



Using Model Theory to Find Decidable and Tractable Description Logics with Concrete Domains

Franz Baader¹ · Jakub Rydval¹

Received: 28 April 2021 / Accepted: 22 February 2022 / Published online: 4 May 2022
© The Author(s) 2022

Abstract

Concrete domains have been introduced in the area of Description Logic to enable reference to concrete objects (such as numbers) and predefined predicates on these objects (such as numerical comparisons) when defining concepts. Unfortunately, in the presence of general concept inclusions (GCIs), which are supported by all modern DL systems, adding concrete domains may easily lead to undecidability. To regain decidability of the DL \mathcal{ALC} in the presence of GCIs, quite strong restrictions, in sum called ω -admissibility, were imposed on the concrete domain. On the one hand, we generalize the notion of ω -admissibility from concrete domains with only binary predicates to concrete domains with predicates of arbitrary arity. On the other hand, we relate ω -admissibility to well-known notions from model theory. In particular, we show that finitely bounded homogeneous structures yield ω -admissible concrete domains. This allows us to show ω -admissibility of concrete domains using existing results from model theory. When integrating concrete domains into lightweight DLs of the \mathcal{EL} family, achieving decidability is not enough. One wants reasoning in the resulting DL to be tractable. This can be achieved by using so-called p-admissible concrete domains and restricting the interaction between the DL and the concrete domain. We investigate p-admissibility from an algebraic point of view. Again, this yields strong algebraic tools for demonstrating p-admissibility. In particular, we obtain an expressive numerical p-admissible concrete domain based on the rational numbers. Although ω -admissibility and p-admissibility are orthogonal conditions that are almost exclusive, our algebraic characterizations of these two properties allow us to locate an infinite class of p-admissible concrete domains whose integration into \mathcal{ALC} yields decidable DLs.

Keywords Description logic · Concrete domains · GCIs · ω -Admissibility · Homogeneity · Finite boundedness · Decidability · P-admissibility · Convexity · ω -Categoricity · Tractability · Constraint satisfaction

✉ Franz Baader
franz.baader@tu-dresden.de

Jakub Rydval
jakub.rydval@tu-dresden.de

¹ Theoretical Computer Science, TU Dresden, Dresden, Germany

1 Introduction

Description Logics (DLs) [10, 12] are a well-investigated family of logic-based knowledge representation languages, which are frequently used to formalize ontologies for application domains such as the Semantic Web [52] or biology and medicine [51]. To define the important notions of such an application domain as formal concepts, DLs state necessary and sufficient conditions for an individual to belong to a concept. These conditions can be Boolean combinations of atomic properties required for the individual (expressed by concept names) or properties that refer to relationships with other individuals and their properties (expressed as role restrictions). For example, the concept of a father that has only daughters can be formalized by the concept description $C_{ex} := \neg Female \sqcap \exists child.Human \sqcap \forall child.Female$, which uses the concept names *Female* and *Human* and the role name *child* as well as the concept constructors negation (\neg), conjunction (\sqcap), existential restriction ($\exists r.D$), and value restriction ($\forall r.D$). The GCIs $Human \sqsubseteq \forall child.Human$ and $\exists child.Human \sqsubseteq Human$ say that humans have only human children, and they are the only ones that can have human children.

DL systems provide their users with reasoning services that allow them to derive implicit knowledge from the explicitly represented one. In our example, the above GCIs imply that elements of our concept C_{ex} also belong to the concept $D_{ex} := Human \sqcap \forall child.Human$, i.e., C_{ex} is subsumed by D_{ex} w.r.t. these GCIs. A specific DL is determined by which kind of concept constructors are available. A major goal of DL research was and still is to find a good compromise between expressiveness and the complexity of reasoning, i.e., to locate DLs that are expressive enough for interesting applications, but still have inference problems (like subsumption) that are decidable and preferably of a low complexity. For the DL \mathcal{ALC} , in which all the concept descriptions used in the above example can be expressed, the subsumption problem w.r.t. GCIs is ExpTime-complete [12].

Classical DLs like \mathcal{ALC} cannot refer to concrete objects and predefined relations over these objects when defining concepts. For example, a constraint stating that parents are strictly older than their children cannot be expressed in \mathcal{ALC} . To overcome this deficit, a scheme for integrating certain well-behaved concrete domains, called admissible, into \mathcal{ALC} was introduced in [2], and it was shown that this integration leaves the relevant inference problems (such as subsumption) decidable. Basically, admissibility requires that the set of predicates of the concrete domain is closed under negation and that the constraint satisfaction problem (CSP) for the concrete domain is decidable. However, in this setting, GCIs were not considered since they were not a standard feature of DLs then,¹ though a combination of concrete domains and GCIs would be useful in many applications. For example, using the syntax employed in [65] and also in the present article, the above constraint regarding the age of parents and their children could be expressed by the GCI $Human \sqcap \exists age, child.age.(x_1 < x_2) \sqsubseteq \perp$, which says that there cannot be a human whose age is smaller than the age of one of his or her children. Here \perp is the bottom concept, which is always interpreted as the empty set, *age* is a feature that maps from the abstract domain populating concepts into the concrete domain of rational numbers, and $<$ is the usual “smaller than” predicate.

A first indication that concrete domains might be harmful for decidability was given in [4], where it was shown that adding transitive closure of roles to the extension of \mathcal{ALC} by an admissible concrete domain based on real arithmetics renders the subsumption problem undecidable. The proof of this result uses a reduction from the Post Correspondence Problem (PCP). It was shown in [63] that this proof can be adapted to the case where transitive closure of roles is replaced by GCIs, and it actually works for considerably weaker concrete domains,

¹ Actually, GCIs were introduced (with a different name) at about the same time as concrete domains [9, 71].

such as the rational numbers \mathbb{Q} or the natural number \mathbb{N} with a unary predicate $=_0$ for equality with zero, a binary equality predicate $=$, and a unary predicate $+_1$ for incrementation. In [6] it is shown, by a reduction from the halting problem of two-register machines, that undecidability even holds without $=$ and $=_0$.

To regain decidability, one option is to impose syntactic restriction on how the DL can interact with the concrete domain [45, 69]. The main idea is here to disallow paths (such as *child age* in our example), which has the effect that concrete domain predicates cannot compare properties (such as the age) of different individuals. Another option is to impose stronger restrictions than admissibility on the concrete domain. After first positive results for specific concrete domains (e.g., a concrete domain over the rational numbers with order and equality [62, 64]), the notion of ω -admissible concrete domains was introduced in [65], and it was shown (by designing a tableau-based decision procedure) that integrating such a concrete domain into \mathcal{ALC} leaves reasoning decidable also in the presence of GCIs. In [6], we generalize the notion of ω -admissibility and the decidability result from concrete domains with only binary predicates as in [65] to concrete domains with predicates of arbitrary arity.

When integrating a concrete domain into a lightweight DL like \mathcal{EL} , one wants to preserve tractability rather than just decidability. To achieve this, the notion of p-admissible concrete domains was introduced in [11] and paths of length > 1 were disallowed in concrete domain restrictions. Regarding the latter condition, note that, in the above example, we have used the path *child age*, which has length 2. The restriction to paths of length 1 means (in our example) that we can no longer compare the ages of different humans, but we can still define concepts like *teenager*, using the GCI $Teenager \sqsubseteq Human \sqcap \exists age. \geq_{10}(x_1) \wedge \leq_{19}(x_1)$, where \geq_{10} and \leq_{19} are unary predicates respectively interpreted as the rational numbers greater equal 10 and smaller equal 19. In a p-admissible concrete domain, satisfiability of conjunctions of atomic formulae and validity of implications between such conjunctions must be tractable. In addition, the concrete domain must be *convex*, which roughly speaking means that a conjunction cannot imply a true disjunction. For example, the concrete domain $(\mathbb{Q}; <, =, >)$ is ω -admissible, but it is not convex since $x < y \wedge x < z$ implies $y < z \vee y = z \vee y > z$, but none of the disjuncts. In [11], two p-admissible concrete domains were exhibited, where one of them is based on \mathbb{Q} with unary predicates $=_p, >_p$ and binary predicates $+_p, =$. To the best of our knowledge, since then no other p-admissible concrete domains have been described in the literature before our work in [8]. Similarly, after the publication of [65] and before our work in [6], no new ω -admissible concrete domains were exhibited. We believe that the reason for this is that it is quite hard to prove ω -admissibility or p-admissibility of a concrete domain without appropriate mathematical tools.

The main contribution of this paper is to develop such tools based on a model-theoretic analysis of the conditions required by these two notions of admissibility. It is based on the conference publications [6] and [8], but differs from them w.r.t. some details and also presents additional results. On the one hand, we show that there is a close relationship between ω -admissibility and well-known notions from model theory. In particular, we prove that finitely bounded homogeneous structures yield ω -admissible concrete domains. This allows us to show ω -admissibility of known such concrete domains (like Allen and RCC8 from [65]; see Example 2) and to locate new ω -admissible concrete domains using existing results from model theory (see Examples 3, 4, and 5). We can even show that some of the relevant properties can be algorithmically tested (see Sect. 6). On the other hand, we devise an algebraic characterization of convexity based on the notion of *square embeddings*, which are embeddings of the second power of a relational structure into itself. We investigate the implications of this characterization further for so-called ω -categorical structures, finitely bounded structures, and numerical structures. Each of these cases provides us with

new examples of p-admissible concrete domains. In particular, we exhibit a new and quite expressive p-admissible concrete domain based on the rational numbers, whose predicates are defined by linear equations over \mathbb{Q} . The paper also investigates the connection between p-admissibility and ω -admissibility. It turns out that only trivial concrete domains can satisfy both properties. However, we can show that convex finitely bounded homogeneous structures, which are p-admissible, can be integrated into \mathcal{ALC} (even without the length 1 restriction on role paths) without losing decidability. Whereas these structures are not ω -admissible, they can be expressed in an ω -admissible concrete domain.

To increase readability of the main text, some of the technical proofs have been moved to an appendix. Due to space constraints, some of the results of [6, 8] are only cited without proof here.

2 Preliminaries

In this section, we introduce the algebraic and logical notions that will be used in the rest of the paper. The set $\{1, \dots, n\}$ is denoted by $[n]$. Given a set A , the *diagonal relation* (or *equality*) on A is defined as the binary relation $\text{Eq}^A := \{(a, a) \mid a \in A\}$. We use the bar notation for tuples; for a tuple \bar{t} indexed by a set I , the value of \bar{t} at the position $i \in I$ is denoted by $\bar{t}[i]$. For a function $f: A^k \rightarrow B$ and n -tuples $\bar{t}_1, \dots, \bar{t}_k \in A^n$, we use the shortcut

$$f(\bar{t}_1, \dots, \bar{t}_k) := (f(\bar{t}_1[1], \dots, \bar{t}_k[1]), \dots, f(\bar{t}_1[n], \dots, \bar{t}_k[n])).$$

From a mathematical point of view, concrete domains are relational structures. A *relational signature* τ is a set of *relation symbols*, each with an associated natural number called *arity*. For a relational signature τ , a *relational τ -structure* \mathfrak{A} (or simply τ -structure or structure) consists of a set A (the *domain*) together with the relations $R^{\mathfrak{A}} \subseteq A^k$ for each relation symbol $R \in \tau$ of arity k . Such a structure \mathfrak{A} is *finite* if its domain A is finite. We often describe structures by listing their domain and relations; e.g., we write $\mathfrak{Q} = (\mathbb{Q}; <)$ for the relational structure whose domain is the set of rational numbers \mathbb{Q} , and which has the usual smaller relation $<$ on \mathbb{Q} as its only relation.²

The *product* of a family $(\mathfrak{A}_i)_{i \in I}$ of τ -structures is the τ -structure $\prod_{i \in I} \mathfrak{A}_i$ over $\prod_{i \in I} A_i$ such that, for each $R \in \tau$ of arity k , we have

$$(\bar{a}_1, \dots, \bar{a}_k) \in R^{\prod_{i \in I} \mathfrak{A}_i} \text{ iff } (\bar{a}_1[i], \dots, \bar{a}_k[i]) \in R^{\mathfrak{A}_i} \text{ for every } i \in I.$$

We denote the binary product of a structure \mathfrak{A} with itself as \mathfrak{A}^2 . An *expansion* of a τ -structure \mathfrak{A} is a σ -structure \mathfrak{B} with $A = B$ such that $\tau \subseteq \sigma$ and $R^{\mathfrak{B}} = R^{\mathfrak{A}}$ for each relation symbol $R \in \tau$. Conversely, we call \mathfrak{A} a *reduct* of \mathfrak{B} . We use the notation $(\mathfrak{A}, R_1, \dots, R_n)$ to describe an expansion of \mathfrak{A} by the relations R_1, \dots, R_n over A ; e.g., (\mathfrak{Q}, \neq) stands for $(\mathbb{Q}; <, \neq)$.

One possibility to obtain an expansion of a τ -structure is to use formulas of *first-order (fo) logic* over the signature τ to define new predicates, where a formula with k free variables defines a k -ary predicate in the obvious way. We say that a first-order formula is k -ary if it has k free variables. For a structure \mathfrak{A} we denote its first-order theory, i.e., the first-order sentences that hold in \mathfrak{A} , with $\text{Th}(\mathfrak{A})$. We assume that equality $=$ as well as the nullary predicate symbol

² By a slight abuse of notation, we use $<$ instead of $<^{\mathfrak{Q}}$ to denote also the interpretation of the predicate symbol $<$ in \mathfrak{Q} .

ff for falsity are always available when building these formulas. Thus, atomic formulas are of the form ff , $x_i = x_j$, and $R(x_1, \dots, x_k)$ for some k -ary $R \in \tau$ and variables x_1, \dots, x_k . In addition to full first-order logic, we also use standard fragments such as the *existential positive* (\exists^+), the *quantifier-free* (qf), and the *primitive positive* (pp) *fragment*. The existential positive fragment consists of formulas built using conjunction, disjunction, and existential quantification only. The quantifier-free fragment consists of Boolean combinations of atomic formulas, and the primitive positive fragment of existentially quantified conjunctions of atomic formulas. A formula is called *equality-free* if it does not contain any occurrence of the default equality predicate $=$.³ Let Σ be a set of first-order formulas and \mathfrak{D} a structure. We say that a relation over D has a Σ -*definition* in \mathfrak{D} if it is of the form $\{\bar{t} \in D^k \mid \mathfrak{D} \models \phi(\bar{t})\}$ for some $\phi \in \Sigma$. We refer to this relation by $\phi^{\mathfrak{D}}$. For example, the formula $y < x \vee y = x$ is existential positive and quantifier-free. Interpreted in the structure \mathbb{Q} , it clearly defines the binary relation \geq on \mathbb{Q} . This shows that \geq is \exists^+ and qf definable in \mathbb{Q} . An example of a pp formula is the formula $x = x$, which defines the unary relation interpreted as the whole domain \mathbb{Q} .

The definition of admissibility of a concrete domain in [2] requires that the constraint satisfaction problem for this structure is decidable. For a fixed τ -structure \mathfrak{B} , the *constraint satisfaction problem* (CSP) for \mathfrak{B} [16] asks whether a given pp τ -sentence is satisfiable in \mathfrak{B} . An *implication* is of the form $\forall \bar{x}. (\phi \Rightarrow \psi)$ where ϕ is a conjunction of atomic τ -formulas, ψ is a disjunction of atomic τ -formulas, and the tuple \bar{x} consists of all the variables occurring in ϕ or ψ . Such an implication is a *Horn implication* if ψ is a single atomic τ -formula. A universal sentence is called *Horn* if it is a conjunction of Horn implications. The CSP for \mathfrak{A} can be reduced in polynomial time to the validity problem for Horn implications since ϕ is satisfiable in \mathfrak{A} iff $\forall \bar{x}. (\phi \Rightarrow \text{ff})$ is not valid in \mathfrak{A} . Conversely, validity of Horn implications in a structure \mathfrak{A} can be reduced in polynomial time to $\text{CSP}(\mathfrak{A}^\neg, \neq)$ where \mathfrak{A}^\neg is the expansion of \mathfrak{A} by the complements of all relations. In fact, the Horn implication $\forall \bar{x}. (\phi \Rightarrow \psi)$ is valid in \mathfrak{A} iff $\phi \wedge \neg\psi$ is not satisfiable in $(\mathfrak{A}^\neg, \neq)$. Note that, in the signature of $(\mathfrak{A}^\neg, \neq)$, $\neg\psi$ can be expressed by an atomic formula.

A *homomorphism* $h: \mathfrak{A} \rightarrow \mathfrak{B}$ for τ -structures \mathfrak{A} and \mathfrak{B} is a mapping $h: A \rightarrow B$ that *preserves* each relation of \mathfrak{A} , i.e., if $\bar{t} \in R^{\mathfrak{A}}$ for some k -ary relation symbol $R \in \tau$, then $h(\bar{t}) \in R^{\mathfrak{B}}$. The homomorphism $h: \mathfrak{A} \rightarrow \mathfrak{B}$ is *strong* if it additionally satisfies the inverse condition: for every k -ary relation symbol $R \in \tau$ and $\bar{t} \in A^k$ we have $h(\bar{t}) \in R^{\mathfrak{B}}$ only if $\bar{t} \in R^{\mathfrak{A}}$. An *embedding* is an injective strong homomorphism. We write $\mathfrak{A} \rightarrow \mathfrak{B}$ ($\mathfrak{A} \hookrightarrow \mathfrak{B}$) if there is a homomorphism (embedding) from \mathfrak{A} to \mathfrak{B} . A *substructure* of \mathfrak{B} is a structure \mathfrak{A} over the domain $A \subseteq B$ such that the inclusion map $i: A \rightarrow B$ is an embedding. Conversely, we call \mathfrak{B} an *extension* of \mathfrak{A} . We denote by $\text{Age}(\mathfrak{B})$ the class of all *finite* structures \mathfrak{A} with $\mathfrak{A} \hookrightarrow \mathfrak{B}$. An *isomorphism* is a surjective embedding. Two structures \mathfrak{A} and \mathfrak{B} are *isomorphic* (written $\mathfrak{A} \cong \mathfrak{B}$) if there exists an isomorphism from \mathfrak{A} to \mathfrak{B} . An *automorphism* is an isomorphism from \mathfrak{A} to \mathfrak{A} . Two structures \mathfrak{A} and \mathfrak{B} are *homomorphically equivalent* if $\mathfrak{A} \rightarrow \mathfrak{B}$ and $\mathfrak{B} \rightarrow \mathfrak{A}$.

If the signature τ of \mathfrak{B} is finite, the constraint satisfaction problem for \mathfrak{B} can also be conveniently formulated using homomorphisms: given a finite τ -structure \mathfrak{A} , decide whether $\mathfrak{A} \rightarrow \mathfrak{B}$. A *solution* for such an instance \mathfrak{A} of the CSP is then simply a homomorphism $h: \mathfrak{A} \rightarrow \mathfrak{B}$ and $\text{CSP}(\mathfrak{B})$ is the class of all finite τ -structures that homomorphically map to \mathfrak{B} . It is easy to see that this definition of the CSP coincides with the one given above. Indeed, deciding whether a CSP instance \mathfrak{A} admits a solution amounts to evaluating a pp sentences in

³ In case the signature τ of a structure contains a symbol that is interpreted as equality in that structure, an equality-free formula can, of course, still use that symbol.

\mathfrak{B} and vice versa [16]. For example, verifying whether the structure $\mathfrak{A} = (\{a_1, a_2, a_3\}; <^{\mathfrak{A}})$ with $<^{\mathfrak{A}} := \{(a_1, a_2), (a_2, a_3), (a_3, a_1)\}$ homomorphically maps into \mathfrak{Q} is the same as checking whether the pp formula $x_1 < x_2 \wedge x_2 < x_3 \wedge x_3 < x_1$ is satisfiable in \mathfrak{Q} .

The CSP for \mathfrak{Q} is tractable since a structure $\mathfrak{A} = (A; <^{\mathfrak{A}})$ can homomorphically be mapped into \mathfrak{Q} iff it does not contain a $<$ -cycle, i.e., there are no $n \geq 1$ and elements $a_1, \dots, a_n \in A$ such that $a_1 <^{\mathfrak{A}} \dots <^{\mathfrak{A}} a_n <^{\mathfrak{A}} a_1$. Testing whether such a cycle exists can be done in non-deterministic logarithmic space since it requires solving the reachability problem in a directed graph (digraph). In the example above, we obviously have a cycle, and thus this instance of $\text{CSP}(\mathfrak{Q})$ has no solution.

The definition of admissibility in [2] actually also requires that the predicates are closed under negation and that there is a predicate for the whole domain. We have already seen that the negation \geq of $<$ is \exists^+ definable in \mathfrak{Q} and that the predicate for the whole domain is pp definable. The negation of this predicate has the pp definition $x < x$. The following lemma implies that the expansion of \mathfrak{Q} by these predicates still has a decidable CSP.⁴

Lemma 1 ([16]) *Let $\mathfrak{C}, \mathfrak{D}$ be structures over the same domain with finite signatures.*

1. *If the relations of \mathfrak{C} have a pp definition in \mathfrak{D} , then $\text{CSP}(\mathfrak{C}) \leq_{\text{PTIME}} \text{CSP}(\mathfrak{D})$.*
2. *If the relations of \mathfrak{C} have an \exists^+ definition in \mathfrak{D} , then $\text{CSP}(\mathfrak{C}) \leq_{\text{NPTIME}} \text{CSP}(\mathfrak{D})$.*

3 Description Logics with Concrete Domains

We assume that the reader is familiar with the basic definitions and results in DL [10, 12], but nevertheless briefly recall the definitions of the two DLs \mathcal{ALC} and \mathcal{EL} relevant for this paper. Then we describe how these DLs can be extended with concrete domains. The integrations of concrete domains into DLs described in the literature [2, 11, 35, 60–62, 65] differ in some details. The approaches described below for \mathcal{ALC} and \mathcal{EL} are close to the ones in [65] and [11], respectively, but not identical, mainly as a matter of convenience of presentation. Reasoning in DLs obtained this way may easily become undecidable, and thus one needs to find conditions that guarantee decidability, and even tractability for the case of \mathcal{EL} .

3.1 Basic Definitions and Undecidability Results

Given countably infinite sets $\mathbb{N}_{\mathcal{C}}$ and $\mathbb{N}_{\mathcal{R}}$ of concept and role names, \mathcal{ALC} concepts are built using the concept constructors top concept (\top), bottom concept (\perp), negation (\neg), conjunction ($C \sqcap D$), disjunction ($C \sqcup D$), existential restriction ($\exists r.C$), and universal restriction ($\forall r.C$). The semantics of the constructors is defined in the usual way (see, e.g., [10, 12]). It assigns to every \mathcal{ALC} concept C a set $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, where $\Delta^{\mathcal{I}}$ is the interpretation domain of the given interpretation \mathcal{I} . The set of \mathcal{EL} concepts is obtained by restricting the available constructors to $\top, C \sqcap D$, and $\exists r.C$. As usual, a TBox is defined to be a finite set of *general concept inclusions* (GCIs) $C \sqsubseteq D$, where C, D are concepts. The interpretation \mathcal{I} is a *model* of such a TBox if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ holds for all GCIs $C \sqsubseteq D$ occurring in it. Given a concept description C and a TBox \mathcal{T} , we say that C is *satisfiable* w.r.t. \mathcal{T} if $C^{\mathcal{I}}$ is non-empty for some model \mathcal{I} of \mathcal{T} . Concept satisfiability w.r.t. GCIs is ExpTime-complete in \mathcal{ALC} [71], but trivial in \mathcal{EL} since \mathcal{EL} concepts are always satisfiable. Given concept descriptions C, D and a TBox \mathcal{T} , we say that C is *subsumed by* D w.r.t. \mathcal{T} (written $C \sqsubseteq_{\mathcal{T}} D$) if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ holds

⁴ The lemma actually only yields an NP decision procedure for this CSP, but it is easy to see that the above polynomial-time cycle-checking algorithm can be adapted such that it also works for the expanded structure.

for all models of \mathcal{T} . Subsumption w.r.t. TBoxes in \mathcal{ALC} is also ExpTime-complete since it is irreducible with concept satisfiability, but tractable (i.e., decidable in polynomial time) in \mathcal{EL} [11, 32].

From an algebraic point of view, a concrete domain is a τ -structure \mathfrak{D} with a relational signature τ without constant symbols. To integrate such a structure into \mathcal{ALC} and \mathcal{EL} , we complement concept and role names with a set of *feature names* N_F , which provide the connection between the abstract domain $\Delta^{\mathcal{T}}$ and the concrete domain D . A *path* is of the form $r f$ or f where $r \in N_R$ and $f \in N_F$. In our example in the introduction, *age* is both a feature name and a path of length 1, and *child age* is a path of length 2.

Definition 1 *Concrete domain restrictions* for a relational τ -structure \mathfrak{D} are concept constructors of the form $\exists p_1, \dots, p_k. \phi(x_1, \dots, x_k)$ or $\forall p_1, \dots, p_k. \phi(x_1, \dots, x_k)$, where p_1, \dots, p_k are paths and ϕ is a first-order τ -formula with free variables x_1, \dots, x_k . The DL $\mathcal{ALC}(\mathfrak{D})$ extends \mathcal{ALC} with concrete domain restrictions where ϕ is allowed to be an arbitrary atomic τ -formula. The DL $\mathcal{EL}(\mathfrak{D})$ is the sublanguage of $\mathcal{ALC}(\mathfrak{D})$ where only the concept constructors of \mathcal{EL} together with existential concrete domain restrictions can be used. Let Σ be a set of first-order τ -formulas and n a natural number. The DL $\mathcal{ALC}_{\Sigma}^n(\mathfrak{D})$ extends \mathcal{ALC} with concrete domain restrictions where ϕ is allowed to be an at most n -ary formula from Σ .

In contrast to previous works on concrete domains [2, 65], we generally allow the use of the equality predicate in concrete domain restrictions, even if it is not explicitly contained in the signature of the concrete domain. This assumption will turn out to be useful later on, and it is basically without loss of generality since virtually all concrete domains considered in the literature can express equality in a way that does not impact on the complexity of reasoning. Our assumption that ff is an atom implies that $\mathcal{EL}(\mathfrak{D})$ can express the bottom concept \perp by the concrete domain restriction $\exists f.f f$. A third difference is that, while features pointing into the concrete domain are functional, we do not allow the use of functional roles in paths. In [2], only functional roles are allowed to occur in paths whereas in [65] both functional and other roles can occur there. For \mathcal{ALC} , this does not really make a difference due to the availability of universal concrete domain restrictions. For \mathcal{EL} , the presence of functional roles would destroy tractability even without concrete domains [11], and thus needs to be avoided anyway.

