finite subsets of Σ_{all} (note that the letters are equally partitioned for all the languages). It is known that $\mathcal{ALC}_{\mathsf{vpl}}$ has 2ExpTime-complete [28] satisfiability problem. Finally, $\mathcal{ALC}^{r\#s\#}_{\mathsf{reg}}$ denotes the restriction of $\mathcal{ALC}_{\mathsf{vpl}}$ in which the only allowed non-regular language is $r^{\#}s^{\#}$ for *fixed* call r and return s.

3 Nominals lead to undecidability

We first establish undecidability of the satisfiability problem for $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$.

A *domino tiling system* is a triple $\mathcal{D} := (\mathrm{Col}, \mathrm{T}, \square)$, where Col is a finite set of *colours*, $\mathrm{T} \subseteq \mathrm{Col}^4$ is a finite set of 4-sided *tiles*, and $\square \in \mathrm{Col}$ is a distinguished colour called *white*. For brevity, we call a tile $(\mathrm{c}_l, \mathrm{c}_d, \mathrm{c}_r, \mathrm{c}_u) \in \mathrm{T}$ (i) *left-border* if $\mathrm{c}_l = \square$, (ii) *down-border* if $\mathrm{c}_d = \square$, (iii) *right-border* $\mathrm{c}_r = \square$, and (iii) *up-border* if $\mathrm{c}_u = \square$. We say that $\mathrm{t} := (\mathrm{c}_l, \mathrm{c}_d, \mathrm{c}_r, \mathrm{c}_u)$ and $\mathrm{t}' := (\mathrm{c}'_l, \mathrm{c}'_d, \mathrm{c}'_r, \mathrm{c}'_u)$ from T are (i) H-*compatible* if $\mathrm{c}_r = \mathrm{c}'_l$, and (ii) V-*compatible* if $\mathrm{c}_u = \mathrm{c}'_d$. We say that $\mathcal{D}$ *covers* $\mathbb{Z}_n \times \mathbb{Z}_m$ (where n, m are positive integers) if there is a mapping $\xi \colon \mathbb{Z}_n \times \mathbb{Z}_m \to \mathrm{T}$ such that for all pairs $(x, y) \in \mathbb{Z}_n \times \mathbb{Z}_m$ with $\xi(x, y) := (\mathrm{c}_l, \mathrm{c}_d, \mathrm{c}_r, \mathrm{c}_u)$ we have:

(TBor) $x = 0$ iff $\mathrm{c}_l = \square$; $x = n{-}1$ iff $\mathrm{c}_r = \square$; $y = 0$ iff $\mathrm{c}_d = \square$; $y = m{-}1$ iff $\mathrm{c}_u{=}\square$
(THori) If $(x{+}1, y) \in \mathbb{Z}_n \times \mathbb{Z}_m$ then $\xi(x, y)$ and $\xi(x{+}1, y)$ are H-compatible.
(TVerti) If $(x, y{+}1) \in \mathbb{Z}_n \times \mathbb{Z}_m$ then $\xi(x, y)$ and $\xi(x, y{+}1)$ are V-compatible.



(a) Visualization of ξ. (b) The encoding of ξ as a $\mathcal{D}$-snake $\mathcal{I}$.

Fig. 1: If Col = {, , $\square$, } and $\mathrm{T} = \mathrm{Col}^4$, the map $\xi := \{(0,0) \mapsto$, $(1,0) \mapsto$, $(2,0) \mapsto$, $(3,0) \mapsto$, $(0,1) \mapsto$, $(1,1) \mapsto$, $(2,1) \mapsto$, $(3,1) \mapsto$, $(0,2) \mapsto$, $(1,2) \mapsto$, $(2,2) \mapsto$, $(3,2) \mapsto$ $\}$ covers $\mathbb{Z}_4 \times \mathbb{Z}_3$.

Intuitively, $\xi : \mathbb{Z}_n \times \mathbb{Z}_m$ can be seen as a rectangle of size $n \times m$ coloured by unit 4-sided tiles (with coordinates corresponding to the left, down, right, and upper colour) from T, where sides of tiles of consecutive squares have matching colours, and borders of the rectangle are white. Consult Figure 1a for more intuitions.

W.l.o.g. we will always assume that T does not contain tiles having more than two white sides. A system $\mathcal{D}$ is *solvable* if there exist positive $n, m \in \mathbb{N}$ for which $\mathcal{D}$ covers $\mathbb{Z}_n \times \mathbb{Z}_m$. The problem of deciding if an input domino tiling system is solvable is undecidable, which can be shown by a minor modification of classical undecidability proofs for tilling problems, see *e.g.* [31, Lemma 3.9].

For a tiling system $\mathcal{D} \coloneqq (\mathrm{Col}, \mathrm{T}, \Box)$ we encode mappings ξ from T to some $\mathbb{Z}_n \times \mathbb{Z}_m$ in interpretations $\mathcal{I}$ as certain r^+-paths ρ from $\mathtt{ld}^{\mathcal{I}}$ to $\mathtt{ru}^{\mathcal{I}}$ passing through $\mathtt{rd}^{\mathcal{I}}$ and $\mathtt{lu}^{\mathcal{I}}$ (using fresh names from $\mathsf{N}^{\mathrm{T}}_{\maltese} \coloneqq \{\mathtt{ld}, \mathtt{rd}, \mathtt{lu}, \mathtt{ru}\}$) composed of elements labelled with fresh concepts names from $\mathrm{C}^{\mathrm{T}}_{\maltese} \coloneqq \{\mathrm{C}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$, see Figure 1b.

Definition 1. *An interpretation $\mathcal{I}$ is a $\mathcal{D}$-*snake *for a tiling system $\mathcal{D}$ if:*

(SPath) *There is an r^+-path ρ that starts in $\mathtt{ld}^{\mathcal{I}}$, then passes through $\mathtt{rd}^{\mathcal{I}}$, then passes through $\mathtt{lu}^{\mathcal{I}}$ and finishes in $\mathtt{ru}^{\mathcal{I}}$.*

(SNoLoop) *No $\mathsf{N}^{\mathrm{T}}_{\maltese}$-named element can r^+-reach itself.*

(SUniqTil) *For every* d *r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ there is precisely one tile* $\mathrm{t} \in \mathrm{T}$ *such that* $\mathrm{d} \in \mathrm{C}^{\mathcal{I}}_{\mathrm{t}}$ *(we say that* d *is* labelled *by a tile* t *or that it* carries t*).*

(SSpecTil) *The $\mathsf{N}^{\mathrm{T}}_{\maltese}$-named elements are unique elements r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ that are labelled by tiles with two white sides. Moreover, we have that (a) $\mathtt{ld}^{\mathcal{I}}$ carries a tile that is left-border and down-border, (b) $\mathtt{rd}^{\mathcal{I}}$ carries a tile that is right-border and down-border, (c) $\mathtt{lu}^{\mathcal{I}}$ carries a tile that is left-border and up-border, (d) $\mathtt{ru}^{\mathcal{I}}$ carries a tile that is right-border and up-border.*

(SHori) *For all elements* d *different from $\mathtt{ru}^{\mathcal{I}}$ that are r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ and labelled by some tile* $\mathrm{t} \coloneqq (\mathrm{c}_l, \mathrm{c}_d, \mathrm{c}_r, \mathrm{c}_u)$*, there exists a tile* $\mathrm{t}' \coloneqq (\mathrm{c}'_l, \mathrm{c}'_d, \mathrm{c}'_r, \mathrm{c}'_u)$ *for which all r-successors* e *of* d *carry the tile* t′ *and: (i)* t, t′ *are* H*-compatible, (ii) if* $\mathrm{c}_d = \Box$ *then* $(\mathrm{c}_r \neq \Box$ *iff* $\mathrm{c}'_d = \Box)$*, and (iii) if* $\mathrm{c}_u = \Box$ *then* $\mathrm{c}'_u = \Box$*.*

(SLen) *There is a unique* N *such that all r^+-paths between $\mathtt{ld}^{\mathcal{I}}$ and $\mathtt{rd}^{\mathcal{I}}$ are of length* $\mathrm{N}-1$*. Moreover, $\mathtt{rd}^{\mathcal{I}}$ is the only element $r^{\mathrm{N}-1}$-reachable from $\mathtt{ld}^{\mathcal{I}}$.*

(SVerti) *For all elements* d *that are r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ and labelled by some* $\mathrm{t} \in \mathrm{T}$ *that is not up-border, we have that (a) there exists a tile* $\mathrm{t}' \in \mathrm{T}$ *such that all elements* e *r^{N}-reachable (for* N *guaranteed by* (SLen)*) from* d *carry* t′*, (b)* t *and* t′ *are* V*-compatible, (c)* t *is left-border (resp. right-border) iff* t′ *is.*

If $\mathcal{I}$ satisfy all but the last two conditions, we call it a $\mathcal{D}$-*pseudosnake*. The key property of our encoding is summarized in the following lemma.

Lemma 2. *A domino tiling system $\mathcal{D}$ is solvable iff there exists a $\mathcal{D}$-snake.*

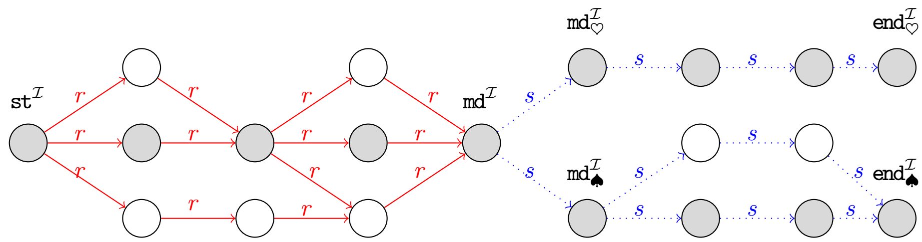
While $\mathcal{D}$-snakes are not directly axiomatizable in $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$, we at least see how to express $\mathcal{D}$-pseudosnakes. See full version of the paper for the proof.

Lemma 3. *For every tiling system $\mathcal{D} \coloneqq (\mathrm{Col}, \mathrm{T}, \Box)$, there is an $\mathcal{ALCO}^{r\#s\#}_{reg}$-concept $\mathrm{C}^{\mathcal{D}}_{\mathcal{S}}$, that employs the role r, individual names from $\mathsf{N}^{\mathrm{T}}_{\maltese}$ and concept names from $\mathrm{C}^{\mathrm{T}}_{\maltese}$, such that for all $\mathcal{I}$ we have that $\mathcal{I}$ is a $\mathcal{D}$-pseudosnake iff $\mathcal{I} \models \mathrm{C}^{\mathcal{D}}_{\mathcal{S}}$.*

Note that the property that pseudosnakes are missing in order to be proper snakes, is the ability to measure. We tackle this issue by introducing "yardsticks".

Definition 4. *Let* T *be a finite and non-empty set, and let* $\mathsf{N}^{\mathrm{T}}_{\text{📏}} := \{\mathit{st}, \mathit{md}, \mathit{md}_{\mathrm{t}}, \mathit{end}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$ *be composed of (pairwise different) individual names. A* T-yardstick *is any interpretation* $\mathcal{I}$ *that satisfies all the conditions listed below.*

(YDifNom) $\mathsf{N}^{\mathrm{T}}_{\text{📏}}$*-named elem. are pairwise-diffr. and* $(r+s)^*$*-reach. from* $\mathit{st}^{\mathcal{I}}$*.*
(YNoLoop) *No* $\mathsf{N}^{\mathrm{T}}_{\text{📏}}$*-named element can* $(r+s)^+$*-reach itself.*
(YMid) $\mathit{md}^{\mathcal{I}}$ *is the* unique *elem. with an* s*-successor that is* r^**-reachable from* $\mathit{st}^{\mathcal{I}}$*.*
(YSuccOfMid) s*-successors of* $\mathit{md}^{\mathcal{I}}$ *are precisely* $\{\mathit{md}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$*-named elems.*
(YReachMidT) *For every* $\mathrm{t} \in \mathrm{T}$ *we have that* $\mathit{md}_{\mathrm{t}}^{\mathcal{I}}$ *can* s^**-reach* $\mathit{end}_{\mathrm{t}}^{\mathcal{I}}$ *but it cannot* s^**-reach* $\mathit{end}_{\mathrm{t}'}^{\mathcal{I}}$ *for all* $\mathrm{t}' \neq \mathrm{t}$*.*
(YEqDst) *The* $\{\mathit{end}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$*-named elements are precisely the elements* $r^{\#}s^{\#}$*-reachable from* $\mathit{st}^{\mathcal{I}}$*.*
(YNoEqDst) *No* $\{\mathit{end}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$*-named element is* $r^{\#}s^{\#}$*-reachable from an element* $(s+r)^+$*-reachable from* $\mathit{st}^{\mathcal{I}}$*.*



An example $\{\heartsuit, \spadesuit\}$-yardstick is depicted above. A "minimal" yardstick would contain the grey nodes only. Lemma 5 justifies the name "yardstick". Intuitively it says that in any T-yardstick $\mathcal{I}$, all s^*-paths from $\mathtt{md}^{\mathcal{I}}$ to all $\mathtt{end}_{\mathrm{t}}^{\mathcal{I}}$ have equal length, to which we refer as the *length* of $\mathcal{I}$.

Lemma 5. *Let* $\mathcal{I}$ *be a* T*-yardstick. Then there exists a unique positive integer* N *such that: (i) for all* $\mathrm{t} \in \mathrm{T}$ *we have that* $\mathit{end}_{\mathrm{t}}^{\mathcal{I}}$ *is* s^{N}*-reachable from* $\mathit{md}^{\mathcal{I}}$*, and (ii) for all* $\mathrm{t} \in \mathrm{T}$ *we have that* $\mathit{end}_{\mathrm{t}}^{\mathcal{I}}$ *is* $s^{\mathrm{N}-1}$*-reachable from* $\mathit{md}_{\mathrm{t}}^{\mathcal{I}}$*.*