To define the semantics of concrete domain restrictions, we assume that an interpretation \mathcal{I} assigns functional binary relations $f^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times D$ to feature names $f \in N_F$, where *functional* means that $(a, d) \in f^{\mathcal{I}}$ and $(a, d') \in f^{\mathcal{I}}$ imply $d = d'$. We extend the interpretation function to paths of the form $p = r f$ by setting

$$(r f)^{\mathcal{I}} = \{(a, d) \in \Delta^{\mathcal{I}} \times D \mid \text{there is } b \in \Delta^{\mathcal{I}} \text{ such that } (a, b) \in r^{\mathcal{I}} \text{ and } (b, d) \in f^{\mathcal{I}}\}.$$

The semantics of concrete domain restrictions is now defined as follows:

$$\begin{aligned} (\exists p_1, \dots, p_k. \phi(x_1, \dots, x_k))^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \text{there are } d_1, \dots, d_k \in D \text{ such that} \\ &\quad (a, d_i) \in p_i^{\mathcal{I}} \text{ for all } i \in [k] \text{ and } \mathfrak{D} \models \phi(d_1, \dots, d_k)\}, \\ (\forall p_1, \dots, p_k. \phi(x_1, \dots, x_k))^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \text{for all } d_1, \dots, d_k \in D \text{ such that} \\ &\quad (a, d_i) \in p_i^{\mathcal{I}} \text{ for all } i \in [k] \text{ we have } \mathfrak{D} \models \phi(d_1, \dots, d_k)\}. \end{aligned}$$

As already mentioned above, the concrete domain restriction $\exists f.f f$ is unsatisfiable, and thus equivalent to \perp . The restriction $\exists f.f.(x_1 = x_2)$ expresses that the value of the feature f must be defined, without putting any constraint on this value.

Adding a concrete domain to a DL can easily lead to undecidability. Clearly, if the CSP of the concrete domain is undecidable, then this transfers to the DL it is integrated in. If the

concrete domain is admissible, i.e., its CSP is decidable and its relations are closed under complements, then concept satisfiability without GCIs is decidable in a variant of \mathcal{ALC} with concrete domains that uses functional roles in paths [2]. But even for very simple concrete domains with decidable CSPs, the presence of GCIs may cause undecidability. For instance, $\mathcal{ALC}(\mathcal{D})$ is undecidable already when \mathcal{D} is a structure over \mathbb{N} that has access to the unary predicate $=_0$, which is interpreted as the singleton set $\{0\}$, and the binary predicate $+_1$, which is interpreted as incrementation (i.e., it consists of the tuples $(m, m + 1)$ for $m \in \mathbb{N}$) [12]. We can improve on this result by showing undecidability for even less expressive concrete domains without the predicate $=_0$.

Proposition 1 ([6]) *If \mathcal{D} is of the form $(D; +_1)$ or $(D; +)$ for $D \in \{\mathbb{Q}, \mathbb{Z}, \mathbb{N}\}$, then concept satisfiability in $\mathcal{ALC}(\mathcal{D})$ w.r.t. TBoxes is undecidable.*

This undecidability results also holds without our assumption that equality is always available, but the proof given in [6] uses functional roles in paths. This proof can, however, easily be adapted to work also without functional roles. One simply must use additional universal quantification (i.e., value restrictions and universal concrete domain restrictions) to ensure that all the successors of an individual w.r.t. a role that was assumed to be functional in the original proof behave the same. More specifically, one must replace each occurrence of a concrete domain restriction of the form $\exists p_1, \dots, p_k.\phi$ for functional paths p_i with the concept description $\exists p_1, \dots, p_k.\phi \sqcap \forall p_1, \dots, p_k.\phi$. Then the functionality restriction on these paths can be removed.

Even for \mathcal{EL} , integrating a decidable concrete domain may cause undecidability if we allow for role paths of length 2. Proving this is, however, more challenging, not only due to the fact that not all Boolean operations are available, but also since the absence of functional roles cannot be compensated by the use of universal quantification. To illustrate the latter point, assume we have a concrete domain with binary predicates P and P' that are disjoint. If g is assumed to be a functional role, then the concept $\exists f, g.f.P(x_1, x_2) \sqcap \exists f, g.f.P'(x_1, x_2)$ is unsatisfiable, but if g is just an arbitrary role, then it is satisfiable since a given individual belonging to the concept may have two different g -successors, one satisfying the P -constraint and the other satisfying the P' -constraint. However, conjoining this concept with the corresponding universal concrete domain restrictions $\forall f, g.f.P(x_1, x_2) \sqcap \forall f, g.f.P'(x_1, x_2)$ yields an unsatisfiable concept again.

To show undecidability for a concrete domain extension of \mathcal{EL} without functional roles, we consider the relational structure $\mathcal{D}_{\text{aff}, \mathbb{Q}^2}$ over \mathbb{Q}^2 , which has, for every affine transformation $\mathbb{Q}^2 \rightarrow \mathbb{Q}^2 : \bar{x} \mapsto A\bar{x} + \bar{b}$, the binary relation $R_{A, \bar{b}} := \{(\bar{x}, \bar{y}) \in (\mathbb{Q}^2)^2 \mid \bar{y} = A\bar{x} + \bar{b}\}$ as a predicate. We will show in Corollary 8 that the CSP for this structure is decidable in polynomial time. Undecidability of subsumption w.r.t. TBoxes in $\mathcal{EL}(\mathcal{D}_{\text{aff}, \mathbb{Q}^2})$ can be shown by a reduction from *2-Dimensional Affine Reachability*, which is undecidable by Corollary 4 in [14]. For this problem, one is given vectors $\bar{v}, \bar{w} \in \mathbb{Q}^2$ and a finite set S of affine transformations from \mathbb{Q}^2 to \mathbb{Q}^2 . The question is then whether \bar{w} can be obtained from \bar{v} by repeated application of transformations from S .

Proposition 2 *Subsumption w.r.t. TBoxes is undecidable in $\mathcal{EL}(\mathcal{D}_{\mathbb{Q}^2, \text{aff}})$.*

Proof We define the reduction of 2-dimensional Affine Reachability to subsumption w.r.t. general TBoxes in $\mathcal{EL}(\mathcal{D}_{\text{aff}, \mathbb{Q}^2})$ as follows. For given vectors $\bar{v}, \bar{w} \in \mathbb{Q}^2$ and affine transformations $S = \{\bar{x} \mapsto M_1\bar{x} + \bar{v}_1, \dots, \bar{x} \mapsto M_k\bar{x} + \bar{v}_k\}$, the TBox \mathcal{T} contains, for every $i \in [k]$, the GCI $\top \sqsubseteq \exists f, g.f.R_{M_i, \bar{v}_i}(x_1, x_2)$. Additionally, \mathcal{T} contains the GCIs $\exists g.L \sqsubseteq L$ and $\exists f, f.R_{Z, \bar{w}}(x_1, x_2) \sqsubseteq L$, where L is a fresh concept name and Z is the 2×2 zero matrix.

Note that $(\bar{x}, \bar{x}) \in R_{Z, \bar{w}}$ iff $\bar{x} = \bar{w}$. Each involved concept is either \top , a concept name, or an existential (concrete domain) restriction, and thus definable in $\mathcal{EL}(\mathcal{D}_{\text{aff}, \mathbb{Q}^2})$. We claim that $\exists f, f.R_{Z, \bar{v}}(x_1, x_2)$ is subsumed by L w.r.t. \mathcal{T} iff \bar{w} can be obtained from \bar{v} through an application of a composition of affine transformations from S .

“ \Leftarrow ”: Suppose that there exists such a composition and let \mathcal{I} be a model of \mathcal{T} . Let a be an arbitrary element of $(\exists f, f.R_{Z, \bar{v}}(x_1, x_2))^{\mathcal{I}}$, i.e., satisfying $f^{\mathcal{I}}(a) = \bar{v}$. Since \mathcal{T} contains $\top \sqsubseteq \exists f, g f.R_{M_i, \bar{v}_i}(x_1, x_2)$ for every $i \in [k]$ and \bar{w} is reachable from \bar{v} through an application of a composition of affine transformations from S , there exists a role path $a \rightarrow_{g^{\mathcal{I}}} \dots \rightarrow_{g^{\mathcal{I}}} b$ to some element b with $f^{\mathcal{I}}(b) = \bar{w}$. Since \mathcal{T} contains the GCI $\exists f, f.R_{Z, \bar{w}}(x_1, x_2) \sqsubseteq L$, we have $b \in L^{\mathcal{I}}$. The GCI $\exists g.L \sqsubseteq L$ then yields $a \in L^{\mathcal{I}}$.

“ \Rightarrow ”: Suppose that $\exists f, f.R_{Z, \bar{v}}(x_1, x_2)$ is subsumed by L w.r.t. \mathcal{T} . Consider the following interpretation \mathcal{I} . The domain of \mathcal{I} is \mathbb{Q}^2 . We define $f^{\mathcal{I}}$ as the identity map on \mathbb{Q}^2 and set $g^{\mathcal{I}} := \{(\bar{x}, \bar{y}) \in (\mathbb{Q}^2)^2 \mid \exists i \in [k] \text{ such that } \bar{y} = M_i \bar{x} + \bar{v}_i\}$. Finally, we set $L^{\mathcal{I}} := \{\bar{w}\} \cup \{\bar{x} \in \mathbb{Q}^2 \mid \text{there exists a role path } \bar{x} \rightarrow_{g^{\mathcal{I}}} \dots \rightarrow_{g^{\mathcal{I}}} \bar{w}\}$. It is easy to check that \mathcal{I} is a model of \mathcal{T} . Since $\bar{v} \in (\exists f, f.R_{Z, \bar{v}}(x_1, x_2))^{\mathcal{I}}$ and $\exists f, f.R_{Z, \bar{v}}(x_1, x_2)$ is subsumed by L w.r.t. \mathcal{T} , we have $\bar{v} \in L^{\mathcal{I}}$. The definition of $L^{\mathcal{I}}$ thus implies that \bar{w} is reachable from \bar{v} through an application of a composition of affine transformations from S . \square

Note that the signature of $\mathcal{D}_{\text{aff}, \mathbb{Q}^2}$ is infinite since there are infinitely many affine transformations on \mathbb{Q}^2 . One might think that this is important for our undecidability proof.

We can show, however, that this is not the case: a fixed finite set of affine transformations is sufficient (see the appendix for a proof).

Corollary 1 *There exists a finite signature reduct \mathcal{D} of $\mathcal{D}_{\text{aff}, \mathbb{Q}^2}$ such that subsumption w.r.t. $T\Box$ es is undecidable in $\mathcal{EL}(\mathcal{D})$.*

3.2 Decidable and Tractable DLs with Concrete Domains

There are two strategies for regaining decidability of DLs with concrete domains in the presence of GCIs: syntactically restricting the interaction of the DL with the concrete domain or limiting the expressiveness of the concrete domain itself. Typically, the former is realized by restricting the length of paths in concrete domain restrictions to 1. We indicate this restriction by writing square brackets around the concrete domain instead of round brackets.

Definition 2 The restriction of $\mathcal{EL}(\mathcal{D})$ and $\mathcal{ALC}(\mathcal{D})$ to paths of length 1 in concrete domain restrictions is respectively denoted by $\mathcal{EL}[\mathcal{D}]$ and $\mathcal{ALC}[\mathcal{D}]$.

For \mathcal{ALC} , this restriction results in decidability [45, 69] for concrete domains that are *admissible* in the sense introduced in [2], i.e., whose predicates are closed under negation and whose CSP is decidable. In the case of \mathcal{EL} , the expectations are a bit higher: the aim there is to regain tractability. To obtain tractability of $\mathcal{EL}[\mathcal{D}]$, the notion of *p-admissible* concrete domains was introduced in [11], and it was shown that subsumption in $\mathcal{EL}[\mathcal{D}]$ is decidable in polynomial time iff \mathcal{D} is *p-admissible*. Before defining this condition below, we introduce a condition, called ω -admissibility, which ensures decidability of $\mathcal{ALC}(\mathcal{D})$ in the presence of GCIs and paths of length > 1 .

3.2.1 ω -Admissible Concrete Domains

The notion of ω -admissibility was introduced in [65] to regain decidability of \mathcal{ALC} with concrete domains in the presence of GCIs. Motivated by binary constraint calculi like Allen’s

interval calculus [1] and the region connection calculus [70], only concrete domains where all predicates are binary were considered in [65]. In [6], the notion and the corresponding decidability result were generalized to concrete domains with predicates of arbitrary arity.

We say that the structure \mathcal{D} has *homomorphism ω -compactness* if the following holds for every countable structure \mathcal{B} : $\mathcal{B} \rightarrow \mathcal{D}$ iff $\mathcal{A} \rightarrow \mathcal{D}$ for every $\mathcal{A} \in \text{Age}(\mathcal{B})$. In [65], the inputs to this condition were not formally restricted to countable structures. However, it is clear that this is what the authors meant because (i) the structures produced by the original tableau algorithm that need to be tested for a homomorphism to the concrete domain are always countable, and (ii) the examples of ω -admissible concrete domains presented in [65] are not homomorphic compact for arbitrarily large cardinalities. A relational τ -structure \mathcal{D} is

- JE if, for every $k \geq 1$, either \mathcal{D} has no k -ary relations or $\bigcup \{R^{\mathcal{D}} \mid R \in \tau, R^{\mathcal{D}} \subseteq D^k\} = D^k$;
- PD if $R^{\mathcal{D}} \cap \tilde{R}^{\mathcal{D}} = \emptyset$ for all pairwise distinct $R, \tilde{R} \in \tau$;
- JD if equality $\text{Eq}^{\mathcal{D}}$ has a (quantifier and equality)-free definition in \mathcal{D} .

Here JE stands for “jointly exhaustive,” PD for “pairwise disjoint,” and JD for “jointly diagonal.” In [6], JD was defined in a more restricted way as $\bigcup \{R^{\mathcal{D}} \mid R \in \tau, R^{\mathcal{D}} \subseteq \text{Eq}^{\mathcal{D}}\} = \text{Eq}^{\mathcal{D}}$, which explains the name. The condition JD was not considered in [65]. We include it here mainly because it makes the comparison with known notions from model theory easier. In the setting considered in the present paper, where concrete domain restrictions always have access to equality, JD is actually needed to ensure decidability. If the equality predicate is dropped from concrete domain restrictions, then the decidability results in [7, 65] do not depend on JD. However, all examples of ω -admissible concrete domains presented in [65] satisfy JD since equality is contained in the signature. In [39], k -ary structures, i.e., structures \mathcal{D} that have only k -ary predicates, are considered that have the k -ary equality relation $\text{Eq}_k^{\mathcal{D}} = \{(d, \dots, d) \in D^k \mid d \in D\}$. For $k \geq 2$, such a structure satisfies JD in the sense introduced above, since binary equality $x = y$ can be defined as $\text{Eq}_k^{\mathcal{D}}(x, y, \dots, y)$.

A relational τ -structure \mathcal{D} is a *patchwork* if it is JDJEPD, and for all finite JEPD τ -structures $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2$ with $e_1: \mathcal{A} \hookrightarrow \mathcal{B}_1, e_2: \mathcal{A} \hookrightarrow \mathcal{B}_2, \mathcal{B}_1 \rightarrow \mathcal{D}$, and $\mathcal{B}_2 \rightarrow \mathcal{D}$, there exist $f_1: \mathcal{B}_1 \rightarrow \mathcal{D}$ and $f_2: \mathcal{B}_2 \rightarrow \mathcal{D}$ with $f_1 \circ e_1 = f_2 \circ e_2$.

Definition 3 The relational structure \mathcal{D} is *ω -admissible* if it has a finite signature, $\text{CSP}(\mathcal{D})$ is decidable, \mathcal{D} has homomorphism ω -compactness, and \mathcal{D} is a patchwork.

The idea is now that one can use disjunctions of atomic formulas of the same arity within concrete domain restrictions. By \vee^+ we denote the set of all τ -formulas of the form $\phi_1(x_1, \dots, x_k) \vee \dots \vee \phi_m(x_1, \dots, x_k)$ where each ϕ_i is a k -ary atomic τ -formula.

The following theorem is shown in [6, 7] by extending the tableau-based decision procedure given in [65] to our more general definition of ω -admissibility.

Theorem 1 Let \mathcal{D} be an ω -admissible τ -structure with at most d -ary relations for some $d \geq 2$. Then concept satisfiability in $\text{ALC}_{\vee^+}^d(\mathcal{D})$ w.r.t. $T\text{Boxes}$ is decidable.

The main motivation for the definition of ω -admissible concrete domains in [65] was that they can capture qualitative calculi of time and space. In particular, it was shown in [65] that Allen’s interval logic [1] as well as the region connection calculus RCC8 [70] can be represented as ω -admissible concrete domains. To the best of our knowledge, no other ω -admissible concrete domains have been exhibited in the literature before our investigations in [6], which we will describe in detail in the next section. Among other things, we prove that

the structure $(\mathbb{Q}; <, =, >)$ is ω -admissible. The “discrete” version $(\mathbb{Z}; <, =, >)$, on the other hand, is not ω -admissible because it lacks homomorphism ω -compactness (see Example 1 below). By the results in [60, 61], $(\mathbb{Z}; <, =, >)$ nevertheless yields a decidable concrete domain extension of \mathcal{ALC} , but proving this requires a more specialized approach than the tableau algorithm provided by the original paper of Lutz and Milićić [65]. This shows that ω -admissibility is not necessary for decidable reasoning in \mathcal{ALC} with concrete domains in the presence of GCIs.

Example 1 The concept $A \in \mathbf{N}_C$ is satisfiable w.r.t. the TBox

$$\mathcal{T} := \{A \sqsubseteq (\exists f, g. <(x_1, x_2)) \sqcap (\exists r. A) \sqcap (\forall f, r. f. <(x_1, x_2)) \sqcap (\forall r. g, g. <(x_1, x_2))\}$$

in $\mathcal{ALC}(\mathbb{Q}; <, =, >)$, and this can be tested using the tableau algorithm from [65] because $(\mathbb{Q}; <, =, >)$ is ω -admissible by Theorem 6. However, A is not satisfiable w.r.t. \mathcal{T} in $\mathcal{ALC}(\mathbb{Z}; <, =, >)$ because its satisfiability would imply the existence of a homomorphism to $(\mathbb{Z}; <, =, >)$ from a structure \mathfrak{B} with domain $B = \{f_n, g_n \mid n \in \mathbb{N}\}$ and relations given by $f_n <^{\mathfrak{B}} f_{n+1} <^{\mathfrak{B}} g_{n+1} <^{\mathfrak{B}} g_n$ for every $n \in \mathbb{N}$. Such a homomorphism cannot exist because the ordering of the integers is not dense. Note that $\mathfrak{A} \rightarrow (\mathbb{Z}; <, =, >)$ for every $\mathfrak{A} \in \text{Age}(\mathfrak{B})$, which shows that $(\mathbb{Z}; <, =, >)$ is not homomorphism compact.

3.2.2 p-Admissible Concrete Domains

The notion of p-admissibility was introduced in [11] to capture precisely those structures \mathfrak{D} for which subsumption in $\mathcal{EL}[\mathfrak{D}]$ is tractable. Clearly, this requires the CSP of \mathfrak{D} to be decidable in polynomial time. However, this is not sufficient since even for a concrete domain \mathfrak{D} with tractable CSP disjunction may be expressible in $\mathcal{EL}[\mathfrak{D}]$, which then leads to intractability [11]. To avoid this source of intractability, the concrete domain must be convex. Unfortunately, the definition of convexity given in [11] was ambiguous, and what is really needed in the setting considered in [11] is what we call guarded convexity in [8]. However, in the setting considered in the present paper, where equality is always available in concrete domain restrictions, we will see that convexity rather than guarded convexity is the adequate notion.

We say that a τ -structure \mathfrak{D} is *convex* if the following holds: whenever a conjunction of atomic τ -formulas implies a disjunction of atomic τ -formulas in \mathfrak{D} , then it already implies one of the disjuncts. Note that this definition does not say anything about which variables may occur in the left- and right-hand sides of such implications. *Guarded convexity* requires this condition to hold only for guarded implications, where all variables occurring on the right-hand side must also occur on the left-hand side.

To illustrate the difference between convexity and guarded convexity, let us consider the structure $\mathfrak{N} = (\mathbb{N}; E, O)$ in which the unary predicates E and O are respectively interpreted as the even and odd natural numbers. This structure is not convex since $\forall x, y. (E(x) \Rightarrow E(y) \vee O(y))$ holds in \mathfrak{N} , but neither $\forall x, y. (E(x) \Rightarrow E(y))$ nor $\forall x, y. (E(x) \Rightarrow O(y))$ does. However, the first implication is not guarded, and it is easy to see that \mathfrak{N} is in fact guarded convex. Note that, whereas $\forall x, y. (E(x) \Rightarrow E(y) \vee O(y))$ holds in \mathfrak{N} , the subsumption $\exists f. E(x_1) \sqsubseteq_{\emptyset} \exists g. E(x_1) \sqcup \exists g. O(x_1)$ does not hold in the extension of $\mathcal{EL}[\mathfrak{N}]$ with disjunction since the feature g need not have a value. However, as we have pointed out above, $\exists g, g. (x_1 = x_2)$ expresses that the value of g must be defined. Thus, $\exists g, g. (x_1 = x_2) \sqsubseteq_{\emptyset} \exists g. E(x_1) \sqcup \exists g. O(x_1)$ is a valid subsumption in $\mathcal{EL}[\mathfrak{N}]$. This can be used to show that this DL is not tractable [11], but only under the assumption that equality can be used in concrete domain

restrictions. Consequently, in the setting of the present paper, convexity should be used in the definition of p-admissibility.

Definition 4 A relational structure \mathfrak{D} is *p-admissible* if it is convex and validity of Horn implications in \mathfrak{D} is decidable in polynomial time.

The main result of [11] concerning concrete domains can now be stated as follows.

Theorem 2 (Baader, Brandt, and Lutz [11]) *Let \mathfrak{D} be a relational structure. Then subsumption in $\mathcal{EL}[\mathfrak{D}]$ is decidable in polynomial time iff \mathfrak{D} is p-admissible.*

Note that the theorem above does not hold for the more expressive logic $\mathcal{EL}(\mathfrak{D})$ where paths of length 2 are allowed in concrete domain constructors. This is because we can show that the concrete domain $\mathfrak{D}_{\text{aff}, \mathbb{Q}^2}$ from Proposition 2 is p-admissible (see Corollary 8). In Sect. 5, we provide an algebraic characterization of convexity. Regarding the tractability condition in the definition of p-admissibility, we have seen in Sect. 2 that it is closely related to the constraint satisfaction problem for \mathfrak{D} and $(\mathfrak{D}^\neg, \neq)$. Characterizing tractability of the CSP in a given structure is a very hard problem. Whereas the Feder-Vardi conjecture [43] has recently been confirmed after 25 years of intensive research in the field by giving an algebraic criterion that can distinguish between *finite* structures with tractable and with NP-complete CSP [34, 72], finding comprehensive criteria that ensure tractability for the case of infinite structures is a wide open problem, though first results for special cases have been found (see, e.g., [18, 25–28, 58, 67]).

3.2.3 ω -Admissibility Versus p-Admissibility

From an application point of view it would be desirable to have concrete domains \mathfrak{D} that preserve tractability if used in $\mathcal{EL}[\mathfrak{D}]$ and decidability if used in $\mathcal{ALC}(\mathfrak{D})$. This would be the case for concrete domains that are both ω -admissible and p-admissible. Unfortunately, for structures over a finite signature, JEPD (required for ω -admissibility) and convexity (required for p-admissibility) do not go well together.

Proposition 3 *Let τ be a finite signature and \mathfrak{D} a relational τ -structure that is both JEPD and convex. Then $R^\mathfrak{D} \in \{\emptyset, D^k\}$ for all k -ary relation symbols $R \in \tau$.*

Proof Assume that $R \in \tau$ is a relation symbol of arity k such that $R^\mathfrak{D} \neq \emptyset$. Then JE yields $D^k = \bigcup_{i=1}^m R_i^\mathfrak{D}$, where R_1, \dots, R_m are all the relation symbols of arity k in τ . Consequently, the implication $\forall x_1, \dots, x_k. (\bigwedge_{i=1}^k x_i = x_i) \Rightarrow (\bigvee_{i=1}^m R_i(x_1, \dots, x_k))$ holds in \mathfrak{D} , and thus convexity implies that there is an i , $1 \leq i \leq m$, such that $\forall x_1, \dots, x_k. (\bigwedge_{i=1}^k x_i = x_i) \Rightarrow R_i(x_1, \dots, x_k)$ holds in \mathfrak{D} . This means that $R_i^\mathfrak{D} = D^k$. Since we have assumed that $R^\mathfrak{D} \neq \emptyset$ and R is of arity k , PD yields that $R = R_i$, and thus we are done. \square

If τ contains a symbol R that is interpreted as equality Eq^D on D , then this proposition implies that $\text{Eq}^D = R^\mathfrak{D} = D \times D$, which can only be the case if $|D| \leq 1$. The proof of Proposition 3 makes use of our assumption that equality is always available when building formulas. But even without that assumption, concrete domains \mathfrak{D} that are both p- and ω -admissible would have a rather restricted form. In that case, there always exists a finite partition $\{V_1, \dots, V_m\}$ of D such that the only non-empty k -ary relations of \mathfrak{D} are of the form $V_{i_1} \times \dots \times V_{i_k}$ for $i_1, \dots, i_k \in [m]$ [5].

Finally, let us point out another notable difference between the two conditions, namely that p-admissibility permits infinite signatures whereas ω -admissibility does not. It turns out that

finiteness of the signature is a necessary part of ω -admissibility to achieve decidability. If we allowed the signature of \mathfrak{D} to be infinite, then we would have the following counterexample. Let \mathfrak{D} be the structure over \mathbb{Z} with the relations $+_k^{\mathfrak{D}} = \{(x, y) \in \mathbb{Z}^2 \mid y = x + k\}$ for every $k \in \mathbb{Z}$. It is easy to see that $\text{CSP}(\mathfrak{D})$ can be solved in polynomial time and that \mathfrak{D} has homomorphism ω -compactness. Moreover, one can show, using the results in Sect. 4 (Proposition 4 and Theorem 5), that \mathfrak{D} is a patchwork. However, we have seen in Proposition 1 that concept satisfiability w.r.t. GCIs is undecidable already in $\mathcal{ALC}(\mathbb{Z}; +_1)$.