Proof. Fix $\mathrm{t}_\star \in \mathrm{T}$. By (YEqDst) we know that $\mathtt{st}^{\mathcal{I}}$ $r^{\#}s^{\#}$-reaches $\mathtt{end}_{\mathrm{t}_\star}^{\mathcal{I}}$, and let $\rho := \rho_1 \ldots \rho_{2\mathrm{N}+1}$ be a path witnessing it. We claim that N is the desired length of $\mathcal{I}$. First, note that N is greater than 0 by (YDifNom). Second, by the semantics of $r^{\#}s^{\#}$, for all $i \leq \mathrm{N}$ we have $(\rho_i, \rho_{i+1}) \in r^{\mathcal{I}}$ and $(\rho_{\mathrm{N}+i}, \rho_{\mathrm{N}+i+1}) \in s^{\mathcal{I}}$. Thus $\rho_{\mathrm{N}+1}$ is r^*-reachable from $\mathtt{st}^{\mathcal{I}}$ and has an s-successor. These two facts imply (by (YMid)) that $\rho_{\mathrm{N}+1}$ is equal to $\mathtt{md}^{\mathcal{I}}$. It remains to show that all the paths leading from $\mathtt{md}^{\mathcal{I}}$ to some $\mathtt{end}_{\mathrm{t}}$ are of length N. Towards a contradiction, assume that there is $\mathrm{t}' \in \mathrm{T}$ and an integer $\mathrm{M} \neq \mathrm{N}$ such that $\mathtt{md}^{\mathcal{I}}$ s^{M}-reaches $\mathtt{end}_{\mathrm{t}'}^{\mathcal{I}}$ via a path $\rho' := \rho'_1 \ldots \rho'_{\mathrm{M}}$. We stress that $\rho'_1 = \mathtt{md}^{\mathcal{I}}$ and $\rho'_{\mathrm{M}} = \mathtt{end}_{\mathrm{t}'}^{\mathcal{I}}$ (by design of ρ'), and $\rho'_2 = \mathtt{md}_{\mathrm{t}'}^{\mathcal{I}}$ (by a conjunction of (YSuccOfMid) and (YReachMidT)). To conclude the proof, it suffices to resolve the following two cases.

– Suppose that $\mathrm{M} < \mathrm{N}$. Then $\rho_{\mathrm{N}+1-\mathrm{M}}$ $(r^{\mathrm{M}}s^{\mathrm{M}})$-reaches (thus $r^{\#}s^{\#}$-reaches) $\mathtt{end}_{\mathrm{t}'}^{\mathcal{I}}$, as witnessed by the path $\rho_{\mathrm{N}+1-\mathrm{M}} \ldots \rho_{\mathrm{N}}\rho'$. Moreover $\rho_{\mathrm{N}+1-\mathrm{M}}$ is r^+-reachable from $\mathtt{st}^{\mathcal{I}}$, witnessed by the path $\rho_1 \ldots \rho_{\mathrm{N}+1-\mathrm{M}}$ (note that its length is positive by the inequality $\mathrm{M} < \mathrm{N}$). This contradicts (YNoEqDst).

– Suppose that $\mathrm{M} > \mathrm{N}$. Consider the path $\rho_1 \ldots \rho_{\mathrm{N}} \rho'_1 \ldots \rho'_{\mathrm{N}}$. By design, such a path witnesses that $\mathtt{st}^{\mathcal{I}}$ $(r^{\mathrm{N}} s^{\mathrm{N}})$-reaches (and thus also $r^{\#} s^{\#}$-reaches) ρ'_{N}. By (YEqDst) we infer that ρ'_{N} is then $\{\mathtt{end}_{\mathrm{t}} \mid \mathrm{t} \in \mathrm{T}\}$-named. As $\rho'_2 = \mathtt{md}^{\mathcal{I}}_{\mathrm{t}'}$ s^{+}-reaches ρ'_{N}, we infer that $\rho'_{\mathrm{N}} = \mathtt{end}^{\mathcal{I}}_{\mathrm{t}'}$ (otherwise we would have a contradiction with (YReachMidT)). But then $\mathtt{end}^{\mathcal{I}}_{\mathrm{t}'}$ s^{+}-reaches itself via a path $\rho'_{\mathrm{N}} \ldots \rho_{\mathrm{M}}$, which is of positive length due to $\mathrm{M} > \mathrm{N}$. A contradiction with (YNoLoop).

This establishes Property (i). The satisfaction of Property (ii) is now immediate.

As the next step of our construction, we establish existence of arbitrary long yardsticks, and axiomatize them with an $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$-concept. Indeed:

Lemma 6. *For every finite non-empty set* T *and a positive integer* N*, there exists a* T*-yardstick of length* N*. Moreover, there exists an* $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$*-concept* $\mathrm{C}^{\mathrm{T}}_{\text{📏}}$*, that employs only role names* r, s *and individual names from* $\mathsf{N}^{\mathrm{T}}_{\text{📏}}$*, such that for all interpretations* $\mathcal{I}$ *we have that* $\mathcal{I}$ *is a* T*-yardstick if and only if* $\mathcal{I} \models \mathrm{C}^{\mathrm{T}}_{\text{📏}}$.

We next put pseudosnakes and yardsticks together, obtaining metricobras. The intuition behind their construction is fairly simple: (i) we take a disjoint union of a pseudosnake and a yardstick, (ii) we then connect (via the role s) every element carrying a tile t with the interpretation of the corresponding nominal $\mathtt{md}_{\mathrm{t}}$, and finally (iii) we synchronize the length of the underlying yardstick, say N, with the length of the path between the interpretations of $\mathtt{ld}$ and $\mathtt{rd}$. After such "merging", retrieving (SHori) is relatively easy: rather than testing if every N-reachable element from some d carries a suitable tile t (for an a priori unknown N) we can check instead whether d can $r^{\#} s^{\#}$-reach the interpretation of $\mathtt{end}_{\mathrm{t}}$.

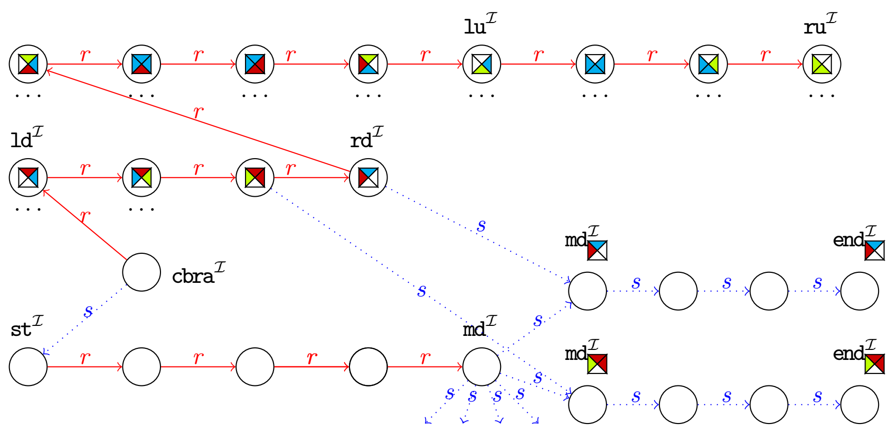


Fig. 2: A fragment of an example $\mathcal{D}$-metricobra representing ξ from Figure 1. The upper part corresponds to a $\mathcal{D}$-snake, and the lower part corresponds to a T-yardstick. The distances between named elements are important.

Definition 7. *Let* $\mathcal{D} \coloneqq (\mathrm{Col}, \mathrm{T}, \square)$ *be a domino tiling system and* $\mathtt{cbra}$ *be an individual name. An interpretation* $\mathcal{I}$ *is a* $\mathcal{D}$-metricobra *if it satisfies:*

(MInit) *$\mathcal{I}$ is a $\mathcal{D}$-pseudosnake and a* T*-yardstick, and $\mathtt{cbra}^{\mathcal{I}}$ has precisely two successors: one r-successors, namely $\mathtt{ld}^{\mathcal{I}}$, and one s-successor, namely $\mathtt{st}^{\mathcal{I}}$.*

(MTile) *For every tile* $\mathrm{t} \in \mathrm{T}$ *and every element* $\mathrm{d} \in \Delta^{\mathcal{I}}$ *that is r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ we have that* d *carries a tile* $\mathrm{t} \in \mathrm{T}$ *if and only if* d *has a unique s-successor and such a successor is equal to $\mathtt{md}_{\mathrm{t}}^{\mathcal{I}}$.*

(MSync) *Let* t *be the tile of $\mathtt{rd}^{\mathcal{I}}$. Then (a) $\mathtt{cbra}^{\mathcal{I}}$ $r^{\#}s^{\#}$-reaches $\mathtt{end}_{\mathrm{t}}^{\mathcal{I}}$ and cannot $r^{\#}s^{\#}$-reach any of $\mathtt{end}_{\mathrm{t}'}^{\mathcal{I}}$ for* $\mathrm{t}' \neq \mathrm{t}$*, (b) $\mathtt{cbra}^{\mathcal{I}}$ cannot $r^{\#}s^{\#}$-reach an elem. that s^{+}-reaches $\mathtt{end}_{\mathrm{t}}^{\mathcal{I}}$, (c) no elem. r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ $r^{\#}s^{\#}$-reaches $\mathtt{end}_{\mathrm{t}}^{\mathcal{I}}$.*

(MVerti) *For all elements* d *that are r^*-reachable from $\mathtt{ld}^{\mathcal{I}}$ and carry a tile* $\mathrm{t} \in \mathrm{T}$ *that is not up-border, we have that there exists a tile* $\mathrm{t}' \in \mathrm{T}$ *such that (a)* t *and* t' *are* V*-compatible, (b)* t *is left-border (resp. right-border) iff* t' *is, and (c)* d *can $r^{\#}s^{\#}$-reach $\mathtt{end}_{\mathrm{t}'}$ but cannot reach $r^{\#}s^{\#}$-reach $\mathtt{end}_{\mathrm{t}''}$ for all* $\mathrm{t}'' \neq \mathrm{t}'$*.*

We first provide an $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$-axiomatization of $\mathcal{D}$-metricobras.

Lemma 8. *There exists an $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$-concept $\mathrm{C}^{\mathcal{D}}$ such that for all interpretations $\mathcal{I}$ we have that $\mathcal{I}$ is a $\mathcal{D}$-metricobra if and only if $(\mathrm{C}^{\mathcal{D}})^{\mathcal{I}} = \{\mathtt{cbra}^{\mathcal{I}}\}$.*

Second, we relate $\mathcal{D}$-snakes and $\mathcal{D}$-metricobras as follows.

Lemma 9. *Every $\mathcal{D}$-metricobra is also a $\mathcal{D}$-snake. Moreover, if a $\mathcal{D}$-snake exists then so does a $\mathcal{D}$-metricobra.*

By collecting all previous lemmas we infer the main theorem of the paper:

Theorem 10. *A tiling system $\mathcal{D}$ is solvable iff the $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$-concept $\mathrm{C}^{\mathcal{D}}$ is satisfiable. Thus, the concept satisfiability problem of $\mathcal{ALCO}^{r\#s\#}_{\mathsf{reg}}$ is undecidable.*

4 Querying in $\mathcal{ALC}_{\mathsf{vpl}}$

We next address the problem of query entailment under logical constraints. The $\mathbb{C}$-*enriched Positive Existential Queries* (abbreviated as $\mathbb{C}$-PEQs) are defined with:

$$q, q' ::= \bot \mid \mathrm{A}(x) \mid r(x,y) \mid \mathcal{L}(x,y) \mid q \vee q' \mid q \wedge q',$$

where $\mathrm{A} \in \mathbf{N_C}$, $r \in \mathbf{N_R}$, $\mathcal{L} \in \mathbb{C}$, and x, y are variables from a countably infinite set $\mathbf{N_V}$. The semantics is defined as expected, *e.g.* $\mathcal{L}(x,y)$ evaluates to true under a variable assignment η if and only if $\eta(x)$ can $\mathcal{L}$-reach $\eta(y)$ in $\mathcal{I}$. The $\emptyset$-PEQs (or Positive Existential Queries) are well-known generalizations of (unions of) *conjunctive queries*, *i.e.* PEQs in which disjunction is allowed only at the outermost level. The $\mathbb{REG}$-PEQs (or Positive Regular Path Queries) are among the most popular query languages nowadays [14,30]. Finally, $\mathbb{VPL}$-PEQs recently received some attention in [27]. An interpretation $\mathcal{I}$ *satisfies* a query q (written $\mathcal{I} \models q$), if there exists an assignment η of variables (a *match*) from q to $\Delta^{\mathcal{I}}$ under which q evaluates to true. A concept C entails a query q (written $\mathrm{C} \models q$) if all models of C satisfy q. In the $\mathbb{C}$-*PEQ entailment problem* for a DL $\mathcal{L}$ we ask, given an $\mathcal{L}$-concept C and a $\mathbb{C}$-PEQ q, whether $\mathrm{C} \models q$ holds.

By existing results on querying $\mathcal{ALC}$ [17, Lemma 8] and by the tree model property of $\mathcal{ALC}_{\mathsf{vpl}}$ [28, Sec. 4.1], we obtain:

Theorem 11. *The entailment problem of* $\mathbb{REG}$*-PEQs over* $\mathcal{ALC}_{\mathsf{vpl}}$*-concepts is complete for* 2ExpTime.

Unfortunately, the relatively positive results of Theorem 11 do not transfer beyond the class of $\mathbb{REG}$-PEQs, especially if atoms of the form $r^{\#}s^{\#}(x,y)$ are present in the query. To justify this claim, we are going to provide a reduction from the *Octant Tiling Problem* [7, Sec 3.1]. Roughly speaking, the ontology in our reduction will define a grid labelled with tiles, while the query counterpart will serve as a tool to detect mismatches in its lower triangle (a.k.a. octant) part.

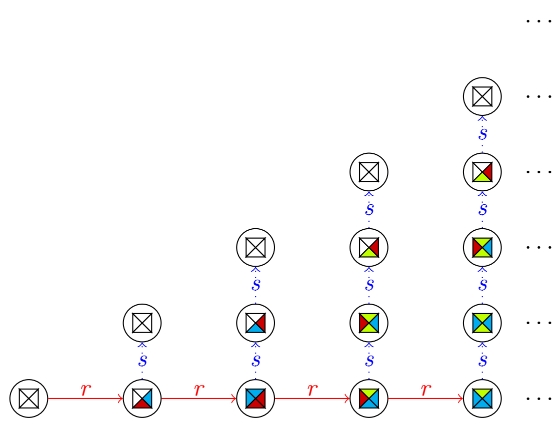


Fig. 3: Visualisation of an octant-based interpretation.

Let $\mathcal{D} := (\mathrm{Col}, \mathrm{T}, \square)$ be a domino tiling system (defined as in Section 3), and let us call the set $\mathbb{O} := \{(x,y) \mid x,y \in \mathbb{N}, 0 \leq y \leq x\}$ *the octant*. It is convenient for our reduction to assume that T contains an all-border white tile $\boxtimes$, and all other tiles from T are not right-border and not down-border. We say that $\mathcal{D}$ *covers* $\mathbb{O}$ if there exists a mapping $\xi\colon \mathbb{O} \to \mathrm{T}$ such that for all pairs $(x,y) \in \mathbb{O}$ satisfy:

(OBord) $\xi(0,0) = \boxtimes$, and $\xi(1,0) \neq \boxtimes$.

(OHori) The tiles $\xi(x,y)$ and $\xi(x{+}1,y)$ are H-compatible. In addition, whenever $\xi(x,y) = \boxtimes$ holds, the tile $\xi(x{+}1,y)$ is left- and up-border.