4 A Model-Theoretic Analysis of ω -Admissibility

We introduce several model-theoretic properties of relational structures and show their connection with ω -admissibility. This allows us to formulate sufficient conditions for ω -admissibility using well-known notions from model theory, and thus to use existing model-theoretic results to find new ω -admissible concrete domains. We start with the notion of ω -categoricity in a countable signature, which is sufficient to obtain homomorphism ω -compactness. Next, we consider homogeneous structures with a finite relational signature, which induce ω -categorical patchworks with a finite signature. This provides us with patchworks with a finite signature that also have homomorphism ω -compactness. What is still missing is decidability of the CSP. This can be achieved by restricting the attention to finitely bounded structures since their CSP is always in NP. Thus, finitely bounded homogeneous structures yield ω -admissible concrete domains. Alternatively, we consider homogeneous structures with a finite relational signature for which we can show by some other means that the CSP is decidable. In this setting, the induced patchwork has a decidable CSP if the structure is a so-called core. Conversely, we prove that every ω -admissible structure is equivalent to a particular homogeneous core in the sense that they both provide the same concrete domain extension of \mathcal{ALC} . The last part of this section investigates closure properties for homogeneity and finite boundedness.

4.1 Homomorphism ω -Compactness via ω -Categoricity

We start by introducing ω -categoricity since it gives us homomorphism ω -compactness “for free.” A structure is *ω -categorical* if its first-order theory has exactly one countable model up to isomorphism. For example, it is well known that \mathfrak{Q} is, up to isomorphism, the only countable dense linear order without lower or upper bound. This result, which clearly implies that \mathfrak{Q} is ω -categorical, is due to Cantor.

For every structure \mathfrak{A} , its automorphisms form a permutation group with composition as binary operation, which we denote by $\text{Aut}(\mathfrak{A})$ (see Theorem 1.2.1 in [50]). Every relation R with a first-order definition in \mathfrak{A} is easily seen to be preserved by $\text{Aut}(\mathfrak{A})$, i.e., $\bar{r} \in R$ implies $h(\bar{r}) \in R$ for every $h \in \text{Aut}(\mathfrak{A})$. For ω -categorical structures, the other direction holds as well.

Theorem 3 (Engeler, Ryll-Nardzewski and Svenonius [50]) *For a countable structure \mathfrak{A} with a countable signature, the following are equivalent:*

1. \mathfrak{A} is ω -categorical.
2. For every k , only finitely many k -ary relations are first-order definable in \mathfrak{A} .
3. Every relation over A preserved by $\text{Aut}(\mathfrak{A})$ is first-order definable in \mathfrak{A} .

The following corollary to this theorem establishes the first important link between model theory and ω -admissibility.

Corollary 2 (Lemma 3.1.5 in [16]) *Every countable ω -categorical structure with a countable signature has homomorphism ω -compactness.*

4.2 Patchworks via Homogeneity and the Amalgamation Property

We show that, in order to obtain patchworks with homomorphism ω -compactness, it is sufficient to consider homogeneous structures. A structure \mathfrak{A} is *homogeneous* if every isomorphism between finite substructures of \mathfrak{A} extends to an automorphism of \mathfrak{A} . We say that a τ -structure admits *quantifier elimination* if for every first-order τ -formula there exists a quantifier-free τ -formula that defines the same relation over this structure.

Theorem 4 ([50]) *A countable relational structure with a finite signature is homogeneous iff it is ω -categorical and admits quantifier elimination.*

The structure \mathfrak{Q} is homogeneous. This can be shown directly without using the theorem, but we will also see later that \mathfrak{Q} admits quantifier elimination. Given finite substructures \mathfrak{B} and \mathfrak{C} of \mathfrak{Q} and an isomorphism between them, we know that B consists of finitely many elements p_1, \dots, p_n and C of the same number of elements q_1, \dots, q_n such that $p_1 < \dots < p_n$, $q_1 < \dots < q_n$, and the isomorphism maps p_i to q_i (for $i = 1, \dots, n$). It is now easy to see that $<$ is also a dense linear order without lower or upper bound on the sets $\{p \mid p < p_1\}$ and $\{q \mid q < q_1\}$, and thus there is an order isomorphism between these sets. The same is true for the pairs of sets $\{p \mid p_i < p < p_{i+1}\}$ and $\{q \mid q_i < q < q_{i+1}\}$, and for the pair $\{p \mid p_n < p\}$ and $\{q \mid q < q_n\}$. Using the isomorphisms between these pairs, we can clearly put together an isomorphism from \mathfrak{Q} to \mathfrak{Q} that extends the original isomorphism from \mathfrak{B} to \mathfrak{C} .

Countable homogeneous structures can be obtained as Fraïssé limits of so-called amalgamation classes. A class \mathcal{K} of relational τ -structures has the *amalgamation property* (AP) if, for all $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2 \in \mathcal{K}$ with $e_1: \mathfrak{A} \hookrightarrow \mathfrak{B}_1$ and $e_2: \mathfrak{A} \hookrightarrow \mathfrak{B}_2$ there exists $\mathfrak{C} \in \mathcal{K}$ with $f_1: \mathfrak{B}_1 \hookrightarrow \mathfrak{C}$ and $f_2: \mathfrak{B}_2 \hookrightarrow \mathfrak{C}$ such that $f_1 \circ e_1 = f_2 \circ e_2$.

Theorem 5 (Fraïssé [50]) *For a class \mathcal{K} of finite τ -structures over a countable signature τ , the following are equivalent:*

1. $\mathcal{K} = \text{Age}(\mathfrak{D})$ for a countable homogeneous structure \mathfrak{D} .
2. \mathcal{K} contains countably many structures up to isomorphism, is closed under isomorphisms and taking substructures, and has AP.

The homogeneous structure \mathfrak{D} in 1. is unique up to isomorphism and called the Fraïssé limit of \mathcal{K}

If \mathcal{K} satisfies 2. in Theorem 5, then we call it an *amalgamation class*. In general, amalgamation classes are required to satisfy one additional condition called the *joint embedding property* (JEP) [50], which we will introduce in Sect. 5. However, since in our case the signature does not contain function symbols, JEP is actually implied by AP and closure under taking substructures.

For our running example $\mathfrak{Q} = (\mathbb{Q}; <)$, the class $\text{Age}(\mathfrak{Q})$ consists of all finite linear orders, and thus by Fraïssé's theorem this class of structures is an amalgamation class. In addition, \mathfrak{Q} is the Fraïssé limit of this class.

Proposition 4 below shows that there is a close connection between AP and the patchwork property. Its proof uses the following lemma.

Lemma 2 *Let $\mathfrak{A}, \mathfrak{B}$ be JEPD τ -structures, $f: \mathfrak{A} \rightarrow \mathfrak{B}$ a homomorphism, and ϕ a k -ary (quantifier and equality)-free formula. Then, $\mathfrak{A} \models \phi(\bar{a})$ iff $\mathfrak{B} \models \phi(f(\bar{a}))$ for every $\bar{a} \in A^k$.*

Proof First, we show that f preserves complements of relations of \mathfrak{A} . Let R be an ℓ -ary relation symbol in τ . Since \mathfrak{A} is JEPD, for every $\bar{a} \in A^\ell$ with $\bar{a} \notin R^{\mathfrak{A}}$, there exists exactly one $\tilde{R} \in \tau \setminus \{R\}$ with $\bar{a} \in \tilde{R}^{\mathfrak{A}}$. This implies $f(\bar{a}) \in \tilde{R}^{\mathfrak{B}}$ since f is a homomorphism. It follows that $f(\bar{a}) \notin R^{\mathfrak{B}}$ because \mathfrak{B} is PD.

Without loss of generality, we assume that ϕ is in DNF, i.e., ϕ is of the form $\phi_1 \vee \dots \vee \phi_n$ where each ϕ_i is a conjunction of possibly negated atomic formulas of the form $R(\bar{x})$ for $R \in \tau$. Since f preserves such atomic formulas and their negations, it also preserves their conjunctions, and thus also disjunctions of such conjunctions. \square

Proposition 4 *A JDJEPD structure \mathfrak{D} is a patchwork iff $\text{Age}(\mathfrak{D})$ has AP.*

Proof For simplicity, every statement indexed by i is supposed to hold for both $i \in \{1, 2\}$. Let τ be the signature of \mathfrak{D} .

“ \Leftarrow ”: Suppose that $\text{Age}(\mathfrak{D})$ has AP. Let $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$ be finite JEPD τ -structures with $e_i: \mathfrak{A} \hookrightarrow \mathfrak{B}_i$ and $h_i: \mathfrak{B}_i \rightarrow \mathfrak{D}$. We must show that there exist $f_i: \mathfrak{B}_i \rightarrow \mathfrak{D}$ with $f_1 \circ e_1 = f_2 \circ e_2$. Let $\tilde{\mathfrak{A}}_1$ and $\tilde{\mathfrak{A}}_2$ be the substructures of \mathfrak{D} on $(h_1 \circ e_1)(A)$ and $(h_2 \circ e_2)(A)$, respectively. Clearly both $\tilde{\mathfrak{A}}_1$ and $\tilde{\mathfrak{A}}_2$ are JDJEPD, because they are substructures of \mathfrak{D} . Note that JD is witnessed in both $\tilde{\mathfrak{A}}_1$ and $\tilde{\mathfrak{A}}_2$ by an identical formula $\phi(x, y)$ inherited from \mathfrak{D} . We claim that there exists an isomorphism from $\tilde{\mathfrak{A}}_1$ to $\tilde{\mathfrak{A}}_2$ which commutes with $h_1 \circ e_1$ and $h_2 \circ e_2$. Consider the map $g: \tilde{A}_1 \rightarrow \tilde{A}_2$ given by $g((h_1 \circ e_1)(a)) := (h_2 \circ e_2)(a)$. By Lemma 2, for every all $a_1, a_2 \in A$, we have

$$\begin{aligned} (h_1 \circ e_1)(a_1) = (h_1 \circ e_1)(a_2) &\text{ iff } \mathfrak{D} \models \phi((h_1 \circ e_1)(a_1), (h_1 \circ e_1)(a_2)) \\ &\text{ iff } \mathfrak{D} \models \phi(a_1, a_2) \\ &\text{ iff } \mathfrak{D} \models \phi((h_2 \circ e_2)(a_1), (h_2 \circ e_2)(a_2)) \\ &\text{ iff } (h_2 \circ e_2)(a_1) = (h_2 \circ e_2)(a_2). \end{aligned}$$

This means that g is well defined and injective. Let $R \in \tau$ be an arbitrary symbol and ℓ its arity. Since \mathfrak{A} and $\tilde{\mathfrak{A}}_1$ are JEPD, by Lemma 2, $h_1 \circ e_1$ preserves the complements of all relations of \mathfrak{A} . Thus, for every $\bar{t} \in A^\ell$, if $(h_1 \circ e_1)(\bar{t}) \in R^{\tilde{\mathfrak{A}}_1}$, then $\bar{t} \in R^{\mathfrak{A}}$ and consequently $(h_2 \circ e_2)(\bar{t}) \in R^{\tilde{\mathfrak{A}}_2}$. This means that g is a homomorphism from $\tilde{\mathfrak{A}}_1$ to $\tilde{\mathfrak{A}}_2$. Since $\tilde{\mathfrak{A}}_1$ and $\tilde{\mathfrak{A}}_2$ are JEPD, by Lemma 2, g also preserves the complements of all relations of $\tilde{\mathfrak{A}}_1$. Hence g is an isomorphism that additionally satisfies $g \circ h_1 \circ e_1 = h_2 \circ e_2$ by its definition. Let \mathfrak{B}_1 and \mathfrak{B}_2 be the substructures of \mathfrak{D} on $h_1(B_1)$ and $h_2(B_2)$, respectively. Now consider the inclusions $\tilde{e}_i: \tilde{\mathfrak{A}}_i \hookrightarrow \mathfrak{B}_i$. Since $\text{Age}(\mathfrak{D})$ has AP, there exists $\mathfrak{C} \in \text{Age}(\mathfrak{D})$ together with $\tilde{f}_i: \mathfrak{B}_i \hookrightarrow \mathfrak{C}$ and $e: \mathfrak{C} \hookrightarrow \mathfrak{D}$ such that $\tilde{f}_1 \circ \tilde{e}_1 = \tilde{f}_2 \circ \tilde{e}_2 \circ g$. We define the homomorphisms $f_i: \mathfrak{B}_i \rightarrow \mathfrak{D}$ by $f_i := e \circ \tilde{f}_i \circ h_i$. Then, for every $a \in A$, we have

$$\begin{aligned} (f_1 \circ e_1)(a) &= (e \circ \tilde{f}_1 \circ h_1 \circ e_1)(a) \\ &= (e \circ \tilde{f}_1 \circ \tilde{e}_1 \circ h_1 \circ e_1)(a) \\ &= (e \circ \tilde{f}_2 \circ \tilde{e}_2 \circ g \circ h_1 \circ e_1)(a) \\ &= (e \circ \tilde{f}_2 \circ \tilde{e}_2 \circ h_2 \circ e_2)(a) \\ &= (e \circ \tilde{f}_2 \circ h_2 \circ e_2)(a) \\ &= (f_2 \circ e_2)(a). \end{aligned}$$

Note that, as inclusions, the mappings \tilde{e}_i are the identity on the elements for which they are defined. The above identities show that \mathfrak{D} is a patchwork.

“ \Rightarrow ”: Suppose that \mathfrak{D} is a patchwork. Let $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$ be finite τ -structures with $e_i: \mathfrak{A} \hookrightarrow \mathfrak{B}_i$ and $h_i: \mathfrak{B}_i \hookrightarrow \mathfrak{D}$. Since \mathfrak{B}_1 and \mathfrak{B}_2 are isomorphic to substructures of \mathfrak{D} , they are clearly JDJEPD. Thus, as \mathfrak{D} is a patchwork, there exist homomorphisms $f_i: \mathfrak{B}_i \rightarrow \mathfrak{D}$ with $f_1 \circ e_1 = f_2 \circ e_2$. Let $\phi(x, y)$ be a formula witnessing JD in both \mathfrak{B}_1 and \mathfrak{B}_2 that is inherited from \mathfrak{D} . By Lemma 2, the operations f_i preserve the complements of all relations of \mathfrak{B}_i , and, for all $b_1, b_2 \in B_i$, we have

$$f_i(b_1) = f_i(b_2) \text{ iff } \mathfrak{D} \models \phi(f_i(b_1), f_i(b_2)) \text{ iff } \mathfrak{D} \models \phi(b_1, b_2) \text{ iff } b_1 = b_2.$$

This means that the operations f_i are embeddings. We obtain AP for $\text{Age}(\mathfrak{D})$ by choosing \mathfrak{C} to be the substructure of \mathfrak{D} on $f_2(B_1) \cup f_1(B_2)$. □

4.3 JDJEPD for ω -Categorical Structures

To apply Proposition 4, we need the structure to be JDJEPD. Given an ω -categorical τ -structure \mathfrak{A} , we can obtain JDJEPD by replacing the original relations with appropriate first-order definable ones, using the results of Theorem 3. The orbit of a tuple $\bar{a} \in A^k$ under the natural action of $\text{Aut}(\mathfrak{A})$ on A^k is the set $\{g(\bar{a}) \mid g \in \text{Aut}(\mathfrak{A})\}$. By Theorem 3, the set of all at most k -ary relations definable in \mathfrak{A} is finite for every $k \in \mathbb{N}$. Since every such set is closed under intersections, it contains finitely many minimal non-empty relations. Since every relation over A that is preserved by all automorphisms of \mathfrak{A} is first-order definable in \mathfrak{A} , these minimal elements are precisely the orbits of tuples over A under the natural action of $\text{Aut}(\mathfrak{A})$.

Definition 5 For a given arity bound $d \geq 2$, the d -ary decomposition of the ω -categorical τ -structure \mathfrak{A} , denoted by $\mathfrak{A}^{\leq d}$, is the relational structure over A whose relations are all orbits of at most d -ary tuples over A under $\text{Aut}(\mathfrak{A})$. We denote the signature of $\mathfrak{A}^{\leq d}$ by $\tau^{\leq d}$.

It is easy to see that $\mathfrak{A}^{\leq d}$ is JDJEPD, and that every at most d -ary relation over A first-order definable in \mathfrak{A} can be obtained as a disjunction of atomic $\tau^{\leq d}$ -formulas.

As an example, consider the ω -categorical structure \mathfrak{Q} . The orbits of k -tuples of elements of \mathbb{Q} can be defined by quantifier-free formulas that are conjunctions of atomic formulas of the form $x_i = x_j$ or $x_i < x_j$. For example, the orbit of the tuple $(2, 3, 2, 5)$ consists of all tuples $(q_1, q_2, q_3, q_4) \in \mathbb{Q}^4$ that satisfy the formula $x_1 < x_2 \wedge x_1 = x_3 \wedge x_2 < x_4$ if x_i is replaced by q_i for $i = 1, \dots, 4$. The first-order definable k -ary relations in \mathfrak{Q} are obtained as unions of these orbits, where the defining formula is then the disjunction of the formulas defining the respective orbits. Since these formulas are quantifier-free, this also shows that \mathfrak{Q} admits quantifier elimination.

We have seen that, to obtain JDJEPD, we actually need to take the d -ary decomposition of a given ω -categorical structure, rather than the structure itself. Fortunately, homogeneity transfers from \mathfrak{D} to $\mathfrak{D}^{\leq d}$.

Proposition 5 Let \mathfrak{D} be a countable homogeneous structure with a finite relational signature τ . Then $\mathfrak{D}^{\leq d}$ is homogeneous for every d that exceeds or is equal to the maximal arity of a symbol from τ .

Proof By Theorem 4, \mathfrak{D} has quantifier elimination. Note that the relations of $\mathfrak{D}^{\leq d}$ and \mathfrak{D} are first-order interdefinable, which implies $\text{Aut}(\mathfrak{D}^{\leq d}) = \text{Aut}(\mathfrak{D})$ by Theorem 3. This

shows in particular that $\mathfrak{D}^{\leq d}$ is ω -categorical. Every first-order $\tau^{\leq d}$ -formula ϕ defines a relation in $\mathfrak{D}^{\leq d}$ that has a first-order definition ϕ' in \mathfrak{D} . We can assume that ϕ' is quantifier-free due to Theorem 4. We replace every atomic formula $\psi(\bar{x})$ in ϕ' by $\bigvee_{i=1}^n R_i(\bar{x})$ with $R_1, \dots, R_n \in \tau^{\leq d}$, where $R_1^{\mathfrak{D}^{\leq d}} \cup \dots \cup R_n^{\mathfrak{D}^{\leq d}}$ is the unique decomposition of $\psi^{\mathfrak{D}}$ into orbits of k -tuples over D under $\text{Aut}(\mathfrak{D})$. The resulting formula is a quantifier-free first-order definition of $\phi^{\mathfrak{D}^{\leq d}}$ in $\mathfrak{D}^{\leq d}$. Thus $\mathfrak{D}^{\leq d}$ has quantifier elimination as well, which means that it is homogeneous due to Theorem 4. \square

4.4 Finitely Bounded Structures have a Decidable CSP

Above, we have described model-theoretic properties that provide us with all the ingredients needed for ω -admissibility, except for decidability of the CSP. Finding model-theoretic conditions that guarantee decidability of the CSP for infinite structures is a very broad topic with many open questions. Here we focus on a well-known condition that ensures that the CSP is decidable in NP and the first-order theory in PSPACE.

For a class \mathcal{N} of τ -structures (called *bounds* or *forbidden patterns*), we denote by $\text{Forb}_e(\mathcal{N})$ the class of all finite τ -structures not embedding any member of \mathcal{N} . Following the terminology in [20], we say that a relational structure \mathfrak{A} is *finitely bounded* if its signature is finite and $\text{Age}(\mathfrak{A}) = \text{Forb}_e(\mathcal{N})$ for a finite set of bounds \mathcal{N} . The following lemma provides a second, arguably more practical, definition of finite boundedness.

Lemma 3 *A relational structure \mathfrak{A} with a finite signature τ is finitely bounded iff $\text{Age}(\mathfrak{A})$ is the class of all finite models of some universal τ -sentence $\Phi(\mathfrak{A})$.*

Proof “ \Rightarrow ”: Let $\text{Age}(\mathfrak{A}) = \text{Forb}_e(\mathcal{N})$ for $\mathcal{N} = \{\mathfrak{C}_1, \dots, \mathfrak{C}_k\}$. For every $i, 1 \leq i \leq k$, we can write down a quantifier-free formula $\phi_{\mathfrak{C}_i}$ with free variables c_1, \dots, c_{n_i} , where $\{c_1, \dots, c_{n_i}\}$ is the domain of \mathfrak{C}_i , that describes \mathfrak{C}_i up to isomorphism. Then we set

$$\Phi(\mathfrak{A}) := \bigwedge_{1 \leq i \leq k} \forall c_1, \dots, c_{n_i}. \neg \phi_{\mathfrak{C}_i}(c_1, \dots, c_{n_i}).$$

“ \Leftarrow ”: Given a universal τ -sentence $\Phi(\mathfrak{A})$, we define \mathcal{N} as the set of all finite τ -structures \mathfrak{C} of size at most n that do not satisfy $\Phi(\mathfrak{A})$, where n is the number of variables in $\Phi(\mathfrak{A})$. Then $\text{Age}(\mathfrak{A}) = \text{Forb}_e(\mathcal{N})$ clearly holds. \square

The structure \mathfrak{Q} is finitely bounded. To show this, we can use the set \mathcal{N} consisting of the following four structures: the self-loop, the 2-cycle, the 3-cycle, and two isolated vertices. We must show that $\text{Age}(\mathfrak{Q}) = \text{Forb}_e(\mathcal{N})$. Clearly, none of the structures in \mathcal{N} embeds into a linear order, which shows $\text{Age}(\mathfrak{Q}) \subseteq \text{Forb}_e(\mathcal{N})$. Conversely, assume that \mathfrak{A} is an element of $\text{Forb}_e(\mathcal{N})$. We must show that $<^{\mathfrak{A}}$ is a linear order. Since \mathcal{N} contains the self-loop, we have $(a, a) \notin <^{\mathfrak{A}}$ for all $a \in A$, which shows that $<^{\mathfrak{A}}$ is irreflexive. For distinct elements $a, b \in A$, we must have $a <^{\mathfrak{A}} b$ or $b <^{\mathfrak{A}} a$ since otherwise the structure consisting of two isolated vertices could be embedded into \mathfrak{A} . This shows that any two distinct elements are comparable w.r.t. $<^{\mathfrak{A}}$. To show that $<^{\mathfrak{A}}$ is transitive, assume that $a <^{\mathfrak{A}} b$ and $b <^{\mathfrak{A}} c$ holds. Since the 2-cycle does not embed into \mathfrak{A} , a and c must be distinct, and are thus comparable. We cannot have $c >^{\mathfrak{A}} a$ since then we could embed the 3-cycle into \mathfrak{A} . Consequently, we must have $a <^{\mathfrak{A}} c$, which proves transitivity. This shows that \mathfrak{A} is a linear order. As formula $\Phi(\mathfrak{Q})$ we can take the conjunction of the usual axioms defining linear orders.

Finitely bounded structures are useful in the context of this paper due to the following proposition.

Proposition 6 *Let \mathfrak{D} be a finitely bounded structure.*

1. $\text{CSP}(\mathfrak{D})$ is in NP.
2. If \mathfrak{D} is homogeneous, then $\text{Th}(\mathfrak{D})$ is in PSPACE.

The first result is stated in [16, 19], and the second result is stated in [56, 59]. We include a full proof of both in the appendix.

Proposition 6 applies not only to a given finitely bounded homogeneous structure \mathfrak{D} , but also to its d -ary decomposition $\mathfrak{D}^{\leq d}$. This is a direct consequence of the following result.

Proposition 7 *Let \mathfrak{A} be a finitely bounded homogeneous structure and \mathfrak{B} a structure with the same domain and finitely many relations that are first-order definable in \mathfrak{A} . Then \mathfrak{B} is a reduct of a finitely bounded homogeneous structure.*

Proof Let $\tilde{\mathfrak{A}}$ be the expansion of \mathfrak{A} by the relations of \mathfrak{B} , where we assume that the signatures of \mathfrak{A} and \mathfrak{B} are disjoint. By Theorem 4, each of the new relations has a quantifier-free definition in \mathfrak{A} . Consequently, we can choose any universal sentence $\Phi(\mathfrak{A})$ for $\text{Age}(\mathfrak{A})$ and extend it with universal sentences defining the relations of \mathfrak{B} , which yields a universal sentence that shows finite boundedness of $\tilde{\mathfrak{A}}$. The structure $\tilde{\mathfrak{A}}$ is homogeneous since an isomorphism between two finite substructures of $\tilde{\mathfrak{A}}$ induces an isomorphism between their reducts to the signature of \mathfrak{A} , which extends to an automorphism of \mathfrak{A} by homogeneity of \mathfrak{A} . This is also an automorphism of $\tilde{\mathfrak{A}}$ since automorphisms preserve first-order definable relations. Now we are done as \mathfrak{B} is a reduct of $\tilde{\mathfrak{A}}$. □

4.5 Finitely Bounded Homogeneous Structures Yield ω -Admissible Concrete Domains

We are now ready to formulate the main results of this section.

Theorem 6 *Let \mathfrak{D} be a finitely bounded homogeneous relational structure with at most d -ary relations for some $d \geq 2$. Then $\mathfrak{D}^{\leq d}$ is ω -admissible.*

Proof It follows directly from the definition of d -ary decompositions that $\mathfrak{D}^{\leq d}$ is JDJEPD. By Proposition 5, $\mathfrak{D}^{\leq d}$ is homogeneous. By Theorem 4, $\mathfrak{D}^{\leq d}$ is ω -categorical. Thus \mathfrak{D} has homomorphism ω -compactness by Corollary 2. By Theorem 5, $\text{Age}(\mathfrak{D}^{\leq d})$ has AP. Thus $\mathfrak{D}^{\leq d}$ is a patchwork by Proposition 4. By Proposition 7, Lemma 1, and Proposition 6, $\text{CSP}(\mathfrak{D}^{\leq d})$ is in NP. Hence $\mathfrak{D}^{\leq d}$ is ω -admissible. □

This theorem, together with Theorem 1, immediately yields decidability for concept satisfiability in $\mathcal{ALCC}_{\sqrt{t}}^d(\mathfrak{D}^{\leq d})$. The following corollary shows that we can even allow for arbitrary first-order definable relations with arity bounded by d in the concrete domain. The idea for proving this result is to reduce concept satisfiability in $\mathcal{ALCC}_{f_0}^d(\mathfrak{D})$ to concept satisfiability in $\mathcal{ALCC}_{\sqrt{t}}^d(\mathfrak{D}^{\leq d})$. We know that every at most d -ary relation over D first-order definable in \mathfrak{D} can be obtained as a disjunction of atomic formulas built using the signature of $\mathfrak{D}^{\leq d}$. What still needs to be shown is that, given a first-order formula in the signature of \mathfrak{D} with at most d free variables, this disjunction can effectively be computed.