(OVerti) If $(x,y{+}1) \in \mathbb{O}$ then $\xi(x,y)$ and $\xi(x,y{+}1)$ are V-compatible.

Note that $\mathcal{D}$ covers $\mathbb{O}$ if and only if it covers $\mathbb{N} \times \mathbb{N}$, which is a consequence of (OBord) and the specific use of white colour by tiles in T. The *octant tiling problem* asks to decide, for an input domino tiling system $\mathcal{D}$, whether $\mathcal{D}$ covers the octant. This problem can easily be shown undecidable, as discussed in [7].

We again employ concepts from $\mathrm{C}^{\mathrm{T}}_{🧩}$, and the non-regular language $r^{\#}s^{\#}$. We call a pointed interpretation $(\mathcal{I},(0,0))$ *octant-based* if (i) $\Delta^{\mathcal{I}} = \mathbb{O}$, (ii) $r^{\mathcal{I}} = \{((n,0),(n{+}1,0)) \mid n \in \mathbb{N}\}$, (iii) $s^{\mathcal{I}} = \{((n,m),(n,m{+}1)) \mid n,m \in \mathbb{N}, m < n\}$, and (iv) for every $\mathrm{e} \in \Delta^{\mathcal{I}}$ there is a unique $\mathrm{t} \in \mathrm{T}$ for which $\mathrm{e} \in \mathrm{C}^{\mathcal{I}}_{\mathrm{t}}$. Consult Figure 3 for a visualization. An octant-based $\mathcal{I}$ naturally encodes a mapping $\xi\colon \mathbb{O} \to \mathrm{T}$ defined as $(n,m) \mapsto \mathrm{t}$ for the unique tile carried by (n,m). For convenience, we say that $\mathcal{I}$ $\mathcal{D}$*-semicovers* (resp. $\mathcal{D}$*-covers*) the octant if such a map ξ satisfies (OBord) and (OVerti) (resp. all (OBord), (OVerti), and (OHori)). We analogously speak about *grid-based* pointed interpretations, which are defined similarly to octant-based interpretations above, with the exception that their domains are $\mathbb{N} \times \mathbb{N}$ and the condition $m < n$ is removed from Item (iii).

Violations of the condition (OHori) by an octant-based interpretation will be detected with a $\mathbb{VPL}$-PEQ $q^{\mathcal{D}}_{\blacktriangle}$ (to be defined next), which we visualize as follows.

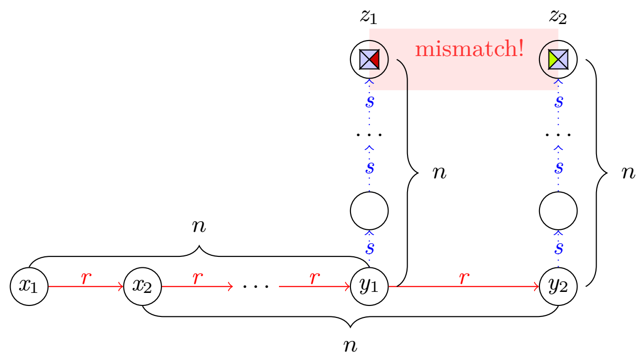


Fig. 4: Visualisation of the query $q^{\mathcal{D}}_{\blacktriangle}(x_1, x_2, y_1, y_2, z_1, z_2)$. The variables z_1, z_2 are mapped to elements that carry tiles violating (OHori); the fact that x_1 and x_2 lie in consecutive columns is handled by means of r-connectedness of y_1, y_2; finally, equi-height of z_1 and z_2 is ensured with non-regular atoms $r^{\#}s^{\#}(x_i, z_i)$.

After the informal explanation, we provide the formal definition of $q^{\mathcal{D}}_{\blacktriangle}$.

$$q^{\mathcal{D}}_{\blacktriangle} \coloneqq \bigvee_{\mathrm{t},\mathrm{t}' \text{ violating (OHori)}} \big[r(x_1, x_2) \wedge r^*(x_2, y_1) \wedge r(y_1, y_2) \wedge s^*(y_1, z_1) \wedge s^*(y_2, z_2)$$
$$\wedge r^{\#}s^{\#}(x_1, z_1) \wedge r^{\#}s^{\#}(x_2, z_2) \wedge \mathrm{C}_{\mathrm{t}}(z_1) \wedge \mathrm{C}_{\mathrm{t}'}(z_2) \big]$$

By routine case analysis with a bit of calculations, we can show that:

Lemma 12. *Let* $\mathcal{D} \coloneqq (\mathrm{Col}, \mathrm{T}, \square)$ *be a domino tilling system. If* $\mathcal{D}$ *covers the octant, then there exist octant-based and grid-based interpretations* $\mathcal{D}$*-covering the octant. Moreover, for all octant-based or grid-based* $\mathcal{I}$ *that* $\mathcal{D}$*-semicover the octant,* $\mathcal{I} \not\models q^{\mathcal{D}}_{\blacktriangle}$ *if and only if* $\mathcal{I}$ *actually* $\mathcal{D}$*-covers the octant.*

It is routine to define a $\mathcal{ALC}_{\mathsf{reg}}$-concept $\mathrm{C}^{\mathrm{T}}_{\mathrm{semicov}}$ stating that the starting element carries $\boxtimes$, that every element carries exactly one tile, and that the tiles of s-connected elements are V-compatible. Expanding $\mathrm{C}^{\mathrm{T}}_{\mathrm{semicov}}$ with an $\mathcal{ALC}_{\mathsf{reg}}$-concept expressing that any element has an r-successor and an s-successor, leads to a concept $\mathrm{C}^{\mathcal{D}}_{\blacksquare}$. This concept is especially useful as it defines grids that $\mathcal{D}$-semicovers the octant. (We note that the use of grids is crucial here, as $\mathcal{ALC}_{\mathsf{reg}}$ cannot define octant-based structures but our queries look only at octants.)

The main property of our reduction is established below.

Lemma 13. *Let* $\mathcal{D} \coloneqq (\mathrm{Col}, \mathrm{T}, \square)$ *be a domino tilling system. Then* $\mathrm{C}^{\mathcal{D}}_{\blacksquare} \not\models q^{\mathcal{D}}_{\blacktriangle}$ *if and only if there is a grid-based interpretation* $\mathcal{I}$ *such that* $\mathcal{I} \models \mathrm{C}^{\mathcal{D}}_{\blacksquare}$ *and* $\mathcal{I} \not\models q^{\mathcal{D}}_{\blacktriangle}$. *Thus* $\mathrm{C}^{\mathcal{D}}_{\blacksquare} \not\models q^{\mathcal{D}}_{\blacktriangle}$ *if and only if* $\mathcal{D}$ *covers the octant.*

The concept $\mathrm{C}^{\mathcal{D}}_{\blacksquare}$ can be equivalently expressed as an $\mathcal{ALC}$-TBox (*cf.* [2, Sec. 2.2.1]). By a combination of previously presented lemmas we thus infer:

Theorem 14. *The* $\mathbb{VPL}$*-PEQs entailment problem for* $\mathcal{ALC}_{\mathsf{reg}}$ *is undecidable. This holds already for* $\{r, r^*, s, s^*, r^{\#}s^{\#}\}$*-PEQ entailment over* $\mathcal{ALC}$*-TBoxes.*

5 The case of functionality

As the next result, we show that $\mathcal{ALCF}_{\mathsf{vpl}}$, *i.e.* the visibly-pushdown extension of $\mathcal{ALC}_{\mathsf{reg}}$ with functionality, has a decidable concept satisfiability problem. To do so, we provide a rather straightforward translation from $\mathcal{ALCF}_{\mathsf{vpl}}$ into $\mathcal{ALC}_{\mathsf{vpl}}$, based on [9, Sec. 3]. Note that it is not a priori obvious that such a reduction works in the case of $\mathcal{ALC}_{\mathsf{vpl}}$, as visibly-pushdown languages are not, in contrast to regular languages, compositional. Moreover, functionality violates the crucial decidability condition of "unique diamond-path property" from [28, Def. 11].

An $\mathcal{ALCF}_{\mathsf{vpl}}$-concept is in *Scott's normal form* if it conforms to the pattern:

$$\mathrm{A} \sqcap \forall\Sigma^*. \left(\mathrm{D} \sqcap \prod_i (\mathrm{A}_i \to \exists\mathcal{L}.\mathrm{B}_i) \sqcap \prod_j (\mathrm{A}_j \to \forall\mathcal{L}.\mathrm{B}_j)\right) \sqcap \prod_{r \in \mathsf{R}} \mathsf{func}(r),$$

where (possibly decorated) A, B are concept names, D is an $\mathcal{ALC}$-concept that does not contain existential and universal restrictions, Σ is the set of all role names occurring in the original concept, and R is a finite set of role names from $\mathbf{N_R}$. By a routine renaming technique by Scott [32] we know that for any $\mathcal{ALCF}_{\mathsf{vpl}}$-concept there exists a polynomially computable, equisatisfiable $\mathcal{ALCF}_{\mathsf{vpl}}$-concept is Scott's normal form. The *closure* $\mathsf{cl}(\mathrm{C})$ of an $\mathcal{ALC}_{\mathsf{vpl}}$-concept C is the set composed of all concept names appearing in C, all existential restrictions of the form $\exists r.\mathrm{D}$, $\exists\mathcal{L}.\mathrm{D}$ from C, universal restrictions $\forall r.\mathrm{D}$, $\forall\mathcal{L}.\mathrm{D}$ appearing in C, and their (single) negations. The size of $\mathsf{cl}(\mathrm{C})$ is polynomial w.r.t the size of C.

Lemma 15. *An* $\mathcal{ALCF}_{\mathsf{vpl}}$*-concept* C *in Scott's normal is satisfiable iff*

$$\mathsf{trans}(\mathrm{C}) \coloneqq \mathrm{C}^- \sqcap \prod_{r \in \mathsf{R}} \prod_{\mathrm{D} \in \mathsf{cl}(\mathrm{C})} \forall\Sigma^*. ((\exists r.\mathrm{D}) \to (\forall r.\mathrm{D}))$$

is satisfiable, where C^- *is obtained by removing functionality restriction from* C.

The presented translation is clearly polynomial w.r.t the input concept. Thus, by 2ExpTime-completeness of $\mathcal{ALC}_{\mathsf{vpl}}$ [28, Thm. 18–19], we conclude:

Theorem 16. *The satisfiability problem for* $\mathcal{ALCF}_{\mathsf{vpl}}$ *is 2ExpTime-complete.*

We think that decidability of extensions of $\mathcal{ALC}_{\mathsf{vpl}}$ with "local" functionality statements or qualified number restrictions ($\mathcal{Q}$) (under the unary encoding of counters), can be established by means of a similar translation.

6 Seemingly innocent Self operator

We conclude the paper by showing yet another negative result. This time we tackle the Self operator, a modelling feature supported by two profiles of the OWL 2 Web Ontology Language [24,26] and $\mathcal{SROIQ}$. Recall that the Self operator allows us to specify the situation when an element is related to it*self* by a binary relationship, *i.e.* we interpret the concept $\exists r.\mathsf{Self}$ in an interpretation $\mathcal{I}$ as the set of all those elements d for which (d, d) belongs to $r^{\mathcal{I}}$. In what follows, we provide a reduction from an undecidable problem of non-emptiness of the intersection of deterministic one-counter automata (DOCA) [34, p. 75]. Such an automata model is similar to pushdown automata, but its stack alphabet is single-letter only. The Self operator will be especially useful to introduce "disjunction" to paths.

Let Σ be an alphabet and $\mathtt{w} := (\mathtt{a}_1, \star_1) \ldots (\mathtt{a}_n, \star_n)$ be a word over $\Sigma \times \{c, r, i\}$. We call the word $\pi_1(\mathtt{w}) := \mathtt{a}_1 \ldots \mathtt{a}_n$ the *projection* of $\mathtt{w}$. An important property of DOCA is that they can be made visibly one-counter in the following sense.

Lemma 17. *For any DOCA $\mathcal{A}$ over Σ, we can construct a VOCA $\tilde{\mathcal{A}}$ over $\tilde{\Sigma} := (\Sigma \times \{c\}, (\Sigma \times \{i\}) \cup \{\mathtt{x}\}, \Sigma \times \{r\})$ where $\mathtt{x}$ is a fresh internal letter, such that all words in $\mathcal{L}(\tilde{\mathcal{A}})$ have the form $\tilde{\mathtt{a}_1}\mathtt{x}\tilde{\mathtt{a}_2}\mathtt{x} \ldots \mathtt{x}\tilde{\mathtt{a}_n}$ for $\tilde{\mathtt{a}_1}, \ldots, \tilde{\mathtt{a}_n} \in \Sigma \times \{c, i, r\}$, and $\mathcal{L}(\mathcal{A}) = \{\pi_1(\tilde{\mathtt{w}}) \mid \tilde{\mathtt{w}} := \tilde{\mathtt{a}_1} \ldots \tilde{\mathtt{a}_n},\ \ \tilde{\mathtt{a}_1}\mathtt{x} \ldots \mathtt{x}\tilde{\mathtt{a}_n} \in \mathcal{L}(\tilde{\mathcal{A}})\}$ holds.*

We fix a finite alphabet $\Sigma \subseteq \mathbf{N_R}$. Moreover, fix two deterministic one-counter automata $\mathcal{A}_1$ and $\mathcal{A}_2$ over Σ, as well as deterministic one-counter automata $\mathcal{C}_1$ and $\mathcal{C}_2$ recognizing the complement of their languages (they can be constructed as DOCA are closed under complement). Finally, we construct their visibly-one-counter counterparts $\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2, \tilde{\mathcal{C}}_1, \tilde{\mathcal{C}}_2$ over the pushdown alphabet $\tilde{\Sigma}$, as provided by Lemma 17. We stress that the letter $\mathtt{x}$, playing the role of a "separator", is identical for all of the aforementioned visibly-one-counter automata. We also point out that the non-emptiness of $\mathcal{L}(\tilde{\mathcal{A}}_1) \cap \mathcal{L}(\tilde{\mathcal{A}}_2)$ is not equivalent to the non-emptiness of $\mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_1)$, as the projection of a letter $\mathtt{a} \in \tilde{\Sigma}$ may be used by $\mathcal{A}_1$ and $\mathcal{A}_2$ in different contexts (*e.g.* both as a call or as a return).

We are going to encode words accepted by one-counter automata by means of word-like interpretations. A pointed interpretation $(\mathcal{I}, \mathrm{d})$ is Σ*-friendly* if for every element $\mathrm{e} \in \Delta^{\mathcal{I}}$ that is $\mathtt{x}^*$-reachable from d in $\mathcal{I}$ there exists a unique letter $\mathtt{a} \in \Sigma$ so that e carries $\tilde{\mathtt{a}}$-self-loops for all $\tilde{\mathtt{a}} \in \tilde{\Sigma}$ with $\pi_1(\tilde{\mathtt{a}}) = \mathtt{a}$, and no self-loops for all other letters in $\tilde{\Sigma}$ (also including the "separator letter" $\mathtt{x}$).