Corollary 3 *Let \mathfrak{D} be a reduct of a finitely bounded homogeneous relational structure with at most d -ary relations for some $d \geq 2$. Then concept satisfiability in $\mathcal{ALCC}_{f_0}^d(\mathfrak{D})$ w.r.t. $T\text{Boxes}$ is decidable.*

Proof Let τ be the signature of \mathfrak{D} . We claim that satisfiability of $\mathcal{ALC}_{fo}^d(\mathfrak{D})$ concepts w.r.t. TBoxes can be reduced to satisfiability of $\mathcal{ALC}_{\forall+}^d(\mathfrak{D}^{\leq d})$ concepts w.r.t. TBoxes. For this purpose, we need to replace first-order τ -formulas ϕ in concrete domain constructors $\forall p_1, \dots, p_k.\phi$ or $\exists p_1, \dots, p_k.\phi$ with disjunctions ψ of atomic formulas in the signature $\tau^{\leq d}$ of $\mathfrak{D}^{\leq d}$. By Theorem 4 together with Theorem 3, the (finitely many) relations in $\tau^{\leq d}$ have quantifier-free definitions in \mathfrak{D} . Since d and \mathfrak{D} are fixed, we can make a list consisting of the quantifier-free definitions for each of them in constant time. Given a first-order τ -formulas ϕ with k free variables, let ψ_1, \dots, ψ_m be the quantifier-free definitions in \mathfrak{D} for all the k -ary relations of $\tau^{\leq d}$ that we have listed before. We test, for every $i \in [m]$, whether $\mathfrak{D} \models \exists \bar{y}.\phi(\bar{y}) \wedge \psi_i(\bar{y})$, which is possible in PSPACE by Proposition 6. By selecting those $\psi_{i_1}, \dots, \psi_{i_s}$ that tested positively, we know that, for every $\bar{a} \in D^k$, $\mathfrak{D} \models \phi(\bar{a})$ iff $\mathfrak{D} \models \bigvee_{r=1}^s \psi_{i_r}(\bar{a})$. We replace each $\psi_{i_r}(\bar{y})$ with $R(\bar{y})$, where R is the unique k -ary relation symbol from $\tau^{\leq d}$ for which $\mathfrak{D} \models \psi_{i_r}(\bar{a})$ iff $\mathfrak{D}^{\leq d} \models R(\bar{a})$. This yields the desired formula ψ that replaces ϕ . Now the claim follows from Theorem 6 and Theorem 1. \square

Example 2 The examples for ω -admissible concrete domains given in [65] were RCC8 and Allen’s interval algebra, for which the patchwork property is proved “by hand” in [65]. Given our Theorem 6, we obtain these results as a consequence of known results from model theory. It was shown in [21] that RCC8 has a representation by a homogeneous structure \mathfrak{R} with a finite relational signature (see Theorem 2 in [21]). Since $\text{Age}(\mathfrak{R})$ has a finite universal axiomatization (see Definition 3 in [21]), \mathfrak{R} is finitely bounded. For Allen’s interval algebra, it was shown in [48] that it has a representation by a homogeneous structure \mathfrak{A} with a finite relational signature (see the second example on page 270 in [48]). Since $\text{Age}(\mathfrak{A})$ has a finite universal axiomatization (see the composition table from Figure 4 in [1]), \mathfrak{A} is finitely bounded. The structure $\mathfrak{Q} = (\mathbb{Q}; <)$ we used as our running example also satisfies the preconditions of Theorem 6, and thus Corollary 3 yields decidability of $\mathcal{ALC}_{fo}^d(\mathfrak{Q})$ with TBoxes. For \mathfrak{Q} extended just with $>, \leq, \geq, =, \neq$, decidability was proved in [62], using an automata-based procedure. Our results show that there is also a tableau-based decision procedure for this logic.

4.6 Homogeneous Cores with Decidable CSP Yield ω -Admissible Concrete Domains

Here, we consider the situation where we have a homogeneous relational structure \mathfrak{D} with finitely many at most d -ary relations that is not necessarily finitely bounded, but which we can show (by some other means) to have a decidable CSP. In this setting, we obtain decidability for $\mathcal{ALC}_{\exists+}^d(\mathfrak{D})$ under the additional assumption that \mathfrak{D} is a core. A structure \mathfrak{D} is a *core* if every endomorphism of \mathfrak{D} is a self-embedding of \mathfrak{D} . It is easy to see that this applies to our running example $\mathfrak{Q} = (\mathbb{Q}; <)$. The structure $\tilde{\mathfrak{Q}} := (\mathbb{Q}; \leq)$, on the other hand, is not a core because it has the trivial endomorphism $x \mapsto 0$ that is not a self-embedding of $\tilde{\mathfrak{Q}}$. Among ω -categorical structures, cores are characterized by the following condition. A countable ω -categorical structure \mathfrak{D} with a countable signature is a core if and only if every relation with an existential definition in \mathfrak{D} has an \exists^+ definition in \mathfrak{D} [24]. Consider the binary inequality relation \neq over \mathbb{Q} which clearly has an existential definition in both \mathfrak{Q} and $\tilde{\mathfrak{Q}}$. Since the complements of the basic relations defined by $=$ and $<$ in \mathfrak{Q} have a positive quantifier-free definition in \mathfrak{Q} , every relation with an existential definition in \mathfrak{Q} has an \exists^+ definition in \mathfrak{Q} . Thus, \mathfrak{Q} is a core according to the characterization of ω -categorical cores from above. However, there can be no \exists^+ definition of \neq in $\tilde{\mathfrak{Q}}$ because relations with an \exists^+ definition are always preserved by all endomorphisms and $x \mapsto 0$ does not preserve \neq .

If \mathfrak{D} is a homogeneous core, then the orbits of tuples over D under $\text{Aut}(\mathfrak{D})$ are pp definable in \mathfrak{D} [15]. As an easy consequence of this fact, we obtain the following sufficient condition for ω -admissibility.

Theorem 7 *Let \mathfrak{D} be a homogeneous core with finitely many at most d -ary relations for some $d \geq 2$ and decidable CSP. Then $\mathfrak{D}^{\leq d}$ is ω -admissible.*

Proof It follows directly from the definition of d -ary decompositions that $\mathfrak{D}^{\leq d}$ is JDJEPD. By Proposition 5, $\mathfrak{D}^{\leq d}$ is homogeneous. By Theorem 4, $\mathfrak{D}^{\leq d}$ is ω -categorical. Thus \mathfrak{D} has homomorphism ω -compactness by Lemma 2. By Theorem 5, $\text{Age}(\mathfrak{D}^{\leq d})$ has AP. Thus, $\mathfrak{D}^{\leq d}$ is a patchwork by Proposition 4. By the results of [15], orbits of tuples over D under $\text{Aut}(\mathfrak{D})$ (i.e., the relations of $\mathfrak{D}^{\leq d}$) are pp definable in \mathfrak{D} . Thus, Lemma 1 yields $\text{CSP}(\mathfrak{D}^{\leq d}) \leq_{\text{PTIME}} \text{CSP}(\mathfrak{D})$. Hence, $\mathfrak{D}^{\leq d}$ is ω -admissible. \square

Let \mathfrak{D} be a structure as in the above theorem. By showing that concept satisfiability in $\mathcal{ALC}_{\exists+}^d(\mathfrak{D})$ can be reduced to concept satisfiability in $\mathcal{ALC}_{\forall+}^d(\mathfrak{D}^{\leq d})$, we obtain the following decidability result.

Corollary 4 *Let \mathfrak{D} be a homogeneous core with finitely many at most d -ary relations for some $d \geq 2$ and a decidable CSP. Then concept satisfiability in $\mathcal{ALC}_{\exists+}^d(\mathfrak{D})$ w.r.t. TBoxes is decidable.*

Proof Since satisfiability of $\mathcal{ALC}_{\forall+}^d(\mathfrak{D}^{\leq d})$ concepts w.r.t. TBoxes is decidable by Theorems 7 and 1, it is sufficient to reduce concept satisfiability w.r.t. TBoxes in $\mathcal{ALC}_{\exists+}^d(\mathfrak{D})$ to this problem. As in the proof of Corollary 3, we do this by showing how existential positive formulas ϕ occurring in concrete domain constructors can be replaced by disjunctions ψ of atomic formulas in the signature of $\mathfrak{D}^{\leq d}$. By the results of [15], the relations of $\mathfrak{D}^{\leq d}$ have pp definitions in \mathfrak{D} . Since d and \mathfrak{D} are fixed, we can make a list consisting of the pp definitions for each of them in constant time. Given an existential positive τ -formula ϕ with $k \leq d$ free variables, let ψ_1, \dots, ψ_m be the pp definitions in \mathfrak{D} for all the k -ary relations of $\mathfrak{D}^{\leq d}$ that we have listed before. Since $\text{CSP}(\mathfrak{D})$ is decidable, we can decide for $i \in [m]$ whether $\mathfrak{D} \models \exists \bar{y}.(\psi_i \wedge \phi)(\bar{y})$. In fact, deciding whether an existential positive sentence is true in \mathfrak{D} only differs from solving $\text{CSP}(\mathfrak{D})$ in a non-deterministic step that deals with disjunction. By selecting those $\psi_{i_1}, \dots, \psi_{i_s}$ that tested positively, we know that $\mathfrak{D} \models \phi(\bar{a})$ iff $\mathfrak{D} \models \bigvee_{r=1}^s \psi_{i_r}(\bar{a})$ holds for every $\bar{a} \in D^k$. Now we replace each $\psi_{i_r}(\bar{y})$ with $R(\bar{y})$, where R is the unique k -ary relation symbol from the signature of $\mathfrak{D}^{\leq d}$ that satisfies $\mathfrak{D} \models \psi_{i_r}(\bar{a})$ iff $\mathfrak{D}^{\leq d} \models R(\bar{a})$. This yields the desired formula ψ , which completes the reduction. \square

4.7 Coverage of the Developed Sufficient Conditions

The next example demonstrates that Theorem 7 and Corollary 4 cover structures to which Theorem 6 and Corollary 3 do not apply. In fact, since the latter consider finitely bounded structures, whose CSP is in NP by Proposition 6, they cannot provide us with ω -admissible concrete domains whose CSP has a higher complexity. Theorem 7 and Corollary 4 make no assumption on the complexity of the CSP: they only require that the CSP is decidable. However, for these results to apply, the structure needs to be a homogeneous core.

Example 3 The paper [44] provides us with examples of structures that are homogeneous cores and whose CSP is considerably more complex than NP. Such structures are called CSP monsters in [44]. To be more precise, Theorem 8 in [44] shows that, for every complexity

class \mathcal{C} for which there exist $\text{CONP}^{\mathcal{C}}$ -complete problems, there exists a homogeneous structure $\mathfrak{H}_{\mathcal{C}}$ with a finite signature such that $\text{CSP}(\mathfrak{H}_{\mathcal{C}})$ is $\text{CONP}^{\mathcal{C}}$ -complete. By Theorem 4 together with Theorem 3.6.23 and Proposition 3.6.24 from [16], for every such structure $\mathfrak{H}_{\mathcal{C}}$, there exists an up to isomorphism unique homogeneous core $\mathfrak{C}_{\mathcal{C}}$ with the property that $\mathfrak{H}_{\mathcal{C}}$ maps homomorphically to $\mathfrak{C}_{\mathcal{C}}$ and vice versa. In particular, this implies that $\text{CSP}(\mathfrak{H}_{\mathcal{C}}) = \text{CSP}(\mathfrak{C}_{\mathcal{C}})$. It follows from Theorem 7 that these structures yield ω -admissible concrete domains whose CSPs have arbitrarily high complexity. Recall that all previously known examples of ω -admissible concrete domains were finitely bounded (see Example 2), and thus their CSPs are in NP by Proposition 6. However, already $\mathfrak{C}_{\text{NExpTime}}$ cannot possibly be even a reduct of a finitely bounded structure due to Proposition 6 because $\text{NP} \subsetneq \text{NExpTime} \subseteq \text{coNP}^{\text{NExpTime}}$. Consequently, the homogeneous cores induced by the CSP monsters of [44] provide us with previously unknown ω -admissible concrete domains that are not covered by Theorem 6 and Corollary 3.

Next, we investigated the coverage of Theorem 7. This theorem states that every homogeneous core with a finite signature and a decidable CSP yields an ω -admissible structure via its d -ary decomposition. The following two results show that, if we are interested in extensions of \mathcal{ALC} of the form $\mathcal{ALC}_{\forall}^d(\mathfrak{D})$, then ω -admissible structures yields the same extensions of \mathcal{ALC} as homogeneous cores with decidable CSPs.

Theorem 8 *Let \mathfrak{B} be an ω -admissible τ -structure. Then there exists an (up to isomorphism) unique countable homogeneous τ -structure \mathfrak{A} that is a core with decidable CSP and embeds the same countable structures as \mathfrak{B} , i.e., $\mathfrak{C} \hookrightarrow \mathfrak{A}$ iff $\mathfrak{C} \hookrightarrow \mathfrak{B}$ for every countable structure \mathfrak{C} .*

Proof Since \mathfrak{B} is JDJEPD and a patchwork, $\text{Age}(\mathfrak{B})$ has AP by Proposition 4. Since τ is finite, $\text{Age}(\mathfrak{B})$ contains only countably many structures up to isomorphism, and thus is an amalgamation class. By Theorem 5, there exists a countable homogeneous structure \mathfrak{A} with $\text{Age}(\mathfrak{A}) = \text{Age}(\mathfrak{B})$. Next, we show that $\mathfrak{C} \hookrightarrow \mathfrak{A}$ iff $\mathfrak{C} \hookrightarrow \mathfrak{B}$ holds for every countable structure \mathfrak{C} .

“ \Leftarrow ”: Let \mathfrak{C} be a countable τ -structure that embeds into \mathfrak{B} . By Theorem 4, \mathfrak{A} is ω -categorical. It is known that ω -categorical structures satisfy an even stronger property than homomorphism ω -compactness, which we refer to as *embedding ω -compactness* (Lemma 3.1.5 in [16]). This property guarantees an embedding from a given countable structure if there exists an embedding from every structure in its age. Since $\text{Age}(\mathfrak{C}) \subseteq \text{Age}(\mathfrak{B}) = \text{Age}(\mathfrak{A})$, we conclude that $\mathfrak{C} \hookrightarrow \mathfrak{A}$.

“ \Rightarrow ”: Let \mathfrak{C} be a countable τ -structure and $e : \mathfrak{C} \hookrightarrow \mathfrak{A}$ be an embedding. If we can show that there is an embedding $f : \mathfrak{A} \hookrightarrow \mathfrak{B}$, then we are done since we can use the composition of e and f as embedding from \mathfrak{C} to \mathfrak{B} . Since \mathfrak{A} is countable, \mathfrak{B} has homomorphism compactness, and $\text{Age}(\mathfrak{B}) = \text{Age}(\mathfrak{A})$, there exists a homomorphism $f : \mathfrak{A} \rightarrow \mathfrak{B}$. We show that f is in fact an embedding.

We claim that \mathfrak{A} is JEPD since \mathfrak{B} is so. In fact, assume the \mathfrak{A} is not PD. Then there are distinct k -ary relations R_1, R_2 and a k -tuple \bar{a} such that $\bar{a} \in R_1^{\mathfrak{A}} \cap R_2^{\mathfrak{A}}$. Thus, the substructure of \mathfrak{A} consisting of the elements of \bar{a} is an element of $\text{Age}(\mathfrak{A})$ that is not PD. But then $\text{Age}(\mathfrak{B}) = \text{Age}(\mathfrak{A})$ contains a structure that is not PD, which yields a contradiction since \mathfrak{B} is PD. The fact that JE transfers from \mathfrak{B} to \mathfrak{A} can be shown similarly. If $\phi(x, y)$ is the formula witnessing that \mathfrak{B} is JD, then one can also show in a similar way that this formula witnesses JD of \mathfrak{A} as well.

Since we now know that both \mathfrak{A} and \mathfrak{B} are JEPD, we can apply Lemma 2, which yields that the homomorphism f preserves also the complements of all relations of \mathfrak{A} . In addition, it

preserves the formula ϕ witnessing JD. Thus, the following holds for all $a_1, a_2 \in A$: $a_1 = a_2$ iff $\mathfrak{A} \models \phi(a_1, a_2)$ iff $\mathfrak{B} \models \phi(f(a_1), f(a_2))$ iff $f(a_1) = f(a_2)$. Thus f is an embedding, which concludes the proof of “ \Rightarrow ”.

Since \mathfrak{A} is JDJEPD, every endomorphism of \mathfrak{A} is a self-embedding of \mathfrak{A} , which can be shown by a similar argument as above. Thus \mathfrak{A} is a core.

Decidability of the CSP transfers from \mathfrak{B} to \mathfrak{A} since the two CSPs coincide. If \mathfrak{F} is a finite structure with $\mathfrak{F} \rightarrow \mathfrak{A}$, then the image \mathfrak{C} of \mathfrak{F} in \mathfrak{A} is a finite (and thus countable) structure such that $\mathfrak{F} \rightarrow \mathfrak{C} \hookrightarrow \mathfrak{A}$. But then $\mathfrak{C} \hookrightarrow \mathfrak{B}$, and thus $\mathfrak{F} \rightarrow \mathfrak{B}$. The inclusion in the other direction can be shown in the same way. \square

Since the structures \mathfrak{A} and \mathfrak{B} in the theorem have the same signature, the DLs $\mathcal{ALC}_{\vee+}^d(\mathfrak{A})$ and $\mathcal{ALC}_{\vee+}^d(\mathfrak{B})$ have the same syntax. We show that they also have the same semantics when it comes to concept satisfiability.

Corollary 5 *Let \mathfrak{A} and \mathfrak{B} be as in Theorem 8 and let d be the largest arity of an atomic τ -formula. Then a concept C is satisfiable w.r.t. a TBox \mathcal{T} in $\mathcal{ALC}_{\vee+}^d(\mathfrak{A})$ iff it is satisfiable in $\mathcal{ALC}_{\vee+}^d(\mathfrak{B})$ w.r.t. \mathcal{T} .*

Proof First note that, since \mathfrak{B} is ω -admissible, the DL $\mathcal{ALC}_{\vee+}^d(\mathfrak{B})$ has the *countable model property*, i.e., a concept C is satisfiable in this logic w.r.t. a TBox \mathcal{T} iff there is a finite model of \mathcal{T} in which C is interpreted as a non-empty set. This is a direct consequence of the proof of Theorem 1 in [7] because the model constructed in this proof is countable. Now suppose that \mathcal{I} is a countable interpretation witnessing that C is satisfiable w.r.t. \mathcal{T} in $\mathcal{ALC}_{\vee+}^d(\mathfrak{B})$. Let \mathfrak{C} be the substructure of \mathfrak{B} on $\{b \in B \mid \text{there is } f \in N_f \text{ and } a \in \Delta^{\mathcal{I}} \text{ such that } (a, b) \in f^{\mathcal{I}}\}$. By Theorem 8, there exists an embedding $e: \mathfrak{C} \hookrightarrow \mathfrak{A}$. Since e is an embedding, we can obtain an interpretation witnessing that C is satisfiable w.r.t. \mathcal{T} in $\mathcal{ALC}_{\vee+}^d(\mathfrak{A})$ from \mathcal{I} by replacing every $(a, d) \in f^{\mathcal{I}}$ with $(a, e(d))$.

The argument used above also works the other way round. Here \mathfrak{A} is a homogeneous core with decidable CSP. The proof of Corollary 4 shows that satisfiability of concepts w.r.t. TBoxes in $\mathcal{ALC}_{\vee+}^d(\mathfrak{A})$ can be reduced to satisfiability of concepts w.r.t. TBoxes in $\mathcal{ALC}_{\vee+}^d(\mathfrak{D})$ for an ω -admissible concrete domain \mathfrak{D} . As above, we can show that this yields the countable model property for $\mathcal{ALC}_{\vee+}^d(\mathfrak{A})$. The rest of the proof is exactly as for the other direction. \square

The following example shows that equi-satisfiability no longer holds if we replace \vee^+ with fo in Corollary 5, i.e., the logics $\mathcal{ALC}_{\text{fo}}^d(\mathfrak{A})$ and $\mathcal{ALC}_{\text{fo}}^d(\mathfrak{B})$ may have a different semantics.

Example 4 The *random graph* is the unique countably infinite homogeneous undirected graph $\mathfrak{G} = (G; E^{\mathfrak{G}})$ such that $\text{Age}(\mathfrak{G})$ consists of all finite undirected graphs [50]. Note that $\text{Age}(\mathfrak{G})$ is defined by the universal sentence $\forall x. (E(x, x) \Rightarrow \text{ff}) \wedge \forall x, y. (E(x, y) \Rightarrow E(y, x))$. Thus, by Lemma 3, \mathfrak{G} is finitely bounded. It also has the *extension property*: if X and Y are disjoint finite subsets of G , then there exists a vertex $v \in G \setminus (X \cup Y)$ that has an edge in \mathfrak{G} to each vertex from X and to none from Y . To see this, let \mathfrak{A} be the extension of the substructure of \mathfrak{G} on $X \cup Y$ by a vertex u that has an edge to each vertex from X and to none from Y . Then there exists an embedding $e: \mathfrak{A} \hookrightarrow \mathfrak{G}$. Since \mathfrak{G} is homogeneous and the map $f: X \cup Y \rightarrow e(X \cup Y), a \mapsto e(a)$ is an isomorphism between its finite substructures, there exists $\tilde{f} \in \text{Aut}(\mathfrak{G})$ extending f . Then $v := \tilde{f}^{-1}(e(u))$ has the desired property.

Consider the direct product \mathfrak{H} of \mathfrak{G} with itself. It is easy to see that $\text{Age}(\mathfrak{G}) = \text{Age}(\mathfrak{H})$. The inclusion from left to right holds since \mathfrak{H} contains an isomorphic copy of \mathfrak{G} , and the one from right to left since $\text{Age}(\mathfrak{G})$ contains all finite undirected graphs. The equality of the two ages implies that $\text{Age}(\mathfrak{H})$ has AP by Theorem 5. Also, by Theorem 3 and Theorem 4, \mathfrak{H} is

ω -categorical because its relations are first-order definable in the so-called full product of \mathfrak{G} with itself and homogeneous structures are closed under building full products. We will introduce the full product and show that it preserves homogeneity in Sect. 4.8.

However, \mathfrak{H} does not have the extension property. To see this, let a, b, c be three distinct vertices in G and set $X := \{(a, b), (b, c)\}$, and $Y := \{(a, c)\}$. Suppose that there exists $(u, v) \in H$ that has an edge in \mathfrak{H} to each vertex from X and to none from Y . By the definition of \mathfrak{H} as the direct product of \mathfrak{G} with itself, there is an edge in \mathfrak{G} from u to a and from v to c . But then there is an edge from (u, v) to (a, c) in \mathfrak{H} , which contradicts to our previous assumption. This implies that \mathfrak{G} and \mathfrak{H} are not isomorphic since the extension property is clearly preserved under isomorphism. We conclude that \mathfrak{H} is not homogeneous since homogeneous structures are uniquely determined up to isomorphism by their age due to Theorem 5. Note that we have just shown with this example that homogeneous structures are not closed under building direct products.

Now consider the expansion \mathfrak{A} of \mathfrak{G} with two new relation symbols R_1, R_2 , where R_1 is interpreted as the diagonal relation Eq^G and R_2 as $G^2 \setminus (\text{Eq}^G \cup E^{\mathfrak{G}})$. Likewise we construct the expansion \mathfrak{B} of \mathfrak{H} with R_1, R_2 . Let \mathfrak{C} be a substructure of \mathfrak{A} and $\tilde{\mathfrak{C}}$ its $\{E\}$ -reduct. Since $\text{Age}(\mathfrak{G}) = \text{Age}(\mathfrak{H})$, there exists an isomorphism f from $\tilde{\mathfrak{C}}$ to some substructure $\tilde{\mathfrak{D}}$ of \mathfrak{H} . Let \mathfrak{D} be the substructure of \mathfrak{B} on the domain \tilde{D} of $\tilde{\mathfrak{D}}$. We claim that f is also an isomorphism from \mathfrak{C} to \mathfrak{D} . We have $\bar{x} \in R_1^{\mathfrak{C}} = \text{Eq}^C$ iff $f(\bar{x}) \in R_1^{\mathfrak{D}} = \text{Eq}^D$ because f is bijective. Moreover, we have

$$\bar{x} \in R_2^{\mathfrak{C}} \text{ iff } \bar{x} \notin (E^{\tilde{\mathfrak{C}}} \cup \text{Eq}^C) \text{ iff } f(\bar{x}) \notin (E^{\tilde{\mathfrak{D}}} \cup \text{Eq}^D) \text{ iff } f(\bar{x}) \in R_2^{\mathfrak{D}}.$$

We conclude that $\text{Age}(\mathfrak{A}) \subseteq \text{Age}(\mathfrak{B})$. Using an analogous argument, we can show $\text{Age}(\mathfrak{A}) \supseteq \text{Age}(\mathfrak{B})$, and thus $\text{Age}(\mathfrak{A}) = \text{Age}(\mathfrak{B})$. Since every homomorphism from a finite structure has a finite range, $\text{Age}(\mathfrak{A}) = \text{Age}(\mathfrak{B})$ implies $\text{CSP}(\mathfrak{A}) = \text{CSP}(\mathfrak{B})$ (see the last paragraph in the proof of Theorem 8).

The following two facts are direct consequences of $R_1^{\mathfrak{A}}$ and $R_2^{\mathfrak{A}}$ being first-order definable in \mathfrak{G} . First, \mathfrak{A} is homogeneous since \mathfrak{G} is homogeneous and first-order definable relations are preserved by automorphisms. Thus, $\text{Age}(\mathfrak{A})$ has AP by Theorem 5. Second, \mathfrak{A} is a reduct of a finitely bounded structure by Proposition 7, and thus $\text{CSP}(\mathfrak{A})$ is in NP by Proposition 6.

By definition, \mathfrak{B} is JDJEPD. Since $\text{Age}(\mathfrak{B}) = \text{Age}(\mathfrak{A})$ has AP, the structure \mathfrak{B} is a patchwork by Proposition 4. By Lemma 2, \mathfrak{B} has homomorphism ω -compactness since it is ω -categorical. This is the case since \mathfrak{H} is ω -categorical and expansions by first-order definable relations do not change the automorphism group. Since \mathfrak{A} and \mathfrak{B} are both countable but not isomorphic, we conclude using Theorem 5 that \mathfrak{B} is ω -admissible but not homogeneous. Since $\text{Age}(\mathfrak{A}) = \text{Age}(\mathfrak{B})$ and \mathfrak{A} is countable and homogeneous, it must be the homogeneous core of \mathfrak{B} from Theorem 8.