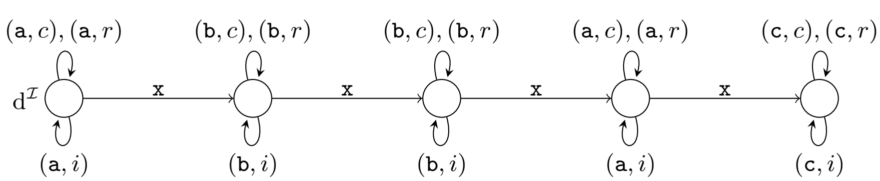


Fig. 5: An example Σ-friendly pointed $(\mathcal{I}, \mathrm{d})$ encoding the word $\mathtt{abbac}$.

Σ-friendly interpretations can easily be axiomatized with an $\mathcal{ALC}^{\mathsf{Self}}$-concept $\mathrm{C}_{\mathrm{fr}}^{\Sigma}$:

$$\mathrm{C}_{\mathrm{fr}}^{\Sigma} \coloneqq \forall \mathtt{x}^*. \bigsqcup_{\mathtt{a}\in\Sigma} \bigsqcap_{\mathtt{b}\neq\mathtt{a},\mathtt{b}\in\Sigma,\pi_1(\tilde{\mathtt{a}})=\mathtt{a},\pi_1(\tilde{\mathtt{b}})=\mathtt{b}} \Big([\exists\tilde{\mathsf{a}}.\mathsf{Self}] \sqcap \neg[\exists\tilde{\mathsf{b}}.\mathsf{Self}] \sqcap \neg[\exists \mathsf{x}.\mathsf{Self}] \Big).$$

Moreover, every $\mathtt{x}^*$-path ρ in a Σ-friendly $\mathcal{I}$ *represents* a word in Σ^* in the following sense: the i-th letter of such a word is $\mathtt{a}$ if and only if the i-th element of the path carries an $(\mathtt{a}, c)$-self-loop. This is well-defined, as every element in Σ-friendly $\mathcal{I}$ carries a $(\mathtt{a}, c)$-self-loop for a unique letter $\mathtt{a} \in \Sigma$. Consult Figure 5.

As a special class of Σ-friendly interpretations we consider Σ-metawords. We say that $(\mathcal{I}, \mathrm{d})$ is a *Σ-metaword* if it is a Σ-friendly interpretation of the domain $\mathbb{Z}_n$ for some positive $n \in \mathbb{N}$, the role name $\mathtt{x}$ is interpreted as the set $\{(i, i{+}1) \mid 0 \leq i \leq n{-}2\}$, and all other role names are either interpreted as $\emptyset$ or are subsets of the diagonal $\{(i, i) \mid i \in \mathbb{Z}_n\}$ (or, put differently, they appear only as self-loops). The example Σ-friendly $\mathcal{I}$ from Figure 5 is actually a Σ-metaword. It is not too hard to see that for every word $\mathtt{w} \in \Sigma^+$ there is a Σ-metaword representing $\mathtt{w}$. A crucial observation regarding Σ-metawords is as follows. If an element starting a Σ-metaword can $\{\tilde{\mathtt{w}}\}$-reach some element (for some $\tilde{\mathtt{w}}$ in the language of $\tilde{\mathcal{A}}_1$), then the path ρ witnessing this fact satisfies $\rho_i = \rho_{i+1}$ for all odd indices i and $\rho_i + 1 = \rho_{i+1}$ for all even indices i. Similar remarks apply to Σ-friendly interpretations but the correspondence is not as elegant anymore.

As the next step of the construction, we are going to decorate Σ-friendly interpretations with extra information on whether or not words represented by paths are accepted by $\mathcal{A}_1$. This is achieved by means of the following concept

$$\mathrm{C}_{\mathcal{A}_1} \coloneqq \mathrm{C}_{\mathrm{fr}}^{\Sigma} \sqcap \forall \mathcal{L}(\tilde{\mathcal{A}}_1).\mathrm{Acc}_{\mathcal{A}_1} \sqcap \forall \mathcal{L}(\tilde{\mathcal{C}}_1).\neg\mathrm{Acc}_{\mathcal{A}_1},$$

for a fresh concept name $\mathrm{Acc}_{\mathcal{A}_1}$. We define $\mathrm{C}_{\mathcal{A}_2}$ analogously. We have that:

Lemma 18. *If $\mathrm{C}_{\mathcal{A}_1}$ is satisfied by a Σ-friendly pointed interpretation $(\mathcal{I}, \mathrm{d})$, then for every element $\mathrm{e} \in \Delta^{\mathcal{I}}$ that is $\mathtt{x}^*$-reachable from d via a path ρ we have that $\mathrm{e} \in (\mathrm{Acc}_{\mathcal{A}_1})^{\mathcal{I}}$ iff the Σ-word represented by ρ belongs to $\mathcal{L}(\mathcal{A}_1)$. Moreover, after reinterpreting the concept $\mathrm{Acc}_{\mathcal{A}_1}$, every Σ-metaword becomes a model of $\mathrm{C}_{\mathcal{A}_1}$.*

Lemma 19. $\mathrm{C}_{\mathcal{A}_1} \sqcap \mathrm{C}_{\mathcal{A}_2} \sqcap \exists \mathtt{x}^*. (\mathrm{Acc}_{\mathcal{A}_1} \sqcap \mathrm{Acc}_{\mathcal{A}_2})$ *is satisfiable iff* $\mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2) \neq \emptyset$.

By the undecidability of the non-emptiness problem for intersection of one-counter languages [34, p. 75], we conclude the last theorem of the paper.

Theorem 20. *The concept satisfiability problem for $\mathcal{ALC}_{\mathsf{vpl}}^{\mathsf{Self}}$ is undecidable, even if only visibly-one-counter languages are allowed in concepts.*

We stress that there is nothing special about DOCA used in the proof. In fact, any automata model would satisfy our needs as long as it would (i) have undecidable non-emptiness problem for the intersection of languages, (ii) enjoy the analogue of Lemma 17, and (iii) be closed under complement. We leave it as an open problem to see if there exists a *single* visibly-pushdown language $\mathcal{L}$ that makes the concept satisfiability of $\mathcal{ALC}_{\mathsf{reg}+\mathcal{L}}^{\mathsf{Self}}$ undecidable. Note that our proof heavily relied on the availability of multiple visibly-one-counter languages.

Acknowledgements

This work was supported by the ERC Consolidator Grant No. 771779 (DeciGUT). Snake and cobra icons were downloaded from Icons8 and Flaticon.

The author would like to thank Reijo Jaakkola for many inspiring discussions; Sebastian Rudolph for pinpointing [9] and suggesting many improvements, especially related to the previous versions of Section 4; Witold Charatonik for very careful proofreading and his pedantic approach to writing; as well as Alessio Mansutti and Emanuel Kieroński for their help polishing the introduction.