It follows from our proof that \mathfrak{H} does not have the extension property that the concept A is satisfiable w.r.t. the TBox

$$\{A \sqsubseteq \exists f. (x_1 = x_1 \wedge \forall x, y, z \exists u. E(u, x) \wedge E(u, y,) \wedge \neg E(u, z))\}$$

in $\mathcal{ALC}_{\text{fo}}^2(\mathfrak{A})$, but not in $\mathcal{ALC}_{\text{fo}}^2(\mathfrak{B})$.

4.8 Closure Properties of Finitely Bounded Homogeneous Structures

We have seen above that finitely bounded homogeneous structures provide us with ω -admissible concrete domains. Closure properties allow us to construct new ω -admissible concrete domains from ones satisfying these properties.

For instance, when modeling concepts in a DL with a concrete domain, it is often useful to be able to refer to specific elements c of the domain, i.e., to have unary predicate symbols $=_c$ that are interpreted as $\{c\}$. For example, when using the ω -admissible concrete domain Ω of our running example, one can compare two numbers (e.g., describing the ages of two individuals), but one cannot state that the value of a feature must be equal to some fixed number (e.g., that a person’s age is 17). For a finitely bounded homogeneous structure (such as Ω), adding *finitely many* such singleton predicates is harmless since the class of reducts of finitely bounded homogeneous structures is closed under expansion by finitely many predicates of the form $=_c$.

Proposition 8 ([19]) *Let \mathfrak{A} be a finitely bounded homogeneous structure. Any expansion of \mathfrak{A} by a relation of the form $\{c\}$ for $c \in A$ is a reduct of a finitely bounded homogeneous structure.*

We have seen in Proposition 7 that this class is also closed under taking expansions by first-order definable relations.

It would also be useful to be able to refer to predicates of different concrete domains (say RCC8 and Allen) when defining concepts. This can sometimes be achieved by using the disjoint union. The *union* of a family $(\mathfrak{A}_i)_{i \in I}$ of τ -structures is the τ -structure $\bigcup_{i \in I} \mathfrak{A}_i$ over $\bigcup_{i \in I} A_i$ such that $R^{\bigcup_{i \in I} \mathfrak{A}_i} = \bigcup_{i \in I} R^{\mathfrak{A}_i}$ for each $R \in \tau$. This union is called *disjoint* if $A_i \cap A_j = \emptyset$ for all distinct $i, j \in I$.

In [3], it was shown that admissible concrete domains are closed under disjoint union. We can show the corresponding result for finitely bounded homogeneous structures. In our definition of the disjoint union, we have assumed that the component structures $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ have the same signature, but disjoint domains. In [3], the signatures of the structures are assumed to be disjoint as well (as is, e.g., the case for RCC8 and Allen). The case of disjoint signatures can, however, be reduced to the case of a common signature: we simply expand the structures to the union of their signatures by interpreting relation symbols not belonging to the respective signature as the empty set. Since empty relations can be defined by first-order formulas, such an expansion by empty relations leaves homogeneity and finite boundedness intact (see Proposition 7). A proof of the following proposition can be found in the appendix.

Proposition 9 *Let $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ be finitely bounded homogeneous structures over a common signature τ , but with disjoint domains. Then their disjoint union $\bigcup_{i=1}^k \mathfrak{A}_i$ is a reduct of a finitely bounded homogeneous structure.*

Using disjoint union to refer to several concrete domain works well if the paths employed in concrete domain constructors contain only functional roles, which is the case considered in [3], but it is not appropriate if non-functional roles occur in paths, as in the present paper. This is illustrated by the following example.

Example 5 If we want to refer to time and location of an event, we can use the disjoint union of RCC8 and Allen, employing two feature names *time* and *location*. If *succ* is a functional role, then the concept description

$$\text{Event} \sqcap \exists \text{succ}.\text{Event} \sqcap \exists \text{time}, \text{succ time}.\text{Before}(x_1, x_2) \sqcap \exists \text{location}, \text{succ location}.\text{EC}(x_1, x_2)$$

describes an event e that takes place before its unique successor event e' , which happens in a region that is externally connected to e . However, if *succ* is not functional, then the above concept description does not express that e has a successor event e' that satisfies both the temporal and the spatial constraint. Instead, there could be two different successor events, one satisfying the temporal constraint and the other the spatial one.

To overcome this problem, we propose to use the so-called full product [16].⁵ Let $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ be relational structures with disjoint signatures τ_1, \dots, τ_k , and let $=_1, \dots, =_k$ be fresh binary symbols such that, for every $i \in [k]$, $=_i$ is interpreted as Eq^{A_i} over \mathfrak{A}_i . We assume in the following that the relation $=_i$ is part of the signature of \mathfrak{A}_i . This assumption is without loss of generality since the equality predicate is first-order definable, and thus extending a homogeneous structure with an explicit relation symbol for it leaves the structure finitely bounded and homogeneous (see Proposition 7).

The *full product* of $\mathfrak{A}_1, \dots, \mathfrak{A}_k$, denoted by $\mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$, has as its domain the Cartesian product $A := A_1 \times \dots \times A_k$ and as its signature the union of the signatures τ_i . The relations of $\mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$ are defined by $R^{\mathfrak{A}} := \{(\bar{a}_1, \dots, \bar{a}_n) \in A^n \mid (\bar{a}_1[i], \dots, \bar{a}_n[i]) \in R^{\mathfrak{A}_i}\}$ for every $i \in [k]$ and every n -ary relation $R \in \tau_i$.

Taking the full product of structures preserves homogeneity and finite boundedness, and thus the prerequisites for Theorem 6 and Corollary 3 to apply (see the appendix for a proof).

Proposition 10 *Let $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ be structures with disjoint relational signatures τ_1, \dots, τ_k such that, for $i \in [k]$, τ_i contains the symbol $=_i$, which is defined in \mathfrak{A}_i as Eq^{A_i} .*

1. *If $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ are homogeneous, then $\mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$ is also homogeneous.*
2. *If $\mathfrak{A}_1, \dots, \mathfrak{A}_k$ are finitely bounded, then $\mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$ is also finitely bounded.*

Coming back to Example 5, we can use a feature *time&location* that maps into the full product of Allen and RCC8 to describe an event e that has some successor event e' (among possibly others) such that e takes place before e' and the regions where e and e' happen are externally connected:

$\text{Event} \sqcap \exists \text{succ.Event} \sqcap \exists \text{time\&location, succ time\&location.}(\text{Before}(x_1, x_2) \wedge \text{EC}(x_1, x_2)).$

5 A Model-Theoretic Analysis of p-Admissibility

Recall that a structure \mathfrak{D} is p-admissible if it is convex and validity of Horn implications in \mathfrak{D} is tractable. As argued at the end of Sect. 3.2.2, developing algebraic conditions that characterize tractability is way beyond the scope of this paper. For this reason, we will concentrate on algebraic conditions that ensure convexity. We will see, however, that for finitely bounded convex structures we obtain tractability for free.

5.1 Convexity via Square Embeddings

Convex structures can be characterized using the square embedding condition introduced in the next theorem. Basically, this condition says that the square of every finite substructure of \mathfrak{B} embeds into \mathfrak{B} . However, since we allow the signature to be infinite, the exact formulation of the property is a bit more complicated. Note that the direction “2 \Rightarrow 1” is a slightly more general version of a result commonly known as McKinsey’s lemma [49].

Theorem 9 *For a structure \mathfrak{B} with a (not necessarily finite) relational signature τ , the following are equivalent:*

1. \mathfrak{B} is convex.

⁵ We have seen in Example 4 that employing the usual direct product of the structures does not work since it does not preserve homogeneity.

2. $\mathcal{K} = \text{Age}(\mathfrak{B})$ satisfies the square embedding property: for every finite $\sigma \subseteq \tau$ and every $\mathfrak{A} \in \mathcal{K}$, there is $\mathfrak{C} \in \mathcal{K}$ such that the σ -reducts of \mathfrak{A}^2 and \mathfrak{C} coincide.

Proof “2 \Rightarrow 1”: Suppose to the contrary that the implication $\forall x_1, \dots, x_n. (\phi \Rightarrow \psi)$ is valid in \mathfrak{B} , where ϕ is a conjunction of atomic formulas and ψ is a disjunction of atomic formulas ψ_1, \dots, ψ_k , but we also have $\mathfrak{B} \not\models \forall x_1, \dots, x_n. (\phi \Rightarrow \psi_i)$ for every $i \in [k]$. Without loss of generality, we assume that $\phi, \psi_1, \dots, \psi_k$ all have the same free variables x_1, \dots, x_n , some of which might not influence their truth value. For every $i \in [k]$, there exists a tuple $\bar{t}_i \in B^n$ such that

$$\mathfrak{B} \models \phi(\bar{t}_i) \wedge \neg \psi_i(\bar{t}_i). \tag{*}$$

We show by induction on i that, for every $i \in [k]$, there exists a tuple $\bar{s}_i \in B^n$ that satisfies the *induction hypothesis*

$$\mathfrak{B} \models \phi(\bar{s}_i) \wedge \neg \bigvee_{\ell \in [i]} \psi_\ell(\bar{s}_i). \tag{†}$$

In the *base case* ($i = 1$), it follows from (*) that $\bar{s}_1 := \bar{t}_1$ satisfies (†).

In the *induction step* ($i \rightarrow i + 1$), let $\bar{s}_i \in B^n$ be any tuple that satisfies (†). Let $\sigma \subseteq \tau$ be the finite set of relation symbols occurring in the implication $\forall x_1, \dots, x_n. (\phi \Rightarrow \psi)$, and let \mathfrak{A}_i be the substructure of \mathfrak{B} on the set $\{\bar{s}_i[1], \bar{t}_{i+1}[1], \dots, \bar{s}_i[n], \bar{t}_{i+1}[n]\}$. Then $\mathfrak{A}_i \models \phi(\bar{s}_i)$ and $\mathfrak{A}_i \models \phi(\bar{t}_{i+1})$, and thus $\mathfrak{A}_i^2 \models \phi(\bar{s}_i \times \bar{t}_{i+1})$ where $\bar{s}_i \times \bar{t}_{i+1} := ((\bar{s}_i[1], \bar{t}_{i+1}[1]), \dots, (\bar{s}_i[n], \bar{t}_{i+1}[n]))$.

By 2., there exists a structure $\mathfrak{C}_i \in \text{Age}(\mathfrak{B})$ whose σ -reduct coincides with \mathfrak{A}_i^2 , which implies that $\mathfrak{C}_i \models \phi(\bar{s}_i \times \bar{t}_{i+1})$. Let f_i be the embedding of \mathfrak{C}_i into \mathfrak{B} . Since ϕ is a conjunction of atomic σ -formulas and f_i is a homomorphism, we have that $\mathfrak{B} \models \phi(f_i(\bar{s}_i \times \bar{t}_{i+1}))$. Suppose that $\mathfrak{B} \models \psi_{i+1}(f_i(\bar{s}_i \times \bar{t}_{i+1}))$. Since f_i is an embedding, we obtain $\mathfrak{C}_i \models \psi_{i+1}(\bar{s}_i \times \bar{t}_{i+1})$, and thus $\mathfrak{A}_i \models \psi_{i+1}(\bar{t}_{i+1})$. This implies $\mathfrak{B} \models \psi_{i+1}(\bar{t}_{i+1})$, which contradicts (*). Similarly, we can show that assuming $\mathfrak{B} \models \psi_j(f_i(\bar{s}_i \times \bar{t}_{i+1}))$ for some $j \leq i$ leads to a contradiction with (†). We conclude that $\bar{s}_{i+1} := f_i(\bar{s}_i \times \bar{t}_{i+1})$ satisfies (†).

Since $\mathfrak{B} \models \forall x_1, \dots, x_n. (\phi \Rightarrow \psi)$, the existence of a tuple $\bar{s}_i \in B^n$ that satisfies (†) for $i = k$ leads to a contradiction. This completes the proof of “2 \Rightarrow 1” of our theorem.

Before we proceed with the proof of “1 \Rightarrow 2”, let us take a closer look at the contraposition of the convexity condition. Suppose that we have a conjunction ϕ of atomic formulas and tuples \bar{r} and \bar{s} over B together with disjunctions $\psi_{\bar{r}}$ and $\psi_{\bar{s}}$ of atomic formulas such that $\mathfrak{B} \models (\phi \wedge \neg \psi_{\bar{r}})(\bar{r})$ and $\mathfrak{B} \models (\phi \wedge \neg \psi_{\bar{s}})(\bar{s})$. Then clearly there must exist a tuple \bar{t} over B such that $\mathfrak{B} \models (\phi \wedge \neg \psi_{\bar{r}} \wedge \neg \psi_{\bar{s}})(\bar{t})$ as otherwise $\mathfrak{B} \models \forall x_1, \dots, x_n. (\phi \Rightarrow (\psi_{\bar{r}} \vee \psi_{\bar{s}}))$, but neither $\mathfrak{B} \models \forall x_1, \dots, x_n. (\phi \Rightarrow \psi_{\bar{r}})$ nor $\mathfrak{B} \models \forall x_1, \dots, x_n. (\phi \Rightarrow \psi_{\bar{s}})$ is true (which contradicts convexity).

We are now ready to prove “1 \Rightarrow 2”. Let σ be a finite subset of τ and $\mathfrak{A} \in \text{Age}(\mathfrak{B})$. In addition, let $\{(r_1, s_1), \dots, (r_n, s_n)\}$ be the domain of \mathfrak{A}^2 . Consider the tuples $\bar{r} := (r_1, \dots, r_n)$ and $\bar{s} := (s_1, \dots, s_n)$. Let $\phi(x_1, \dots, x_n)$ be the conjunction of all atomic σ -formulas such that $\mathfrak{A}^2 \models \phi((r_1, s_1), \dots, (r_n, s_n))$, i.e., we consider all atomic σ -formulas built using a relation symbol from σ (or the equality predicate) and containing variables from $\{x_1, \dots, x_n\}$, assign (r_i, s_i) to the variable x_i , and take those atomic σ -formulas for which the corresponding tuple of elements of \mathfrak{A}^2 belongs to the respective relation in \mathfrak{A}^2 .

Clearly, the tuples \bar{r} and \bar{s} both satisfy ϕ in \mathfrak{B} since the projection to a single coordinate is a homomorphism from \mathfrak{A}^2 to \mathfrak{B} . Now let $\psi_{\bar{r}}$ be the disjunction of all atomic σ -formulas that do not hold on the tuple \bar{r} in \mathfrak{B} . Analogously, let $\psi_{\bar{s}}$ be the disjunction of all atomic

σ -formulas that do not hold on the tuple \bar{s} in \mathfrak{B} . Without loss of generality $|A| > 1$, and thus both disjunctions are non-empty.

We have that $\mathfrak{B} \models \phi \wedge \neg\psi_{\bar{r}}(\bar{r})$ and $\mathfrak{B} \models \phi \wedge \neg\psi_{\bar{s}}(\bar{s})$. Since \mathfrak{B} is convex, there must exist a tuple \bar{t} such that $\mathfrak{B} \models \phi \wedge \neg\psi_{\bar{r}}(\bar{t}) \wedge \neg\psi_{\bar{s}}(\bar{t})$. Now consider the map f that sends, for every $i \in [n]$, the tuple (r_i, s_i) to $\bar{t}[i]$. Clearly f is a homomorphism from the σ -reduct of \mathfrak{A} to the σ -reduct of \mathfrak{B} because $\mathfrak{B} \models \phi(\bar{t})$. Moreover, f is an embedding because, if ψ is a single atomic σ -formula, then $\mathfrak{B} \models \psi(\bar{t})$ only if $\mathfrak{B} \models \psi(\bar{r})$ and $\mathfrak{B} \models \psi(\bar{s})$. We define \mathfrak{C} as the substructure of \mathfrak{B} on $f(A^2)$. \square

Using Theorem 9, we can obtain a statement similar to that of Theorem 5, where convexity replaces homogeneity and the square embedding property together with the joint embedding property replaces AP. A class \mathcal{K} of relational τ -structures has the *joint embedding property* (JEP) if, for every $\mathfrak{B}_1, \mathfrak{B}_2 \in \mathcal{K}$ there exists $\mathfrak{C} \in \mathcal{K}$ such that $\mathfrak{B}_i \hookrightarrow \mathfrak{C}$ for $i \in \{1, 2\}$. Recall our definition of the square embedding property from Theorem 9.

Corollary 6 *For a class \mathcal{K} of finite τ -structures, the following are equivalent:*

1. $\mathcal{K} = \text{Age}(\mathfrak{D})$ for a countable convex structure \mathfrak{D} .
2. \mathcal{K} contains countably many structures up to isomorphism, is closed under isomorphisms and building substructures, has JEP, and satisfies the square embedding property.

Proof The direction “1 \Rightarrow 2” is a direct consequence of Theorem 9 since classes of the form $\text{Age}(\mathfrak{D})$ for a relational structure \mathfrak{D} trivially satisfy JEP. The direction “2 \Rightarrow 1” follows from Theorem 6.1.1 in [50] and Theorem 9. In fact, Theorem 6.1.1 in [50] implies that a class of finite relational structures \mathcal{K} that is closed under building substructures and has JEP is of the form $\mathcal{K} = \text{Age}(\mathfrak{D})$ for a countable structure \mathfrak{D} . An application of Theorem 9 then yields convexity of \mathfrak{D} . \square

In contrast to homogeneous structures, countable convex structures are in general *not* uniquely determined up to isomorphism by their age. The random graph can again serve as a counterexample.

Example 6 The random graph \mathfrak{G} introduced in Example 4 is convex since $\text{Age}(\mathfrak{G})$ satisfies the square embedding condition. In fact, since \mathfrak{G} embeds every finite undirected graph, it also embeds \mathfrak{A}^2 for any undirected graph \mathfrak{A} . The direct product \mathfrak{H} of \mathfrak{G} with itself is thus also convex since $\text{Age}(\mathfrak{H}) = \text{Age}(\mathfrak{G})$. However, we have seen in Example 4 that \mathfrak{G} and \mathfrak{H} are not isomorphic. It is not hard to see that \mathfrak{G} is actually p -admissible. Instead of proving this directly, we will show it as a consequence of Theorem 12 below.

Corollary 6 can also be used to construct p -admissible concrete domains from structures whose CSP is in P . This further substantiates our remark at the end of Sect. 3.2.2 that characterizing all p -admissible concrete domains is at least as hard as characterizing all tractable CSPs.

Definition 6 The *canonical database* $\text{DB}(\exists\bar{x}.\phi(\bar{x}))$ for a satisfiable equality-free pp τ -sentence $\exists\bar{x}.\phi(\bar{x})$ is the τ -structure whose domain consists of the quantified variables \bar{x} and whose relations are specified by the quantifier-free part ϕ .

Proposition 11 *For every structure \mathfrak{B} with a finite relational signature τ , there exists a countable convex τ -structure \mathfrak{D} such that $\text{CSP}(\mathfrak{B}) = \text{Age}(\mathfrak{D}) = \text{CSP}(\mathfrak{D})$. Moreover, \mathfrak{D} is p -admissible iff $\text{CSP}(\mathfrak{B})$ is in P .*

Proof It is easy to see that $\mathcal{K} := \text{CSP}(\mathfrak{B})$ satisfies 2. in Corollary 6; for JEP we use the fact that \mathcal{K} is closed under taking disjoint unions, and for the square embedding property we use the fact that \mathcal{K} is closed under taking second powers. By Corollary 6, there exists a countable convex structure \mathfrak{D} such that $\text{Age}(\mathfrak{D}) = \text{CSP}(\mathfrak{B})$. Regarding the claim that $\text{Age}(\mathfrak{D}) = \text{CSP}(\mathfrak{D})$, first note that the inclusion from left to right is trivial. Now, assume that $\mathfrak{A} \in \text{CSP}(\mathfrak{D})$, i.e., there is a homomorphism $f: \mathfrak{A} \rightarrow \mathfrak{D}$. Then the substructure \mathfrak{C} of \mathfrak{D} on $f(A)$ belongs to $\text{Age}(\mathfrak{D}) = \text{CSP}(\mathfrak{B})$. This yields a homomorphism from \mathfrak{C} to \mathfrak{B} , and thus from \mathfrak{A} to \mathfrak{B} . Consequently, $\mathfrak{A} \in \text{CSP}(\mathfrak{B}) = \text{Age}(\mathfrak{D})$.

To show the second claim of the proposition, let $\forall \bar{x}.(\phi \Rightarrow \psi)$ be a Horn implication. Without loss of generality, we assume that ϕ does not contain equality axioms since we can remove them by identifying variables in such equality axioms in ϕ and ψ . We claim that, under this assumption, $\mathfrak{D} \models \exists \bar{x}.(\phi \wedge \neg \psi)$ iff $\mathfrak{D} \models \exists \bar{x}.\phi$ and ψ does not occur as a conjunct in ϕ .

The only if direction is trivial. For the if direction, note that, by a standard result in database theory, $\mathfrak{D} \models \exists \bar{x}.\phi$ iff the canonical database $\text{DB}(\exists \bar{x}.\phi)$ homomorphically maps to \mathfrak{D} [37]. Since ψ does not occur as a conjunct in ϕ , this atomic formula does not hold in $\text{DB}(\exists \bar{x}.\phi)$. This implies that ψ does not hold in \mathfrak{D} since $\text{DB}(\exists \bar{x}.\phi)$ embeds into \mathfrak{D} because $\text{CSP}(\mathfrak{D}) = \text{Age}(\mathfrak{D})$.

We conclude that $\mathfrak{D} \models \forall \bar{x}.(\phi \Rightarrow \psi)$ iff $\text{DB}(\exists \bar{x}.\phi(\bar{x})) \rightarrow \mathfrak{D}$ and ϕ contains ψ as a conjunct. This can be tested in polynomial time iff $\text{CSP}(\mathfrak{D})$ is in P. □

5.2 Convex ω -Categorical Structures

For countably infinite ω -categorical structures, the characterization of convexity of Theorem 9 can be improved to the following simpler statement.

Theorem 10 *For a countably infinite ω -categorical relational structure \mathfrak{B} with a countable signature τ , the following are equivalent:*

1. \mathfrak{B} is convex.
2. \mathfrak{B}^2 embeds into \mathfrak{B} .

Proof The direction “2 \Rightarrow 1” follows immediately from Theorem 9 since $\mathfrak{B}^2 \hookrightarrow \mathfrak{B}$ obviously implies that $\text{Age}(\mathfrak{B})$ satisfies the square embedding property. Note that for this direction, ω -categoricity of \mathfrak{B} is not required.

The proof of “1 \Rightarrow 2” combines the proof of this direction for Theorem 9 with the following two facts, which are implied by ω -categoricity of \mathfrak{B} . First, there exists an embedding from \mathfrak{B}^2 to \mathfrak{B} iff there exists an embedding from \mathfrak{A} to \mathfrak{B} for every $\mathfrak{A} \in \text{Age}(\mathfrak{B}^2)$ (see, e.g., Lemma 3.1.5 in [16]). Second, to deal with the fact that τ may be infinite we can use Theorem 3, which ensures that, for every $k \geq 1$, there are only finitely many inequivalent k -ary formulas over \mathfrak{B} consisting of a single atomic τ -formula. □

In the CSP literature, one can find two interesting examples of countably infinite ω -categorical structures that satisfy Condition 2 of Theorem 10: atomless Boolean algebras and countably infinite-dimensional vector spaces over finite fields. Since the CSP for atomless Boolean algebras is NP-complete [13], this example does not provide us with a p-admissible concrete domain; but the vector space example does.

As shown in [22], the relational representation $\mathfrak{V}_q = (V_q; R^+, R^{s_0}, \dots, R^{s_{q-1}})$ of the countably infinite-dimensional vector space over a finite field $\text{GF}(q)$ is ω -categorical, satisfies $\mathfrak{V}_q^2 \cong \mathfrak{V}_q$, and its CSP is decidable in polynomial time, even if the complements of all

predicates are added. Here R^+ is a ternary predicate corresponding to addition of vectors, and the R^{s_i} are binary predicates corresponding to scalar multiplication of a vector with the element s_i of $\text{GF}(q)$. These properties are preserved if we add finitely many unary predicates R^{e_i} that correspond to unit vectors e_1, \dots, e_k .

Corollary 7 *The structure \mathfrak{V}_q expanded with predicates R^{e_1}, \dots, R^{e_k} for unit vectors e_1, \dots, e_k is p -admissible.*

Proof We have $\mathfrak{V}_q^2 \cong \mathfrak{V}_q$, and thus both structures are vector spaces over $\text{GF}(q)$ of countably infinite dimension. Now if we fix finitely many unit vectors $e_1, \dots, e_k \in V_q$ by expanding \mathfrak{V}_q with the unary predicates R^{e_1}, \dots, R^{e_k} , we can still extend the map which sends (e_i, e_i) to e_i for each $i \in [k]$ to a bijection between bases of both vector spaces. This bijection then naturally extends to an isomorphism from $(\mathfrak{V}_q, R^{e_1}, \dots, R^{e_k})^2$ to $(\mathfrak{V}_q, R^{e_1}, \dots, R^{e_k})$. Thus, Theorem 10 yields convexity of $(\mathfrak{V}_q, R^{e_1}, \dots, R^{e_k})$. The CSP in its expansion by inequality and the complements of all relations can be solved, similarly as in the Gaussian elimination algorithm, by iterated elimination of variables from equations and subsequent search for unsatisfiable equalities and/or inequalities between unit vectors (e.g., $e_1 \neq e_1$ or $e_1 = e_2$) (see [22] for details). This implies that testing validity of Horn implications in $(\mathfrak{V}_q, R^{e_1}, \dots, R^{e_k})$ is tractable. We conclude that $(\mathfrak{V}_q, R^{e_1}, \dots, R^{e_k})$ is p -admissible. \square

For the case $q = 2$, the vectors in V_q are one-sided infinite tuples of zeros and ones containing only finitely many ones, which can be viewed as representing finite subsets of \mathbb{N} . For example, $(0, 1, 1, 0, 1, 0, 0, \dots)$ represents the set $\{1, 2, 4\}$. Thus, if we use \mathfrak{V}_2 as concrete domain, the features assign finite sets of natural numbers to individuals. For example, assume that the feature *dages* assigns the set of ages of daughters to a person, and *sages* the set of ages of sons. Then $\exists \text{dages, sages, zero. } R^+(x_1, x_2, x_3)$ describes persons that, for every age, have either both a son and a daughter of this age, or no child at all of this age. The feature *zero* is supposed to point to the zero vector, which can, e.g., be enforced using the GCI $\top \sqsubseteq \exists \text{zero, zero, zero. } R^+(x_1, x_2, x_3)$. If e_1 is the unit vector $(0, 1, 0, 0, \dots)$ and e_4 is the unit vector $(0, 0, 0, 0, 1, 0, 0, \dots)$, then the concept $\exists \text{one, four, dages. } R^+(x_1, x_2, x_3)$ describes humans that have daughters of age one and four, and of no other age, if the TBox contains the GCI $\top \sqsubseteq \exists \text{one. } R^{e_1}(x_1) \sqcap \exists \text{four. } R^{e_4}(x_1)$.

5.3 Convex Numerical Structures

Outside of the scope of ω -categoricity, we exhibit two new p -admissible concrete domain that are respectively based on the real and the rational numbers, and whose predicates are defined by linear equations. Let $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ be the relational structure over \mathbb{R} that has, for every linear equation system $A\bar{x} = \bar{b}$ over \mathbb{Q} , a relation consisting of all its solutions in \mathbb{R} . We define $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$ as the substructure of $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ on \mathbb{Q} . For example, using the matrix $A = \begin{pmatrix} 2 & 1 & -1 \end{pmatrix}$ and the vector $\bar{b} = (0)$ one obtains the ternary relation $\{(p, q, r) \in \mathbb{Q}^3 \mid 2p + q = r\}$ in $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$. Our proof of the fact these two structures are p -admissible uses the following simple observation about pp definable relations.

Lemma 4 *Let \mathfrak{D} be a structure for which there exists an isomorphism $f: \mathfrak{D}^2 \rightarrow \mathfrak{D}$, and let $\{R_i \mid i \in I\}$ be a family of relations that are pp definable in \mathfrak{D} . Then f is an isomorphism from $(\mathfrak{D}, \{R_i \mid i \in I\})^2$ to $(\mathfrak{D}, \{R_i \mid i \in I\})$.*

Proof By a standard result in model theory, if R is pp definable in \mathfrak{D} , then f is also a homomorphism from $(\mathfrak{D}, R)^2$ to (\mathfrak{D}, R) (e.g., Proposition 5.2.2 in [16]). Since f is bijective, it only

remains to show that f is even an embedding from $(\mathfrak{D}, R)^2$ to (\mathfrak{D}, R) . Let $\phi(x_1, \dots, x_k) := \exists x_{k+1}, \dots, x_\ell. \psi(x_1, \dots, x_\ell)$ be the pp formula that defines R in \mathfrak{D} , where ψ is the quantifier-free part of ϕ . Let $\bar{r} \in R$ be an arbitrary tuple of the form $\bar{r} = f(\bar{r}_1, \bar{r}_2)$ for some $\bar{r}_1, \bar{r}_2 \in D^k$. Then there exists $\bar{s} \in D^{\ell-k}$ such that $\mathfrak{D} \models \psi(\bar{r}[1], \dots, \bar{r}[k], \bar{s}[1], \dots, \bar{s}[\ell-k])$. Since f is surjective, there exist $\bar{s}_1, \bar{s}_2 \in D^{\ell-k}$ such that $\bar{s} = f(\bar{s}_1, \bar{s}_2)$. Since f is an embedding from \mathfrak{D}^2 to \mathfrak{D} , we have $\mathfrak{D} \models \psi(\bar{r}_i[1], \dots, \bar{r}_i[k], \bar{s}_i[1], \dots, \bar{s}_i[\ell-k])$ for both $i \in \{1, 2\}$. This means that $\bar{r}_1, \bar{r}_2 \in R$, which confirms our claim. \square

Theorem 11 *The relational structures $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ and $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$ are p-admissible.*

Proof To prove this theorem for \mathbb{R} , we start with the well-known fact that $(\mathbb{R}; +, 0)^2$ and $(\mathbb{R}; +, 0)$ are isomorphic [55]. Such an isomorphism exists because $(\mathbb{R}; +, 0)^2$ and $(\mathbb{R}; +, 0)$ are both vector spaces over \mathbb{Q} whose dimensions are uncountably infinite and of the same cardinality. Thus every bijective map from a basis of $(\mathbb{R}; +, 0)^2$ to a basis of $(\mathbb{R}; +, 0)$ extends to an isomorphism. Now we simply choose any two bases of $(\mathbb{R}; +, 0)^2$ and $(\mathbb{R}; +, 0)$ such that the first basis contains $(1, 1)$ and the second basis contains 1 . Then we choose an arbitrary bijection from the first basis to the second basis that sends $(1, 1)$ to 1 . This bijection extends to an isomorphism $f : (\mathbb{R}; +, 0, 1)^2 \rightarrow (\mathbb{R}; +, 0, 1)$. It is easy to see that every relation of $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ can be defined in $(\mathbb{R}; +, 0, 1)$ using a pp formula. By Lemma 4, f is an isomorphism from $\mathfrak{D}_{\mathbb{R}, \text{lin}}^2$ to $\mathfrak{D}_{\mathbb{R}, \text{lin}}$, which implies that $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ is convex by Theorem 9.

Now recall that validity of Horn implications in $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ can be tested in polynomial time if the CSP for $(\mathfrak{D}_{\mathbb{R}, \text{lin}}^-, \neq)$ is in P. It is easy to see that f is a homomorphism from $(\mathfrak{D}_{\mathbb{R}, \text{lin}}^-, \neq)^2$ to $(\mathfrak{D}_{\mathbb{R}, \text{lin}}^-, \neq)$. It follows from Corollary 5.10 in [23] that both the CSP and validity of Horn implications in $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ are decidable in polynomial time. We conclude that $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ is p-admissible.

For \mathbb{Q} , we cannot employ the same argument since $(\mathbb{Q}; +, 0)^2$ does not even embed into $(\mathbb{Q}; +, 0)$. Instead, we use the well-known fact that the structures $(\mathbb{Q}; +, 0)$ and $(\mathbb{R}; +, 0)$ satisfy the same first-order-sentences [55] to show that convexity of $\mathfrak{D}_{\mathbb{R}, \text{lin}}$ implies convexity of $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$.

We claim that a stronger statement is true, namely, that $\text{Th}(\mathbb{Q}; +, 0, 1) = \text{Th}(\mathbb{R}; +, 0, 1)$. Let ϕ be an arbitrary first-order sentence in the signature of $(\mathbb{R}; +, 0, 1)$. We obtain the formula $\psi(x)$ in the signature of $(\mathbb{R}; +, 0)$ by replacing the constant 1 in ϕ by a fresh free variable x , i.e., $(\mathbb{R}; +, 0, 1) \models \phi$ iff $(\mathbb{R}; +, 0) \models \psi(1)$. For every $c \in \mathbb{R} \setminus \{0\}$, the map $x \mapsto cx$ is an automorphism of $(\mathbb{R}; +, 0)$ that sends 1 to c . Since $\{x \in \mathbb{R} \mid (\mathbb{R}; +, 0) \models \psi(x)\}$ has a first-order definition in $(\mathbb{R}; +, 0)$, it is preserved by all automorphisms of $(\mathbb{R}; +, 0)$ [50]. Now we distinguish the following two cases. If $(\mathbb{R}; +, 0) \models \psi(0)$, then $(\mathbb{R}; +, 0, 1) \models \phi$ iff $(\mathbb{R}; +, 0) \models \exists x. \psi(x)$. Otherwise $(\mathbb{R}; +, 0, 1) \models \phi$ iff $(\mathbb{R}; +, 0) \models \exists x. (\neg(x = 0) \wedge \psi(x))$. Using an analogous argument we have either $(\mathbb{Q}; +, 0, 1) \models \phi$ iff $(\mathbb{Q}; +, 0) \models \exists x. \psi(x)$ in the case where $(\mathbb{Q}; +, 0) \models \psi(0)$, or $(\mathbb{Q}; +, 0, 1) \models \phi$ iff $(\mathbb{Q}; +, 0) \models \exists x. (\neg(x = 0) \wedge \psi(x))$. Since ϕ was chosen arbitrarily, we conclude that indeed $\text{Th}(\mathbb{Q}; +, 0, 1) = \text{Th}(\mathbb{R}; +, 0, 1)$.

Since the relations of $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$ are definable in $(\mathbb{Q}; +, 0, 1)$ using the same pp formulas as for their counterparts in $\mathfrak{D}_{\mathbb{R}, \text{lin}}$, and $\text{Th}(\mathbb{Q}; +, 0, 1) = \text{Th}(\mathbb{R}; +, 0, 1)$, we conclude that $\mathfrak{D}_{\mathbb{Q}, \text{lin}}$ is p-admissible as well. \square

In Sect. 3 we have introduced the structure $\mathfrak{D}_{\mathbb{Q}^2, \text{aff}}$ and have shown in Proposition 2 that subsumption w.r.t. TBoxes is undecidable in $\mathcal{EL}(\mathfrak{D}_{\mathbb{Q}^2, \text{aff}})$. Using Theorems 2 and 11 we can now show that subsumption w.r.t. TBoxes is tractable in $\mathcal{EL}[\mathfrak{D}_{\mathbb{Q}^2, \text{aff}}]$ since $\mathfrak{D}_{\mathbb{Q}^2, \text{aff}}$ is p-admissible.

Corollary 8 *The relational structure $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$ is p -admissible.*

Proof First, note that the CSP and validity of Horn implications in $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$ can be reduced in linear time to the same problems for $\mathcal{D}_{\mathbb{Q}, \text{lin}}$.

It remains to show that $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$ is convex. Let σ be a finite subset of the signature of $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$ and let $\mathfrak{A} \in \text{Age}(\mathcal{D}_{\mathbb{Q}^2, \text{aff}})$. We may assume without loss of generality that \mathfrak{A} is a substructure of $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$. It is sufficient to show that the σ -reduct of \mathfrak{A}^2 embeds into the σ -reduct of $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$.

For every relation $R_{M, \bar{v}}$ of $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$ we consider the 4-ary relation $\{(\bar{x}[1], \bar{x}[2], \bar{y}[1], \bar{y}[2]) \in \mathbb{Q}^4 \mid \bar{y} = M\bar{x} + \bar{v}\}$ of $\mathcal{D}_{\mathbb{Q}, \text{lin}}$ which we denote by $\tilde{R}_{M, \bar{v}}$. Consider the substructure $\tilde{\mathfrak{A}}$ of $\mathcal{D}_{\mathbb{Q}, \text{lin}}$ on the set $\tilde{A} := \{x \in \mathbb{Q} \mid \text{there is } \bar{x} \in A \text{ such that } x \in \{\bar{x}[1], \bar{x}[2]\}\}$. Let $\tilde{\sigma}$ be the finite subset of the signature of $\mathcal{D}_{\mathbb{Q}, \text{lin}}$ that contains a symbol for every relation $\tilde{R}_{M, \bar{v}}$ for which there exists a symbol in σ interpreted as $R_{M, \bar{v}}$ in $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$. Since $\mathcal{D}_{\mathbb{Q}, \text{lin}}$ is convex, Theorem 9 yields an embedding \tilde{f} from the $\tilde{\sigma}$ -reduct of $\tilde{\mathfrak{A}}^2$ to the $\tilde{\sigma}$ -reduct of $\mathcal{D}_{\mathbb{Q}, \text{lin}}$. Let f be the mapping from A^2 to \mathbb{Q}^2 defined as $f(\bar{x}_1, \bar{x}_2) := (\tilde{f}(\bar{x}_1[1], \bar{x}_2[1]), \tilde{f}(\bar{x}_1[2], \bar{x}_2[2]))$. It is well defined by the definition of $\tilde{\mathfrak{A}}$. Let $(\bar{x}_1, \bar{x}_2), (\bar{y}_1, \bar{y}_2) \in A^2$ be arbitrary. Then, by the definition of \tilde{A} , $(\bar{x}_1[1], \bar{x}_2[1]), (\bar{x}_1[2], \bar{x}_2[2]), (\bar{y}_1[1], \bar{y}_2[1]), (\bar{y}_1[2], \bar{y}_2[2]) \in \tilde{A}^2$ and, for every affine transformation $\bar{x} \mapsto M\bar{x} + \bar{v}$, we have the following chain of equivalent statements.

$$\underbrace{\begin{pmatrix} \bar{x}_1 \\ \bar{y}_1 \end{pmatrix}, \begin{pmatrix} \bar{x}_2 \\ \bar{y}_2 \end{pmatrix}}_{\in R_{M, \bar{v}}} \text{ iff } \underbrace{\begin{pmatrix} \bar{x}_1[1] \\ \bar{x}_1[2] \\ \bar{y}_1[1] \\ \bar{y}_1[2] \end{pmatrix}, \begin{pmatrix} \bar{x}_2[1] \\ \bar{x}_2[2] \\ \bar{y}_2[1] \\ \bar{y}_2[2] \end{pmatrix}}_{\in \tilde{R}_{M, \bar{v}}} \text{ iff } \underbrace{\begin{pmatrix} \tilde{f}(\bar{x}_1[1], \bar{x}_2[1]) \\ \tilde{f}(\bar{x}_1[2], \bar{x}_2[2]) \\ \tilde{f}(\bar{y}_1[1], \bar{y}_2[1]) \\ \tilde{f}(\bar{y}_1[2], \bar{y}_2[2]) \end{pmatrix}}_{\in \tilde{R}_{M, \bar{v}}} \text{ iff } \underbrace{\begin{pmatrix} f(\bar{x}_1, \bar{x}_2) \\ f(\bar{y}_1, \bar{y}_2) \end{pmatrix}}_{\in R_{M, \bar{v}}}$$

The first equivalence follows from the definition of $\tilde{R}_{M, \bar{v}}$, the second from the fact that \tilde{f} is an embedding, and the third from the definition of f . These equivalences also hold for the equality predicate which can be written as $R_{E, \bar{z}}$ for E the 2×2 identity matrix and $\bar{z} = (0, 0)$. It follows that f is an embedding from the σ -reduct of \mathfrak{A}^2 to the σ -reduct of $\mathcal{D}_{\mathbb{Q}^2, \text{aff}}$. \square

5.4 Convex Structures with Forbidden Substructures

Finitely bounded structures also provide us with interesting examples of convex structures. In this setting, convexity already implies tractability.

Theorem 12 *For a finitely bounded structure \mathfrak{B} , the following are equivalent:*

1. \mathfrak{B} is convex,
2. $\text{Age}(\mathfrak{B})$ is defined by a universal Horn sentence,
3. \mathfrak{B} is p -admissible.

Proof “1 \Rightarrow 2”: Using the logical reformulation of finite boundedness in Lemma 3, we know that \mathfrak{B} is finitely bounded if its signature is finite and there is a universal first-order sentence Φ such that $\text{Age}(\mathfrak{B})$ consists precisely of the finite models of Φ . We bring Φ into prenex normal form, and then transform its quantifier-free part in conjunctive normal form. This shows that we can assume that Φ is a conjunction of implications (in the sense defined in Sect. 2). Note that a universal sentence holds in a relational structure iff it holds in each of its finite substructures. In particular, we have $\mathfrak{B} \models \Phi$. For every implication in Φ where the conclusion consists of at least two atomic formulas we apply the definition of convexity and

reduce Φ to a universal Horn sentence Φ' such that $\mathfrak{B} \models \Phi'$. This implies that Φ' holds in all elements of $\text{Age}(\mathfrak{B})$. In addition, by the construction of Φ' , the original formula Φ is a logical consequence of Φ' . Thus, if a finite τ -structure satisfies Φ' , it also satisfies Φ , and thus belongs to $\text{Age}(\mathfrak{B})$. This shows that Φ' defines $\text{Age}(\mathfrak{B})$.

“2 \Rightarrow 3”: We first show that \mathfrak{B} is convex using Theorem 9. We set $\sigma := \tau$ and select an arbitrary finite substructure \mathfrak{A} of \mathfrak{B} . Let $\forall \bar{x}. (\phi_i \Rightarrow \psi_i)$ be one of the Horn implications whose conjunction Φ over $i \in [\ell]$ defines $\text{Age}(\mathfrak{B})$. Let \bar{t} be a tuple over A^2 such that $\mathfrak{A}^2 \models \phi_i(\bar{t})$ for some $i \in [\ell]$ and let k be its arity. Then \bar{t} is of the form $((x_1, y_1), \dots, (x_k, y_k))$ such that $\mathfrak{B} \models \phi_i(x_1, \dots, x_k)$ and $\mathfrak{B} \models \phi_i(y_1, \dots, y_k)$. Since the substructure of \mathfrak{B} on $\{x_1, \dots, x_k, y_1, \dots, y_k\}$ satisfies $\forall \bar{x}. (\phi_i \Rightarrow \psi_i)$, we have $\mathfrak{B} \models \psi_i(x_1, \dots, x_k) \wedge \psi_i(y_1, \dots, y_k)$, and thus $\mathfrak{A}^2 \models \psi_i(\bar{t})$. Since the tuple \bar{t} and the index $i \in [\ell]$ were chosen arbitrarily, we know that $\mathfrak{A}^2 \models \forall \bar{x}. (\phi_i \Rightarrow \psi_i)$ for all $i \in [\ell]$. Thus, we have $\mathfrak{A}^2 \models \Phi$, which implies $\mathfrak{A}^2 \in \text{Age}(\mathfrak{B})$. We have shown that $\text{Age}(\mathfrak{B})$ is closed under taking second powers, which is a strong form of the square embedding property from Theorem 9.

Regarding tractability, note that the structure \mathfrak{B} satisfies a given Horn implication $\forall \bar{x}. (\phi \Rightarrow \psi)(\bar{x})$ iff this implication is satisfied by all elements of $\text{Age}(\mathfrak{B})$. This is the case iff the universal Horn sentence Φ that defines $\text{Age}(\mathfrak{B})$ implies the Horn implication $\forall \bar{x}. (\phi \Rightarrow \psi)(\bar{x})$. It is well known that the entailment problem is decidable in polynomial time for Horn implications [41].

“3 \Rightarrow 1”: This direction is trivial. □

This theorem yields the following two examples of p-admissible concrete domains.

Example 7 The random graph \mathfrak{G} is p-admissible since its age can be defined by the universal Horn sentence $\forall x. (E(x, x) \Rightarrow \text{ff}) \wedge \forall x, y. (E(x, y) \Rightarrow E(y, x))$.

The structure $(\mathbb{Q}; >)$ is not convex, as otherwise Theorem 9 would imply that it contains incomparable elements since the square of this linear order is not linear. In the universal sentence defining $\text{Age}(\mathbb{Q}; >)$ (see Lemma 3), the totality axiom $\forall x, y. (x < y \vee x = y \vee x > y)$ is the culprit since it is not Horn. If we remove this axiom, we obtain the theory of strict partial orders.

It is well known that there exists a unique countable homogeneous strict partial order \mathfrak{P} [68], whose age is defined by the universal Horn sentence $\forall x, y, z. (x < y \wedge y < z \Rightarrow x < z) \wedge \forall x. (x < x \Rightarrow \text{ff})$. Thus, \mathfrak{P} is finitely bounded and convex. Using \mathfrak{P} as a concrete domain means that the feature values satisfy the theory of strict partial orders, but not more. One can, for instance, use this concrete domain to model preferences of people; e.g., *Italian* $\sqcap \exists \text{pizzapref}, \text{pastapref}. (x_1 > x_2)$ is a concept describing Italians that like pizza more than pasta. Using \mathfrak{P} here means that preferences may be incomparable. As we have seen above, adding totality would break convexity and thus p-admissibility.

By combining Corollary 3 with Theorem 12, we can obtain non-trivial p-admissible concrete domains \mathfrak{D} for which subsumption in $\mathcal{ALC}(\mathfrak{D})$ is decidable. Note that, according to Proposition 3, such a non-trivial structure \mathfrak{D} cannot be ω -admissible, but it is the reduct of the ω -admissible structure $\mathfrak{D}^{\leq d}$.

Corollary 9 *Let \mathfrak{D} be a finitely bounded convex structure that is a reduct of a finitely bounded homogeneous structure. Then subsumption w.r.t. TBoxes is tractable in $\mathcal{EL}[\mathfrak{D}]$ and decidable in $\mathcal{ALC}(\mathfrak{D})$.*

Examples of infinitely many non-trivial structures satisfying the condition stated in this corollary will be given in the next subsection.

5.5 Convex Structures with Forbidden Homomorphic Images

Beside finitely bounded structures, the literature also considers structures whose age can be described by a finite set of forbidden homomorphic images [38, 53]. For a class \mathcal{F} of τ -structures, $\text{Forb}_h(\mathcal{F})$ stands for the class of all finite τ -structures that do not contain a homomorphic image of any member of \mathcal{F} . A structure is *connected* if its so-called Gaifman graph is connected. The *Gaifman graph* of a structure \mathfrak{A} is the undirected graph (A, E) such that there is an edge in E between two elements $a, a' \in A$ iff they occur together in a tuple from a relation of \mathfrak{A} .

Theorem 13 (Cherlin, Shelah, and Shi [38]) *Let \mathcal{F} be a finite family of connected relational structures with a finite signature τ . Then there exists an ω -categorical τ -structure $\text{CSS}(\mathcal{F})$ that is a reduct of a finitely bounded homogeneous structure and $\text{Age}(\text{CSS}(\mathcal{F})) = \text{Forb}_h(\mathcal{F})$.*

We can show that the structures of the form $\text{CSS}(\mathcal{F})$ provided by this theorem are always p-admissible.

Proposition 12 *Let \mathcal{F} be a finite family of connected relational structures with a finite signature τ . Then $\text{CSS}(\mathcal{F})$ is p-admissible.*

Proof Let $\mathfrak{B} := \text{CSS}(\mathcal{F})$. By Theorem 13, we have $\mathfrak{A} \in \text{Age}(\mathfrak{B})$ iff \mathfrak{A} does not contain a homomorphic image of any $\mathfrak{F} \in \mathcal{F}$ as a substructure. If we can show $\text{Age}(\mathfrak{B}^2) \subseteq \text{Age}(\mathfrak{B})$, then it follows from Theorem 9 that $\text{CSS}(\mathcal{F})$ is convex. Suppose that there exists $\mathfrak{C} \in \text{Age}(\mathfrak{B}^2)$ such that $\mathfrak{C} \notin \text{Age}(\mathfrak{B})$. Then there exists $\mathfrak{F} \in \mathcal{F}$ such that $\mathfrak{F} \rightarrow \mathfrak{C}$. Since the projection to a single component is a homomorphism, this shows that there is a homomorphism $\mathfrak{F} \rightarrow \mathfrak{B}$. But then the image of \mathfrak{F} under this homomorphism is a finite substructure of \mathfrak{B} that does not belong to $\text{Forb}_h(\mathcal{F})$, which contradicts the fact that $\text{Age}(\mathfrak{B}) = \text{Forb}_h(\mathcal{F})$. Thus indeed $\text{Age}(\mathfrak{B}^2) \subseteq \text{Age}(\mathfrak{B})$ and $\text{CSS}(\mathcal{F})$ is convex.

Since there are, up to isomorphisms, only finitely many homomorphic images of each $\mathfrak{F} \in \mathcal{F}$ in \mathfrak{B} , there exists a finite set \mathcal{F}' of finite structures such that $\text{Age}(\mathfrak{B}) = \text{Forb}_e(\mathcal{F}')$, which means that \mathfrak{B} is finitely bounded. Since $\text{CSS}(\mathcal{F})$ is convex, its p-admissibility follows from Theorem 12. \square

Proposition 12 together with the next example provides us with infinitely many countable p-admissible concrete domains satisfying the preconditions of Corollary 9, which all yield a different extension of \mathcal{EL} . The usefulness of these concrete domains for defining interesting concepts is, however, as yet unclear.

Example 8 A directed graph is a *tournament* if every two distinct vertices in it are connected by exactly one directed edge. Henson [47] proved that there are uncountably many homogeneous directed graphs by showing that, for any (not necessarily finite) set \mathcal{N} of finite tournaments (plus the loop and the 2-cycle) such that no member of \mathcal{N} is embeddable into any other member of \mathcal{N} , $\text{Forb}_e(\mathcal{N})$ is an amalgamation class whose Fraïssé limit is a homogeneous directed graph. Furthermore, the Fraïssé limits for two distinct sets of such tournaments are distinct as well. In the literature, such directed graphs are often called *Henson digraphs* [66].

An important observation about Henson digraphs is that $\text{Forb}_e(\mathcal{N}) = \text{Forb}_h(\mathcal{N})$ holds for any set \mathcal{N} of finite tournaments plus the loop and the 2-cycle. The inclusion $\text{Forb}_h(\mathcal{N}) \subseteq \text{Forb}_e(\mathcal{N})$ holds since every embedding is a homomorphism. To show the other inclusion, suppose that $\mathfrak{A} \in \text{Forb}_e(\mathcal{N})$. The loop clearly does not homomorphically map to \mathfrak{A} because

every homomorphism from the loop to \mathfrak{A} is an embedding. Since the loop does not homomorphically map to \mathfrak{A} , every homomorphism from the 2-cycle to \mathfrak{A} is an embedding. Thus, the 2-cycle does not homomorphically map to \mathfrak{A} . Since the loop and the 2-cycle do not homomorphically map to \mathfrak{A} , every homomorphism from a tournament to \mathfrak{A} is an embedding. Thus, \mathfrak{A} does not admit any homomorphic image of a structure from \mathcal{N} . We conclude that $\text{Forb}_e(\mathcal{N}) \subseteq \text{Forb}_h(\mathcal{N})$.

For every selection \mathcal{N} of finitely many tournaments that do not embed into each other, the set \mathcal{N} consists of connected structures since tournaments as well as the loop and the 2-cycle are connected. Moreover, if $\mathcal{N}_1, \mathcal{N}_2$ are two distinct such sets, then $\text{Forb}_h(\mathcal{N}_1) \neq \text{Forb}_h(\mathcal{N}_2)$ [66]. Since there are infinitely many such families \mathcal{N} , Theorem 13 and Proposition 12 yield infinitely many non-isomorphic p-admissible and finitely bounded concrete domains that have different ages. Consequently, the ages of these structures are defined by universal Horn sentences that are not equivalent. This implies that, in the extension of \mathcal{EL} with these concrete domains, different subsumptions hold.

To make this more precise, let $\mathfrak{D}_1 := \text{CSS}(\mathcal{N}_1)$ and $\mathfrak{D}_2 := \text{CSS}(\mathcal{N}_2)$. Assume that $\forall \bar{x}. (\phi \Rightarrow \psi)$ is a Horn implication that is satisfied by all elements of $\text{Forb}_h(\mathcal{N}_1) = \text{Age}(\mathfrak{D}_1)$, but for which there is an element \mathfrak{G} of $\text{Forb}_h(\mathcal{N}_2) = \text{Age}(\mathfrak{D}_2)$ that does not satisfy it. We can easily turn the conjunction of atomic formulas ϕ and the atomic formulas ψ into concepts C_ϕ and C_ψ of the DLs $\mathcal{EL}[\mathfrak{D}_1]$ and $\mathcal{EL}[\mathfrak{D}_2]$ by viewing the variables in \bar{x} as features and replacing the conjunct operators \wedge in ϕ by DL conjunction \sqcap . If we additionally ensure that all these features are defined (using GCIs $\top \sqsubseteq \exists x, x.(x_1 = x_2)$ for all x occurring in \bar{x}), then C_ϕ is subsumed by C_ψ w.r.t. these GCIs in $\mathcal{EL}[\mathfrak{D}_1]$, but not in $\mathcal{EL}[\mathfrak{D}_2]$ since one can use $\mathfrak{G} \in \text{Age}(\mathfrak{D}_2)$ to construct a counterexample to the subsumption.