References

1. Alur, R., Madhusudan, P.: Adding Nesting Structure to Words. Journal of ACM **56**(3) (may 2009)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
3. Bárány, V., Löding, C., Serre, O.: Regularity Problems for Visibly Pushdown Languages. In: Durand, B., Thomas, W. (eds.) STACS 2006, 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, France, February 23-25, 2006, Proceedings. Lecture Notes in Computer Science, vol. 3884, pp. 420–431. Springer (2006)
4. Bednarczyk, B., Kieronski, E.: Finite entailment of local queries in the Z family of description logics. In: Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022. pp. 5487–5494. AAAI Press (2022)
5. Bednarczyk, B., Rudolph, S.: Worst-case optimal querying of very expressive description logics with path expressions and succinct counting. In: Kraus, S. (ed.) Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019. pp. 1530–1536. ijcai.org (2019)
6. Bienvenu, M., Calvanese, D., Ortiz, M., Simkus, M.: Nested Regular Path Queries in Description Logics. In: Baral, C., De Giacomo, G., Eiter, T. (eds.) Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014, Vienna, Austria, July 20-24, 2014. AAAI Press (2014)
7. Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G.: Undecidability of the Logic of Overlap Relation over Discrete Linear Orderings. Electronic Notes in Theoretical Computer Science **262**, 65–81 (2010), proceedings of the 6th Workshop on Methods for Modalities (M4M-6 2009)

8. Bruse, F., Lange, M.: A Decidable Non-Regular Modal Fixpoint Logic. In: Haddad, S., Varacca, D. (eds.) 32nd International Conference on Concurrency Theory, CONCUR 2021, August 24-27, 2021, Virtual Conference. LIPIcs, vol. 203, pp. 23:1–23:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2021)
9. Calvanese, D., De Giacomo, G., Rosati, R.: A Note on Encoding Inverse Roles and Functional Restrictions in ALC Knowledge Bases. In: Franconi, E., De Giacomo, G., MacGregor, R.M., Nutt, W., Welty, C.A. (eds.) Proceedings of the 1998 International Workshop on Description Logics (DL'98), IRST, Povo - Trento, Italy, June 6-8, 1998. CEUR Workshop Proceedings, vol. 11. CEUR-WS.org (1998)
10. Calvanese, D., Eiter, T., Ortiz, M.: Regular Path Queries in Expressive Description Logics with Nominals. In: Boutilier, C. (ed.) IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009. pp. 714–720 (2009)
11. Calvanese, D., Ortiz, M., Simkus, M.: Verification of Evolving Graph-structured Data under Expressive Path Constraints. In: Martens, W., Zeume, T. (eds.) 19th International Conference on Database Theory, ICDT 2016, Bordeaux, France, March 15-18, 2016. LIPIcs, vol. 48, pp. 15:1–15:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2016)
12. De Giacomo, G., Lenzerini, M.: Boosting the Correspondence between Description Logics and Propositional Dynamic Logics. In: Hayes-Roth, B., Korf, R.E. (eds.) Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 - August 4, 1994, Volume 1. pp. 205–212. AAAI Press / The MIT Press (1994)
13. Fischer, M.J., Ladner, R.E.: Propositional Dynamic Logic of Regular Programs. Journal of Computer and System Sciences **18**(2), 194–211 (1979)
14. Florescu, D., Levy, A.Y., Suciu, D.: Query Containment for Conjunctive Queries with Regular Expressions. In: Mendelzon, A.O., Paredaens, J. (eds.) Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 1-3, 1998, Seattle, Washington, USA. pp. 139–148. ACM Press (1998)
15. Göller, S.: Computational Complexity of Propositional Dynamic Logics. Ph.D. thesis, University of Leipzig (2008), https://d-nb.info/99245168X
16. Grau, B.C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., Sattler, U.: OWL 2: The Next Step for OWL. Journal of Web Semantics **6**(4), 309–322 (2008)
17. Gutiérrez-Basulto, V., Ibáñez-García, Y., Jung, J.C., Murlak, F.: Answering regular path queries mediated by unrestricted SQ ontologies. Artif. Intell. **314**, 103808 (2023)
18. Harel, D., Paterson, M.: Undecidability of PDL with L $= \{a^{2^i} \mid i \geq 0\}$. Journal of Computer and System Sciences **29**(3), 359–365 (1984)
19. Harel, D., Pnueli, A., Stavi, J.: Further Results on Propositional Dynamic Logic of Nonregular Programs. In: Kozen, D. (ed.) Logics of Programs, Workshop, Yorktown Heights, New York, USA, May 1981. Lecture Notes in Computer Science, vol. 131, pp. 124–136. Springer (1981)
20. Harel, D., Pnueli, A., Stavi, J.: Propositional Dynamic Logic of Context-Free Programs. In: 22nd Annual Symposium on Foundations of Computer Science (SFCS 1981). pp. 310–321. IEEE (1981)
21. Harel, D., Raz, D.: Deciding Properties of Nonregular Programs. SIAM Journal on Computing **22**(4), 857–874 (1993)
22. Harel, D., Singerman, E.: More on Nonregular PDL: Finite Models and Fibonacci-Like Programs. Inf. Comput. **128**(2), 109–118 (1996)

23. Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., Rudolph, S.: OWL 2 Web Ontology Language Primer (Second Edition). World Wide Web Consortium (W3C) (December 2012), https://www.w3.org/TR/owl2-primer/
24. Horrocks, I., Kutz, O., Sattler, U.: The Even More Irresistible SROIQ. In: Doherty, P., Mylopoulos, J., Welty, C.A. (eds.) Proceedings, Tenth International Conference on Principles of Knowledge Representation and Reasoning, Lake District of the United Kingdom, June 2-5, 2006. pp. 57–67. AAAI Press (2006)
25. Koren, T., Pnueli, A.: There Exist Decidable Context Free Propositional Dynamic Logics. In: Proceedings of the Carnegie Mellon Workshop on Logic of Programs. p. 290–312. Springer-Verlag, Berlin, Heidelberg (1983)
26. Krötzsch, M., Rudolph, S., Hitzler, P.: ELP: Tractable Rules for OWL 2. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T.W., Thirunarayan, K. (eds.) The Semantic Web - ISWC 2008, 7th International Semantic Web Conference, ISWC 2008, Karlsruhe, Germany, October 26-30, 2008. Proceedings. Lecture Notes in Computer Science, vol. 5318, pp. 649–664. Springer (2008)
27. Lange, M., Lozes, É.: Conjunctive Visibly-Pushdown Path Queries. In: Kosowski, A., Walukiewicz, I. (eds.) Fundamentals of Computation Theory - 20th International Symposium, FCT 2015, Gdańsk, Poland, August 17-19, 2015, Proceedings. Lecture Notes in Computer Science, vol. 9210, pp. 327–338. Springer (2015)
28. Löding, C., Lutz, C., Serre, O.: Propositional Dynamic Logic with Recursive Programs. Journal of Logical and Algebraic Methods in Programming **73**(1-2), 51–69 (2007)
29. Ortiz, M.: Query Answering in Expressive Description Logics: Techniques and Complexity Results. Ph.D. thesis, Technische Universität Wien (2010)
30. Ortiz, M., Simkus, M.: Reasoning and Query Answering in Description Logics. In: Eiter, T., Krennwallner, T. (eds.) Reasoning Web. Semantic Technologies for Advanced Query Answering - 8th International Summer School 2012, Vienna, Austria, September 3-8, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7487, pp. 1–53. Springer (2012)
31. Pratt-Hartmann, I.: Fragments of First-Order Logic. Oxford University Press (2023)
32. Scott, D.: A Decision Method for Validity of Sentences in Two Variables. Journal of Symbolic Logic **27**, 377 (1962)
33. Sipser, M.: Introduction to the Theory of Computation. Course Technology, Boston, MA, third edn. (2013)
34. Valiant, L.: Decision Procedures for Families of Deterministic Pushdown Automata. Ph.D. thesis, University of Warwick (1973)