A more general class of p-admissible structures can be obtained from connected MMSNP (for monotone monadic strict NP) sentences. Recall the notion of a canonical database from Definition 6.

Definition 7 A *connected (equality-free) MMSNP sentence Φ* over a finite relational signature τ is of the form $\Phi = \exists P_1, \dots, P_n. \forall \bar{x}. \bigwedge_i \neg(\alpha_i \wedge \beta_i)$ where

- P_1, \dots, P_n are unary relation symbols not in τ ,
- each α_i is a conjunction of atomic formulas of the form $R(\bar{x})$ for $R \in \tau$ with free variables \bar{x}_i such that $\text{DB}(\exists \bar{x}_i. \alpha_i)$ is connected,
- each β_i is a conjunction of atomic formulas of the form $P_i(x)$ for $i \in [n]$ and their negations.

Note that, for every family \mathfrak{F} as in Theorem 13, the class $\text{Age}(\text{CSS}(\mathcal{F}))$ consists of all finite models of a particular MMSNP sentence of the form $\forall \bar{x}. \bigwedge_i \neg\alpha_i$ where each α_i encodes a single structure $\mathfrak{F} \in \mathcal{F}$ up to homomorphic equivalence. The following result is thus as a generalization of Theorem 13 to more complicated forbidden patterns involving existentially quantified unary predicates.

Theorem 14 (Theorem 7 in [17]) *For every connected MMSNP sentence Φ over a finite signature τ , there exists an ω -categorical τ -structure \mathfrak{B}_Φ that is a reduct of a finitely bounded homogeneous structure and such that $\text{Age}(\mathfrak{B}_\Phi)$ consists of all finite models of Φ .*

Like Theorem 13, this theorem can be used to produce p-admissible concrete domains. However, in contrast to the setting considered in Theorem 13, connected MMSNP is known to exhibit a complexity dichotomy between P and NP-complete [26]. The following proposition shows that, already within the class of *reducts* of finitely bounded homogeneous structures,

p-admissibility does not only depend on convexity, in contrast to what holds for finitely bounded structures (see Theorem 12).

Proposition 13 *Let Φ be a connected MMSNPN sentence over a finite signature τ . Then \mathfrak{B}_Φ is always convex, and it is p-admissible iff satisfiability of Φ in finite τ -structures can be tested in polynomial time.*

Proof We show convexity using Theorem 9. Let \mathfrak{A} be a finite substructure of \mathfrak{B}_Φ . Then $\mathfrak{A} \models \Phi$ and this is witnessed by sets $P_1, \dots, P_n \subseteq A$. Assume that $\mathfrak{A}^2 \not\models \Phi$. For every $i \in [n]$, we set $P'_i := P_i \times A$. Since $\mathfrak{A}^2 \not\models \Phi$, there exists a tuple \bar{s} over A^2 such that $(\mathfrak{A}^2, P'_1, \dots, P'_n) \models (\alpha_i \wedge \beta_i)(\bar{s})$ for some i . Let \bar{r} be the tuple over A obtained from \bar{s} by taking the projection of each entry in \bar{s} to the first coordinate. By the definition of the product of structures and of the sets P'_i , we obtain $(\mathfrak{A}, P_1, \dots, P_n) \models (\alpha_i \wedge \beta_i)(\bar{r})$, which contradicts our assumption that $\mathfrak{A} \models \Phi$ is witnessed by P_1, \dots, P_n . Thus $\mathfrak{A}^2 \models \Phi$, which shows $\mathfrak{A}^2 \in \text{Age}(\mathfrak{B}_\Phi)$. An application of Theorem 9 thus yields convexity of \mathfrak{B}_Φ .

It remains to determine in which cases we can test validity of Horn implications in \mathfrak{B}_Φ in polynomial time. The proof of Theorem 7 in [17] yields $\text{CSP}(\mathfrak{B}_\Phi) = \text{Age}(\mathfrak{B}_\Phi)$. It can be shown as in the proof of Proposition 11 that testing satisfiability of Horn implications in \mathfrak{B}_Φ reduces in polynomial time to $\text{CSP}(\mathfrak{B}_\Phi)$, which amounts to testing satisfiability of Φ by Theorem 14 because $\text{CSP}(\mathfrak{B}_\Phi) = \text{Age}(\mathfrak{B}_\Phi)$. Hence, testing satisfiability of Horn implications in \mathfrak{B}_Φ can be done in polynomial time iff testing satisfiability of Φ in finite structures can be done in polynomial time. \square

Example 9 Consider the following two connected MMSNPN sentences:

$$\begin{aligned} \Phi_1 &:= \exists P. \forall x, y. \neg(E(x, y) \wedge P(x) \wedge P(y)) \wedge \neg(E(x, y) \wedge \neg P(x) \wedge \neg P(y)) \\ \Phi_2 &:= \exists P. \forall x, y, z. \neg(E(x, y) \wedge E(y, z) \wedge E(z, x) \wedge P(x) \wedge P(y) \wedge P(z)) \\ &\quad \wedge \neg(E(x, y) \wedge E(y, z) \wedge E(z, x) \wedge \neg P(x) \wedge \neg P(y) \wedge \neg P(z)) \end{aligned}$$

It is easy to see that testing satisfiability of Φ_1 in finite structures corresponds to solving the well-known 2-colorability problem, which is known to be tractable. Thus, the structure \mathfrak{B}_{Φ_1} is a p-admissible concrete domain. Satisfiability of Φ_2 in finite structures corresponds to the problem *No-Mono-Tri* (for “no mono-chromatic triangle”), which is known to be NP-complete [24]. Thus, the structure \mathfrak{B}_{Φ_2} is convex, but it is not p-admissible (unless P=NP). More examples of connected MMSNPN sentences can be found in [24].

5.6 (Non-)closure Properties of Finitely Bounded Convex Structures

In contrast to homogeneity, convexity is quite fragile. For example, it is in general not preserved under adding predicates of the form $=_c$, even under the assumption of finite boundedness.

Proposition 14 *Finitely bounded convex structures are not closed under adding singleton predicates $=_c$.*

Proof The unique countable homogeneous strict partial order \mathfrak{P} was introduced and shown to be convex in Example 7. Consider the extension \mathfrak{P}_c of \mathfrak{P} by a smallest element $c \notin P$, i.e., $c < p$ for every $p \in P$. It is easy to see that $\text{Age}(\mathfrak{P}_c) = \text{Age}(\mathfrak{P})$, which means that \mathfrak{P}_c is still finitely bounded and convex. Now consider its expansion \mathfrak{P}'_c by the unary relation $=_c$, which is interpreted as $\{c\}$. The structure \mathfrak{P}'_c is not convex since $\forall x, y. (x = x \wedge$

$=_c(y)) \Rightarrow (y < x \vee =_c(x))$ holds in it, but neither $\forall x, y. (x = x \wedge =_c(y)) \Rightarrow y < x$ nor $\forall x, y. (x = x \wedge =_c(y)) \Rightarrow =_c(x)$. □

When it comes to expansions by first-order definable relations, we clearly run into problems if we allow definitions containing disjunctions of atomic formulas. However, except for very specific situations as in Lemma 4, convexity is not even preserved under taking expansions by pp definable relations.

Proposition 15 *Finitely bounded convex structures are not closed under taking expansions by pp definable relations.*

Proof As shown in [46], there exists a unique countable homogeneous undirected graph \mathfrak{H} that embeds precisely those finite undirected graphs not containing the complete graph on three vertices \mathfrak{K}_3 as an induced subgraph. By Lemma 3, \mathfrak{H} is finitely bounded because $\text{Age}(\mathfrak{H})$ is defined by the following universal Horn sentence:

$$\forall x, y, z. (E(x, y) \wedge E(y, z) \wedge E(z, x) \Rightarrow \text{ff})$$

$$\wedge \forall x, y. (E(x, y) \Rightarrow E(y, x)) \wedge \forall x. (E(x, x) \Rightarrow \text{ff}).$$

By Theorem 12, \mathfrak{H} is also convex. However, the expansion (\mathfrak{H}, \neq) is not convex since $(\mathfrak{H}, \neq) \models \forall x_1, x_2, x_3, x_4. (x_1 \neq x_2 \wedge x_3 \neq x_4) \Rightarrow (x_1 \neq x_3 \vee x_1 \neq x_4)$, but both $x_1 = x_3 \neq x_4$ and $x_1 = x_4 \neq x_3$ is possible in (\mathfrak{H}, \neq) . We claim that \neq can be pp-defined in \mathfrak{H} by the formula

$$\phi(x_1, x_4) = \exists x_2, x_3. (E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_3, x_4)).$$

First, suppose that $\mathfrak{H} \models \phi(h_1, h_4)$ for some $h_1, h_4 \in H$. Then clearly $h_1 \neq h_4$ as otherwise \mathfrak{H} would embed \mathfrak{K}_3 . Second, let h_1, h_4 be arbitrary distinct elements of H . Consider the undirected path \mathfrak{P}_4 with four vertices v_1, v_2, v_3, v_4 . Since \mathfrak{P}_4 does not embed \mathfrak{K}_3 , there exists an embedding $e: \mathfrak{P}_4 \hookrightarrow \mathfrak{H}$. If there is an edge between h_1 and h_4 , then we can take $x_2 = h_4$ and $x_3 = h_1$ to show that $\mathfrak{H} \models \phi(h_1, h_4)$. Otherwise, the substructures of \mathfrak{H} on $\{h_1, h_4\}$ and on $\{e(v_1), e(v_4)\}$ are isomorphic. Since \mathfrak{H} is homogeneous, there exists $\alpha \in \text{Aut}(\mathfrak{H})$ which sends $e(v_1)$ to h_1 and $e(v_4)$ to h_4 . Since $\alpha \circ e$ is a homomorphism, it follows that $(x_1, \dots, x_4) := (\alpha \circ e(v_1), \dots, \alpha \circ e(v_4))$ satisfies the quantifier-free part of ϕ in \mathfrak{H} , and thus $\mathfrak{H} \models \phi(h_1, h_4)$ also in this case. □

Also, convexity is not preserved under taking disjoint unions.

Proposition 16 *Finitely bounded convex structures are not closed under disjoint union.*

Proof Consider a signature with a single unary predicate symbol and a structure $(S; R)$ where S is countably infinite and R is interpreted as the whole domain S . This structure is finitely bounded and convex by Lemma 3 and Theorem 12 since its age is defined by the universal Horn sentence $\forall x. R(x)$. If we build the union of $(S; R)$ with an isomorphic copy of itself over a domain disjoint with S , then we obtain a structure isomorphic to the structure $\mathfrak{N} = (\mathbb{N}; E, O)$, of which we have seen in Sect. 3.2.2 that it is not convex. □

However, convexity is preserved under taking the full product. This is an easy consequence of Theorem 9 combined with the fact that the mapping $((x_1, x_2), (y_1, y_2)) \mapsto ((x_1, y_1), (x_2, y_2))$ is an isomorphism between $\mathfrak{D}_1^2 \boxtimes \mathfrak{D}_2^2$ and $(\mathfrak{D}_1 \boxtimes \mathfrak{D}_2)^2$.

Proposition 17 *Convex structures are closed under taking the full product.*

6 Toward User-Definable Concrete Domains

DL systems that can handle concrete domains allow their users to employ a fixed set of predicates of one or more fixed concrete domains when modeling concepts. They do not provide their users with means for defining new predicates, let alone new concrete domains. Our results in Sect. 4 alleviate the first restriction since Corollary 3 allows the use of first-order definable predicates and Corollary 4 of predicates definable by existential positive formulas. To overcome the second restriction, one would need to provide the user with (i) a mechanism for defining a concrete domain; (ii) an algorithm that checks whether this concrete domain is ω - or p-admissible; and (iii) an automated way of generating the required reasoning procedures for this concrete domain.

For the case of ω -admissible concrete domains, one might think that Theorem 6 provides us with these ingredients. To define a concrete domain satisfying the preconditions of this theorem, one could start with selecting a finite set \mathcal{N} of bounds (or equivalently, by Lemma 3, a universal sentence). The question is then whether \mathcal{N} really induces a finitely bounded structure. The bad news is that this question is in general undecidable.

Proposition 18 *Let τ be a finite relational signature containing at least one binary symbol. The question whether, for a given finite set \mathcal{N} of finite τ -structures, there is a τ -structure \mathfrak{D} such that $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N})$ is in general undecidable.*

Proof It is shown in [33] that the joint embedding property (JEP) is undecidable for classes of undirected graphs definable by finitely many bounds. In addition, it is known that a class of structures definable by finitely many bounds has JEP iff this class is the age of some countable structure (Theorem 6.1.1 in [50]). □

However, to apply Theorem 6, we need AP rather than just JEP. In contrast to JEP, the amalgamation property (AP) is decidable for classes over finite binary signatures defined by finitely many forbidden finite substructures [57].

Theorem 15 *The question whether, for a given finite set \mathcal{N} of finite τ -structures over a finite relational signature τ consisting of binary symbols, there exists a homogeneous τ -structure \mathfrak{D} such that $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N})$ is decidable in Π_2^P .*

Proof According to [57], it is decidable whether $\text{Forb}_e(\mathcal{N})$ has AP. By Theorem 5, this is the case iff there is a homogeneous structure \mathfrak{D} such that $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N})$. A decision procedure for this problem was also described recently in [29]. As mentioned in the proof of Theorem 4 in [29], it is enough to show AP restricted to triples $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$ such that

$$e_i : \mathfrak{A} \hookrightarrow \mathfrak{B}_i \text{ is the identity map and } B_i \setminus A = \{b_i\} \text{ for } i = 1, 2. \tag{\#}$$

Let R_1, \dots, R_n be an enumeration of τ . Suppose that there exists $S \subseteq [n]$ such that the formula

$$\phi_S(y_1, y_2) := \bigwedge_{i \in S} R_i(y_1, y_2) \wedge \bigwedge_{i \in [n] \setminus S} \neg R_i(y_1, y_2) \wedge \bigwedge_{i \in [n]} \neg R_i(y_2, y_1)$$

is not satisfiable in any $\mathfrak{F} \in \mathcal{N}$. Then $\text{Forb}_e(\mathcal{N})$ has AP because, for every triple $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$, we can choose \mathfrak{C} with domain $B_1 \cup B_2$ and relations

$$R_i^{\mathfrak{C}} := \begin{cases} R_i^{\mathfrak{B}_1} \cup R_i^{\mathfrak{B}_2} \cup \{(b_1, b_2)\} & \text{if } i \in S \\ R_i^{\mathfrak{B}_1} \cup R_i^{\mathfrak{B}_2} & \text{if } i \in [n] \setminus S. \end{cases}$$

It is easy to see that $\mathfrak{C} \in \text{Forb}_e(\mathcal{N})$: since $\mathfrak{F} \in \mathcal{N}$ cannot embed into \mathfrak{B}_1 or \mathfrak{B}_2 , the image of an embedding of \mathfrak{F} into \mathfrak{C} would need to contain b_1 and b_2 , but then the formula $\phi_S(y_1, y_2)$ would be satisfiable in \mathfrak{F} .

Now suppose that $\text{Forb}_e(\mathcal{N})$ does not have AP. We define the size of \mathcal{N} as the sum of the sizes of all structures in \mathcal{N} , where the size of a structure is the sum of the cardinalities of the domain and all relations. By the argument in the previous paragraph, for every $S \subseteq [n]$, the formula $\phi_S(y_1, y_2)$ must be satisfiable in some $\mathfrak{F} \in \mathcal{N}$. Consequently, this structure contains a tuple (a_1, a_2) such that $\phi_S(a_1, a_2)$ holds, but $\phi_{S'}(a_1, a_2)$ does not hold for any $S' \neq S$. This shows that, overall structures in \mathcal{N} , there are at least $2^{|\tau|}$ tuples. Since all of them except for one belongs to at least on relation and the cardinality of the structures in \mathcal{N} is at least 1, this shows that the size of \mathcal{N} is at least $2^{|\tau|}$.

To prove that the original problem is in Π_2^P , it is sufficient to show that the complement can be decided by an NP procedure that uses a coNP oracle. Given a finite set of bounds \mathcal{N} , we guess a triple $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$ satisfying (#) and check whether this triple witnesses that $\text{Forb}_e(\mathcal{N})$ does not have AP. According to the proof of Theorem 4 in [29], the size of a smallest counterexample to AP for $\text{Forb}_e(\mathcal{N})$ is bounded by a polynomial in $m \cdot \ell$ where $m := \max_{\mathfrak{F} \in \mathcal{N}} |F|$ and $\ell := 2^{|\tau|}$. Thus, by what we have shown above for the size of \mathcal{N} , we may assume that the size of $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2$ is polynomial in the size of the input \mathcal{N} , which shows that this triple can be guessed within NP. To verify that it is a counterexample to AP, we need to check that

1. $\mathfrak{A}, \mathfrak{B}_1, \mathfrak{B}_2 \in \text{Forb}_e(\mathcal{N})$, and
2. there exists no $\mathfrak{C} \in \text{Forb}_e(\mathcal{N})$ with embeddings $f_i : \mathfrak{B}_i \hookrightarrow \mathfrak{C}, i \in [2]$, such that $f_1|_A = f_2|_A$.

The test in item 1 can be performed by a coNP oracle. In fact, to check whether a finite structure does *not* belong to $\text{Forb}_e(\mathcal{N})$, it is sufficient to guess an embedding from an element of \mathcal{N} into this structure. Clearly, this can be done by an NP procedure.

For item 2, first note that it is clearly sufficient to consider structures \mathfrak{C} such that $C = B_1 \cup B_2$ and where the embeddings f_i are the identity.⁶ There are only polynomially many structures of this kind. In fact, to determine such a structure, we need to decide for the tuples (b_1, b_2) and (b_2, b_1) to which of the binary relations in τ they belong. There are $2^{|\tau|}$ possibilities for each tuple, and we already know that $2^{|\tau|}$ is polynomial in the size of the input. The test whether $\mathfrak{C} \in \text{Forb}_e(\mathcal{N})$ can again be solved by a coNP oracle.

Thus, we have shown that the complement of the problem of deciding AP can be solved by an NP procedure that uses a coNP oracle, which finishes the proof of the lemma. \square

Assume that, in the binary case, the test whether $\text{Forb}_e(\mathcal{N})$ has AP was successful, and let \mathfrak{D} be the corresponding homogeneous structure. Using the results from Sect. 4, we can then transform \mathfrak{D} into an ω -admissible concrete domain through a decomposition of its relations into orbits under $\text{Aut}(\mathfrak{D})$. The required decision procedure for the CSP can then be obtained from the proof of Proposition 6. Thus, for the case of binary signatures, Theorem 6 together with related results in Sect. 4 provides us with the necessary ingredients for enabling user-definable ω -admissible concrete domains. This automated approach can be used to identify RCC8 and Allen as ω -admissible concrete domains because they are both finitely bounded (see Example 2). It is an open question whether the decidability result for AP in [57] can be extended to finite signatures containing non-binary symbols. In the setting relevant for p-admissibility, we can show that the analogous problem is undecidable already for signatures

⁶ The case where $f_1(b_1) = f_2(b_2)$ can only yield a counterexample if \mathfrak{B}_1 and \mathfrak{B}_2 are equal up to renaming of b_1 with b_2 , which can easily be checked.

containing at most binary symbols. This is an easy consequence of the following theorem, whose proof can be found in [30].

Theorem 16 *The question whether the class of all finite models of a given universal Horn sentence in a finite signature τ has JEP is undecidable even if τ is limited to binary symbols.*

This theorem yields the following undecidability result for p -admissibility of finitely bounded structures.

Corollary 10 *The question whether, for a given finite set \mathcal{N} of finite τ -structures over a finite relational signature τ , there exists a p -admissible τ -structure \mathfrak{D} such that $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N})$ is undecidable even if τ is limited to binary symbols.*

Proof By Lemma 3, for every universal Horn sentence Φ in the signature τ , there exists a finite set \mathcal{N}_Φ of finite τ -structures such that $\text{Forb}_e(\mathcal{N}_\Phi)$ is the set of all finite models of Φ . By Corollary 6, there exists a structure \mathfrak{D} with $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N}_\Phi)$ iff $\text{Forb}_e(\mathcal{N}_\Phi)$ has JEP. Moreover, whenever there exists a structure \mathfrak{D} with $\text{Age}(\mathfrak{D}) = \text{Forb}_e(\mathcal{N}_\Phi)$, then \mathfrak{D} is p -admissible by Theorem 12. Thus undecidability follows directly from Theorem 16. \square

There is, however, a simple syntactic restriction that guarantees the existence of a p -admissible structure whose age is described by a given universal Horn sentence.

Proposition 19 *Let Φ be a satisfiable universal Horn sentence over a finite signature such that $\phi(\bar{x})$ is equality-free and $\text{DB}(\exists \bar{x}.\phi(\bar{x}))$ is connected for every conjunct $\forall \bar{x}.\phi(\bar{x}) \Rightarrow \psi(\bar{x})$ in Φ , and let \mathcal{K} be the class of all finite models of Φ . Then there exists a finitely bounded p -admissible structure \mathfrak{D} such that $\mathcal{K} = \text{Age}(\mathfrak{D})$.*

Proof It is easy to see that \mathcal{K} has JEP because, due to the connectedness condition, it is preserved under taking disjoint unions. Then the existence of an appropriate structure \mathfrak{D} follows from Corollary 6 and Theorem 12. \square

The precondition of Proposition 19 is, for instance, satisfied for the class of all finite strict partial orders (see Example 7).

7 Conclusion

The notions of ω -admissibility and p -admissibility were respectively introduced in [65] and [11] to obtain decidable and tractable extensions of DLs by concrete domains. In each of these papers, two examples of concrete domains satisfying the respective restrictions were given. To the best of our knowledge, no other ω -admissible or p -admissible concrete domains had been exhibited in the literature before our investigations in [6] and [8]. This appears to be mainly due to the fact that it is not easy to show the conditions required by ω -admissibility or p -admissibility “by hand”. The main contribution of this work is that it provides us with useful algebraic tools for proving these conditions.

We have shown that ω -admissibility is closely related to well-known notions from model theory such as homogeneity and finite boundedness. Given the fact that a large number of homogeneous structures are known from the literature [66] and that homogeneous and finitely bounded structures play an important rôle in the CSP community, we believe that our work will turn out to be useful for locating new ω -admissible concrete domains.

This is not the first model-theoretic description of a sufficient condition for decidability of reasoning in DLs with concrete domains in the presence of TBoxes. The *existence of homomorphism is definable (EHD)* property was used in [35] to obtain decidability results for DLs with concrete domains. However, the way the concrete domain is integrated into the DL in [35] is different from the classical one employed by us and used in all other papers on DLs with concrete domains. In [35], constraints are always placed along a linear path stemming from a single individual, which is rather similar to the use of constraints in temporal logics [36, 40]. In contrast, in the classical setting of DLs with concrete domains, one can compare feature values of siblings of an individual. Compared to homogeneity and finite boundedness, the EHD property is not as well investigated. To the best of our knowledge, the only article besides [35] where concrete domains satisfying the EHD property are studied in the context of \mathcal{ALC} with GCIs is [60, 61].⁷ There, the authors consider specific concrete domains based on integers equipped with a linear order and provide an exponential upper bound for reasoning using an automata-theoretic algorithm. Interestingly, their upper bound holds not only for constraints along paths, but also for the traditional integration of concrete domain into DLs. The results in [35, 60, 61] demonstrate that ω -admissibility is not necessary for decidable reasoning. However, all known non- ω -admissible concrete domains with the EHD property are based on “discrete” versions of ω -admissible concrete domains, which are patchworks but lack homomorphism ω -compactness, e.g., $(\mathbb{Z}; <, =, >)$ or $\text{Allen}_{\mathbb{Z}}$. Motivated by this observation, we identify homomorphism ω -compactness in its current form as the most obvious “flaw” of the ω -admissibility condition, in the sense that it may be too strong. In fact, the correctness of the tableau algorithm from [65] only requires very specific infinite structures to have a homomorphism to the concrete domain, e.g., their treewidth is bounded by a computable function in the size of the input concept and TBox. But even if we restrict the inputs to homomorphism ω -compactness appropriately, the tableau algorithm from [65] is not correct for “discrete” versions of ω -admissible concrete domains, as illustrated by Example 1. We conclude that, although a modified version of ω -admissibility could in theory be necessary for decidable reasoning in \mathcal{ALC} with concrete domains in the presence of GCIs, showing this might require a non-trivial combination of the methods in [35, 60, 61, 65]. As a closing remark on ω -admissibility, we mention that concept satisfiability w.r.t. GCIs in \mathcal{ALC} with concrete domains is not the only known decision problem whose decidability was shown under the assumption that a given parameterizing class of structures is closed under taking amalgams. For instance, an amalgamation-based approach was used in [31] to show decidability of various decision problems for so-called database-driven systems.

For p-admissibility, we have developed a very useful algebraic tool for showing convexity: the square embedding property. We have shown that this tool can indeed be used to exhibit new p-admissible concrete domains, such as countably infinite vector spaces over finite field, the countable homogeneous partial order, and numerical concrete domains over \mathbb{R} and \mathbb{Q} whose relations are defined by linear equations. The usefulness of these numerical concrete domains for defining concepts should be evident. For the other two we have indicated their potential usefulness by small examples.

We have also shown that, for finitely bounded structures, convexity is equivalent to p-admissibility, and that this corresponds to the finite substructures being definable by a universal Horn sentence. Interestingly, this provides us with infinitely many examples of countable p-admissible concrete domains, which all yield a different extension of \mathcal{EL} : the Henson digraphs. From a theoretical point of view, this is quite a feat, given that before

⁷ though EHD is not used in the proofs in [60, 61].

only two p-admissible concrete domains were known. It is less clear whether these concrete domains are useful for defining concepts.

Finitely bounded structures also provide us with examples of structures \mathfrak{D} that can be used both in the context of \mathcal{EL} and \mathcal{ALC} , in the sense that subsumption is tractable in $\mathcal{EL}[\mathfrak{D}]$ and decidable in $\mathcal{ALC}(\mathfrak{D})$. Finally, we have shown that, when embedding p-admissible concrete domains into \mathcal{EL} , the restriction to paths of length one in concrete domain restrictions (indicated by the square brackets) is needed since there is a p-admissible concrete domain \mathfrak{D} such that subsumption in $\mathcal{EL}(\mathfrak{D})$ is undecidable.

Acknowledgements The authors would like to thank Manuel Bodirsky and Carsten Lutz for helpful discussions. Franz Baader was partially supported by the AI competence center ScaDS.AI Dresden/Leipzig and the Deutsche Forschungsgemeinschaft (DFG), Grant 389792660, within TRR 248. Jakub Rydval was supported by the DFG GRK 1763 (QuantLA).

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

A Proof of Corollary 1

In the proof of Proposition 2, we use concepts of the form $\exists f, gf.R_{M,\bar{v}}(x_1, x_2)$ where $\bar{x} \mapsto M\bar{x} + \bar{v}$ is an arbitrary affine transformation from \mathbb{Q}^2 to \mathbb{Q}^2 . We show that every such concept can be expressed as a conjunction of concepts built using only those affine transformations $\bar{x} \mapsto M\bar{x} + \bar{v}$ where $M \in \{-1, 0, 1\}^{2 \times 2}$ and $\bar{v} \in \{-1, 0, 1\}^2$. This gives us a conservative extension of the concept and the TBox used in the proof of Proposition 2. Thus the statement then follows from Proposition 2.

Note that we can clearly express $\exists f, gf.R_{M,\bar{v}}(x_1, x_2)$ as

$$\exists f, f'.R_{M,\bar{0}}(x_1, x_2) \sqcap \exists f', f''.R_{E_2,\bar{v}}(x_1, x_2) \sqcap \exists f'', gf.(x_1 = x_2)$$

where E_2 is the 2×2 unit matrix and f', f'' are fresh features. Thus, we only really need to express concepts of the form $\exists f, g.R_{E_2,\bar{v}}(x_1, x_2)$ and $\exists f, g.R_{M,\bar{0}}(x_1, x_2)$ where M, \bar{v} are elements of our selected finite set of matrices and vectors.

Consider an arbitrary matrix $M = (m_{i,j})_{i,j \in [2]}$ where, w.l.o.g., $m_{i,j} = p_{i,j}/q_{i,j}$ for an integer $p_{i,j}$ and a positive integer $q_{i,j}$. Then $(\bar{x}, \bar{y}) \in R_{M,\bar{0}}$ iff

$$q_{i,2}p_{i,1}\bar{x}[1] + q_{i,1}p_{i,2}\bar{x}[2] = q_{i,1}q_{i,2}\bar{y}[i] \text{ for } i \in \{1, 2\}. \quad (\dagger)$$

We claim that, for every $n \in \mathbb{Z}$, there exists a concept constructed using our selected set of matrices and vectors that expresses the concept $\exists f, g.R_{A_n, \bar{0}}(x_1, x_2)$ where the affine transformation $\bar{x} \mapsto A_n \bar{x}$ multiplies the first component by n and the second component

by 0, i.e.,

$$A_{n,0} = \begin{pmatrix} n & 0 \\ 0 & 0 \end{pmatrix}.$$

W.l.o.g., $n > 1$, the case $n < 0$ is similar and the case $n \in \{0, 1\}$ is trivial. For every $i \in [n]$, we introduce a fresh feature f_i . Then $\exists f, g.R_{A_{n,0},\bar{0}}(x_1, x_2)$ can be expressed by

$$C_{n,0}^{f,g} := \exists f, f_1.R_{A_1,\bar{0}}(x_1, x_2) \sqcap \exists f_1, f_2.R_{A_2,\bar{0}}(x_1, x_2) \sqcap \dots \sqcap \exists f_{n-1}, f_n.R_{A_n,\bar{0}}(x_1, x_2) \sqcap \exists f_n, g.R_{A_{n+1},\bar{0}}(x_1, x_2)$$

where

$$A_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad A_i = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ for } i \in \{2, \dots, n\}, \quad A_{n+1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

To see this, note that $A_{n+1} \dots A_1 = A_{n,0}$. We assume that the features f_1, \dots, f_n are unique for $C_{n,0}^{f,g}$, i.e., they do not appear in any other concept description.

Analogously, there exists a concept $C_{0,n}^{f,g}$ constructed using our selected set of matrices and vectors which expresses $\exists f, g.R_{A_{0,n},\bar{0}}(x_1, x_2)$ where the affine transformation $\bar{x} \mapsto A_{0,n}\bar{x}$ multiplies the first component by 0 and the second component by n . Again, we assume that each feature in $C_{0,n}^{f,g}$ beside f and g does not appear in any other concept description.

Now, guided by (\dagger), we can express the original concept $\exists f, g.R_{M,\bar{0}}(x_1, x_2)$ by

$$C_{q_1,2p_1,1,0}^{f,f_1,1} \sqcap C_{0,q_1,1p_1,2}^{f,f_1,2} \sqcap C_{q_1,1q_1,2,0}^{g,g_1} \sqcap C_{1,0}^{h_1,f_1,1} \sqcap C_{0,1}^{h_1,f_1,2} \sqcap \exists h_1, g_1.R_{B_1,\bar{0}}(x_1, x_2) \\ \sqcap C_{q_2,2p_2,1,0}^{f,f_2,1} \sqcap C_{0,q_2,1p_2,2}^{f,f_2,2} \sqcap C_{0,q_2,1q_2,2}^{g,g_2} \sqcap C_{1,0}^{h_2,f_2,1} \sqcap C_{0,1}^{h_2,f_2,2} \sqcap \exists h_2, g_2.R_{B_2,\bar{0}}(x_1, x_2)$$

where $f_{1,1}, f_{1,2}, f_{2,1}, f_{2,2}, g_1, g_2, h_1, h_2$ are fresh features and

$$B_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Now consider an arbitrary tuple \bar{v} where, w.l.o.g., $\bar{v}[i] = p_i/q_i$ for an integer p_i and a positive integer q_i . We have that $(\bar{x}, \bar{y}) \in R_{E_2,\bar{v}}$ iff

$$q_i \bar{x}[i] + p_i = q_i \bar{y}[i] \text{ for } i \in \{1, 2\}. \tag{*}$$

For every $n \in \mathbb{Z}$, we can show similarly as above that there exist concepts $D_{n,0}^{f,f}$ and $D_{0,n}^{f,f}$ constructed using our selected finite set of matrices and vectors which express the concepts $\exists f, f.R_{N,(n,0)}(x_1, x_2)$ and $\exists f, f.R_{N,(0,n)}(x_1, x_2)$, respectively, where N is the 2×2 matrix of zeros. One can then express the concept $\exists f, g.R_{E_2,\bar{v}}(x_1, x_2)$ using the concepts $C_{n,0}^{f,g}, C_{0,n}^{f,g}, D_{n,0}^{f,f}$, and $D_{0,n}^{f,f}$ while being guided by (*). \square

A Proof of Proposition 6

Let τ be the signature of \mathfrak{D} . Recall that, since \mathfrak{D} is finitely bounded, by Lemma 3, there exists a universal first-order sentence $\Phi(\mathfrak{D})$ that defines $\text{Age}(\mathfrak{D})$, i.e., a finite τ -structure can be embedded into \mathfrak{D} iff it satisfies $\Phi(\mathfrak{D})$. Since the structure \mathfrak{D} is fixed, this sentence is also fixed, which means that it has constant size. We refer to $\Phi(\mathfrak{D})$ simply by Φ .

For (1), we show that $\text{CSP}(\mathfrak{D})$ is definable in existential second-order logic. Then (1) follows from Fagin’s theorem [42] (see also [54]). Let R_1, \dots, R_ℓ be an enumeration of the symbols in τ . For every $i \in [\ell]$, we introduce a second-order variable S_i of the same arity as R_i . Moreover, we introduce a binary second-order variable \sim . We obtain Φ' from Φ by replacing each atomic formula of the form $R_i(\bar{x})$ for $i \in [\ell]$ in Φ by $S_i(\bar{x})$, and each atomic formula of the form $(x \sim y)$ in Φ by $(x \sim y)$. For every $i \in [\ell]$, let n_i be the arity of R_i , and let Θ_i be the sentence

$$\Theta_i := \bigwedge_{j \in [n_i]} \forall x_1, \dots, x_{n_i}, y. S_i(x_1, \dots, x_{n_i}) \wedge (x_j \sim y) \Rightarrow S_i(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_{n_i}).$$

Now consider the existential second-order sentence Ψ defined as follows:

$$\Psi := \exists \sim. \exists S_1 \dots \exists S_\ell. \forall x, y, z. (x \sim y \wedge y \sim z \Rightarrow x \sim z) \wedge (x \sim y \Rightarrow y \sim x) \wedge (x \sim x) \wedge \Phi' \wedge \bigwedge_{i \in [\ell]} \Theta_i \wedge \forall \bar{x}. R_i(\bar{x}) \Rightarrow S_i(\bar{x}).$$

Let \mathfrak{A} be an instance of $\text{CSP}(\mathfrak{D})$. Suppose that \mathfrak{A} satisfies Ψ . By the definition of Ψ , \sim is an equivalence relation on A and also compatible with the relations S_i for $i \in [\ell]$ (due to the sentences Θ_i). This means that the structure \mathfrak{A}/\sim on the equivalence classes of \sim and with the relations

$$R_i^{\mathfrak{A}/\sim} = \{([x_1]_\sim, \dots, [x_{n_i}]_\sim) \in (A/\sim)^{n_i} \mid (x_1, \dots, x_{n_i}) \in S_i\}$$

is well defined. By the definition of Ψ , we have that $\mathfrak{A} \rightarrow \mathfrak{A}/\sim$ and that $\mathfrak{A}/\sim \models \Phi$. Since Φ defines $\text{Age}(\mathfrak{D})$, we conclude that $\mathfrak{A} \rightarrow \mathfrak{D}$. On the other hand, if there exists a homomorphism $h: \mathfrak{A} \rightarrow \mathfrak{D}$, then $S_i := \{\bar{x} \in A^{n_i} \mid h(\bar{x}) \in R_i^{\mathfrak{D}}\}$ for $i \in [\ell]$ and $\sim := \{(x, y) \in A^2 \mid h(x) = h(y)\}$ witness that Ψ is satisfied in \mathfrak{A} .

For (2), we describe a PSPACE algorithm that decides the first-order theory of \mathfrak{D} . It is based on the algorithm from the proof of Proposition 3.5 in [59], for which an exponential time complexity is shown in [59]. Note that, since D is possibly infinite, we cannot simply substitute all elements from D , one after the other, for a particular quantified variable.

Now, let b_1, b_2, \dots be a countably infinite sequence of pairwise distinct symbols. For a first-order τ -formula ϕ with free variables x_1, \dots, x_n , let $[\phi]_{\mathfrak{D}}$ denote the set of all τ -structures \mathfrak{B} with domain $\{b_1, \dots, b_n\}$ for which there exists an embedding $h: \mathfrak{B} \hookrightarrow \mathfrak{D}$ such that $\mathfrak{D} \models \phi(h(b_1), \dots, h(b_n))$. Every such embedding $h: \mathfrak{B} \hookrightarrow \mathfrak{D}$ represents an injective⁸ substitution of elements from D for the variables x_1, \dots, x_n . We claim that $[\phi]_{\mathfrak{D}}$ does not depend on the choice of h . To see this, consider two embeddings $h_1, h_2: \mathfrak{B} \hookrightarrow \mathfrak{D}$ such that $\mathfrak{D} \models \phi(h_1(b_1), \dots, h_1(b_n))$. For each $i \in [2]$, let \mathfrak{B}_i be the substructure of \mathfrak{D} on the image of $\{b_1, \dots, b_n\}$ under h_i . Consider the map $\tilde{f}: B_1 \rightarrow B_2$ that sends, for every $j \in [n]$, $h_1(b_j)$ to $h_2(b_j)$. Using the definition of an embedding, it is easy to show that \tilde{f} is an isomorphism from \mathfrak{B}_1 to \mathfrak{B}_2 . By assumption, \mathfrak{D} is homogeneous. By homogeneity of \mathfrak{D} , there exists an automorphism f of \mathfrak{D} that extends \tilde{f} . Since ϕ is a first-order formula, $\phi^{\mathfrak{D}}$ is preserved by f , which shows that $\mathfrak{D} \models \phi(h_2(b_1), \dots, h_2(b_n))$ holds as well.

We show by induction on the structure of a first-order τ -formula ϕ with free variables x_1, \dots, x_n that, given a τ -structure \mathfrak{B} with domain $\{b_1, \dots, b_n\}$, it can be decided in PSPACE in the size of ϕ whether $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$. This proves the PSPACE upper bound claimed in the proposition because, if ϕ has no free variables, then testing whether the empty structure is contained in $[\phi]_{\mathfrak{D}}$ is equivalent to answering $\mathfrak{D} \models \phi$.

⁸ In our proof we will ensure that injective substitutions are sufficient, by appropriately identifying variables.

In the *base case*, we consider an atomic formula $\phi(x_1, \dots, x_n)$. Suppose that \mathfrak{B} is a τ -structure with domain $\{b_1, \dots, b_n\}$. If $\mathfrak{B} \models \neg\phi(b_1, \dots, b_n)$, then clearly $\mathfrak{B} \notin [\phi(x_1, \dots, x_n)]_{\mathfrak{D}}$ because embeddings are injective and preserve complements of relations. If $\mathfrak{B} \models \phi(b_1, \dots, b_n)$, then $\mathfrak{D} \models \phi(h(b_1), \dots, h(b_n))$ holds for every embedding $h: \mathfrak{B} \rightarrow \mathfrak{D}$. Consequently, testing whether $\mathfrak{B} \in [\phi(x_1, \dots, x_n)]_{\mathfrak{D}}$ boils down to testing whether $\mathfrak{B} \hookrightarrow \mathfrak{D}$, which is the case iff $\mathfrak{B} \models \Phi$. This can be done in PSPACE in the size of ϕ because it is well known that first-order model checking with a fixed first-order sentence can be done in polynomial time in the size of the input structure.

For the *induction step*, we can restrict the attention to formulas ϕ of the form $\psi_1 \vee \psi_2$, $\neg\psi$ and $\exists x.\psi$. Suppose that ϕ is of the form $\psi_1 \vee \psi_2$ such that the induction hypothesis applies to both ψ_1 and ψ_2 . For each $i \in [2]$, let \mathfrak{B}_i be the substructure of \mathfrak{B} on those b_j s that correspond to the free variables of ψ_i . We claim that $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$ iff $\mathfrak{B} \models \Phi$ and $\mathfrak{B}_i \in [\psi_i]_{\mathfrak{D}}$ for $i = 1$ or $i = 2$. The forward direction is trivial. Now suppose that $\mathfrak{B}_i \in [\psi_i]_{\mathfrak{D}}$ for $i = 1$ or $i = 2$ and $\mathfrak{B} \models \Phi$. Then we have an embedding $h_i: \mathfrak{B}_i \hookrightarrow \mathfrak{D}$ witnessing $\mathfrak{B}_i \in [\psi_i]_{\mathfrak{D}}$, and we also have an embedding $h: \mathfrak{B} \hookrightarrow \mathfrak{D}$. But then $\mathfrak{B}_i \in [\psi_i]_{\mathfrak{D}}$ is also witnessed by $h|_{B_i}$ because $[\psi_i]_{\mathfrak{D}}$ does not depend on the choice of the embedding. This shows that $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$ is witnessed by h . Testing whether $\mathfrak{B}_i \in [\psi_i]_{\mathfrak{D}}$ can be done in PSPACE in the size of ψ_i by the induction hypothesis, and we have already seen in the base case that testing whether $\mathfrak{B} \models \Phi$ can be done in polynomial time in the size of ϕ .

Suppose that ϕ is of the form $\neg\psi$ such that the induction hypothesis applies to ψ . We claim that $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$ iff $\mathfrak{B} \notin [\psi]_{\mathfrak{D}}$. Suppose that there exists $h: \mathfrak{B} \hookrightarrow \mathfrak{D}$ such that $\mathfrak{D} \models \neg\psi(h(b_1), \dots, h(b_n))$. Then there cannot be an embedding $h': \mathfrak{B} \hookrightarrow \mathfrak{D}$ such that $\mathfrak{D} \models \psi(h'(b_1), \dots, h'(b_n))$ because containment in $[\phi]_{\mathfrak{D}}$ does not depend on the choice of the embedding. The backward direction is analogous. By the induction hypothesis, testing whether $\mathfrak{B} \in [\psi]_{\mathfrak{D}}$ can be done in PSPACE in the size of ψ and thus also in the size of ϕ .

Now suppose that ϕ is of the form $\phi(x_1, \dots, x_n) = \exists x_{n+1}.\psi(x_1, \dots, x_n, x_{n+1})$ such that the induction hypothesis applies to ψ . We claim that $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$ iff one of the following is true

1. there exists an extension \mathfrak{B}' of \mathfrak{B} by b_{n+1} such that $\mathfrak{B}' \in [\psi]_{\mathfrak{D}}$,
2. there exists $i \in [n]$ such that $\mathfrak{B} \in [\psi_i]_{\mathfrak{D}}$ holds for the formula ψ_i obtained from ψ by replacing each occurrence of the variable x_{n+1} in ψ by x_i .

First, suppose that $\mathfrak{B} \in [\phi]_{\mathfrak{D}}$ is witnessed by some embedding $h: \mathfrak{B} \hookrightarrow \mathfrak{D}$. Then there exists $d \in D$ such that $\mathfrak{D} \models \psi(h(b_1), \dots, h(b_n), d)$. If d is distinct from $h(b_1), \dots, h(b_n)$, then we are in the case (1) and consider the extension h' of h that maps b_{n+1} to d . We define \mathfrak{B}' as the τ -structure with the domain $\{b_1, \dots, b_{n+1}\}$ such that, for every k -ary symbol $R \in \tau$, we have $t \in R^{\mathfrak{B}'}$ iff $h'(t) \in R^{\mathfrak{D}}$. Clearly h' is an embedding that witnesses $\mathfrak{B}' \in [\psi]_{\mathfrak{D}}$. Otherwise we have $d = h(b_i)$ for some $i \in [n]$. We consider the formula ψ_i from (2). Then h is an embedding that witnesses $\mathfrak{B} \in [\psi_i]_{\mathfrak{D}}$. Since the backward direction is obvious, it remains to show that the tests required by (1) and (2) can be performed in PSPACE.

In case (1), we generate all extensions \mathfrak{B}' of \mathfrak{B} by b_{n+1} and test, using the induction hypothesis, whether $\mathfrak{B}' \in [\psi]_{\mathfrak{D}}$ for some such extension. This can clearly be done in PSPACE because τ is fixed and finite, and for each extension \mathfrak{B}' we can test $\mathfrak{B}' \in [\psi]_{\mathfrak{D}}$ within PSPACE due to the induction hypothesis. In case (2) we guess any such $i \in [n]$ and test, using the induction hypothesis, whether $\mathfrak{B} \in [\psi_i]_{\mathfrak{D}}$. This completes the proof. □

A Proof of Proposition 9

For brevity we write \mathfrak{A} for the disjoint union $\bigcup_{i=1}^k \mathfrak{A}_i$. Let σ be the signature τ extended by a unary symbol D_i for each $i \in [k]$. Consider the σ -expansion \mathfrak{A}' of \mathfrak{A} where $D_i^{\mathfrak{A}'} = A_i$ for each $i \in [k]$.

To show that \mathfrak{A}' is homogeneous, we first observe the following. If, for each $i \in [k]$, f_i is an automorphism of \mathfrak{A}_i , then the map $f : A \rightarrow A$ satisfying $f|_{A_i} := f_i$ is an automorphism of \mathfrak{A}' since it additionally preserves $D_i^{\mathfrak{A}'}$ for each $i \in [k]$. Conversely, if f is an automorphism of \mathfrak{A}' , then $f|_{A_i}$ is an automorphism of \mathfrak{A}_i for each $i \in [k]$. Now, let $f : \mathfrak{B}_1 \rightarrow \mathfrak{B}_2$ be an isomorphism between two finite substructures of \mathfrak{A}' . Since f preserves $D_i^{\mathfrak{B}_1} = B_1 \cap A_i$ for each $i \in [k]$, the restrictions $f|_{B_1 \cap A_i}$ are isomorphisms, and thus extend to automorphism of \mathfrak{A}_i for each $i \in [k]$ by homogeneity of the structures \mathfrak{A}_i . By the observation about automorphisms above, this implies that f itself extends to an automorphism of \mathfrak{A}' .

Next we show that \mathfrak{A}' is finitely bounded. For each $i \in [k]$, let

$$\Phi(\mathfrak{A}_i) = \forall x_{i_1}, \dots, x_{i_{n_i}} \cdot \phi_i(x_{i_1}, \dots, x_{i_{n_i}}) \text{ with } \phi_i \text{ quantifier-free}$$

be a universal sentence that defines $\text{Age}(\mathfrak{A}_i)$. Now consider the universal sentence

$$\begin{aligned} \Phi(\mathfrak{A}') := & \left(\forall x. \bigwedge_{i \neq j} \neg(D_i(x) \wedge D_j(x)) \right) \wedge \left(\forall x. \bigvee_{i=1}^k D_i(x) \right) \\ & \wedge \bigwedge_{i=1}^k \forall x_{i_1}, \dots, x_{i_{n_i}} \cdot \left(\bigwedge_{j=1}^{n_i} D_i(x_{i_j}) \Rightarrow \phi_i(x_{i_1}, \dots, x_{i_{n_i}}) \right). \end{aligned}$$

Let \mathfrak{B} be a finite σ -structure that satisfies $\Phi(\mathfrak{A}')$. By the first line in $\Phi(\mathfrak{A}')$, the unary relations $D_i^{\mathfrak{B}}$ are pairwise disjoint and exhaustive. By the second line in $\Phi(\mathfrak{A}')$, the τ -reduct of the substructure of \mathfrak{B} on $D_i^{\mathfrak{B}}$ is contained in $\text{Age}(\mathfrak{A}_i)$ for each $i \in [k]$. Hence \mathfrak{B} is a substructure of \mathfrak{A}' . Conversely, every finite substructure of \mathfrak{A}' must satisfy $\Phi(\mathfrak{A}')$. This completes the proof as \mathfrak{A} is the τ -reduct of \mathfrak{A}' . □

A Proof of Proposition 10

By proj_i we denote the usual *projection* function $\text{proj}_i : A_1 \times \dots \times A_k \rightarrow A_i$ with $\text{proj}_i(\vec{i}) = \vec{i}[i]$. We use the abbreviation $\mathfrak{A} := \mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$ and denote the signature of \mathfrak{A} by τ .

For (1), let $f : \mathfrak{B}_1 \rightarrow \mathfrak{B}_2$ be an isomorphism between two finite substructures \mathfrak{B}_1 and \mathfrak{B}_2 of \mathfrak{A} . For every $i \in [k]$, we define $\mathfrak{B}_{1,i}$ and $\mathfrak{B}_{2,i}$ as the substructure of \mathfrak{A}_i on $\text{proj}_i(B_1)$ and $\text{proj}_i(B_2)$, respectively. For every $i \in [k]$ and $R \in \tau_i \cup \{=i\}$, the relation $R^{\mathfrak{B}_1}$ is preserved by f . Consider the map $f_i : \mathfrak{B}_{1,i} \rightarrow \mathfrak{B}_{2,i}$ given by $f_i(\vec{i}[i]) := f(\vec{i})[i]$. It is well defined, since for any $\vec{i}, \vec{i}' \in B_1$ with $\vec{i}[i] = \vec{i}'[i]$, we have $(\vec{i}, \vec{i}') \in =_i^{\mathfrak{B}_1}$, which implies $f(\vec{i})[i] = f(\vec{i}')[i]$, because $=_i$ is preserved by f . Since f is an isomorphism, the previous argument can also be read backwards, which implies that f_i is injective. It follows directly from the definition of f_i that it is surjective, because f is surjective. Finally, f_i is an isomorphism since, for every $R \in \tau_i \cup \{=i\}$, we have

$$\begin{aligned}
 (\bar{t}_1[i], \dots, \bar{t}_k[i]) \in R^{\mathfrak{B}_{1,i}} & \text{ iff } (\bar{t}_1[i], \dots, \bar{t}_k[i]) \in R^{\mathfrak{A}_i} \cap \text{proj}_j(B_1)^k \\
 & \text{ iff } (\bar{t}_1, \dots, \bar{t}_k) \in R^{\mathfrak{A}} \cap B_1^k \\
 & \text{ iff } (\bar{t}_1, \dots, \bar{t}_k) \in R^{\mathfrak{B}_1} \\
 & \text{ iff } (f(\bar{t}_1), \dots, f(\bar{t}_k)) \in R^{\mathfrak{B}_2} \\
 & \text{ iff } (f(\bar{t}_1), \dots, f(\bar{t}_k)) \in R^{\mathfrak{A}} \cap B_2^k \\
 & \text{ iff } (f(\bar{t}_1)[i], \dots, f(\bar{t}_k)[i]) \in R^{\mathfrak{A}_i} \cap \text{proj}_j(B_2)^k \\
 & \text{ iff } (f(\bar{t}_1)[i], \dots, f(\bar{t}_k)[i]) \in R^{\mathfrak{B}_{2,i}} \\
 & \text{ iff } (f_i(\bar{t}_1[i]), \dots, f_i(\bar{t}_k[i])) \in R^{\mathfrak{B}_{2,i}}.
 \end{aligned}$$

Each f_i extends to an automorphism f'_i of \mathfrak{A}_i , because \mathfrak{A}_i is homogeneous. Let f' be the map from A to A defined by $f'(\bar{t}) := f'_1(\bar{t}[1]), \dots, f'_k(\bar{t}[k])$. Clearly, f' is bijective because each f'_i is bijective. Let $R \in \tau$ be an arbitrary and n its arity. Then $R \in \tau_i \cup \{=i\}$ for some $i \in [k]$. Since f'_i is an automorphism of \mathfrak{A}_i , for every $(\bar{t}_1, \dots, \bar{t}_n) \in A^n$, we have

$$\begin{aligned}
 (\bar{t}_1, \dots, \bar{t}_n) \in R^{\mathfrak{A}} & \text{ iff } (\bar{t}_1[i], \dots, \bar{t}_n[i]) \in R^{\mathfrak{A}_i} \\
 & \text{ iff } (f'_i(\bar{t}_1[i]), \dots, f'_i(\bar{t}_n[i])) \in R^{\mathfrak{A}_i} \\
 & \text{ iff } (f'(\bar{t}_1)[i], \dots, f'(\bar{t}_n)[i]) \in R^{\mathfrak{A}_i} \\
 & \text{ iff } (f'(\bar{t}_1), \dots, f'(\bar{t}_n)) \in R^{\mathfrak{A}}.
 \end{aligned}$$

Hence, f' is an automorphism of \mathfrak{A} . It follows from the definition of f_i that f' extends f .

For (2) let, for each $i \in [k]$, $\Phi(\mathfrak{A}_i)$ be the universal sentence that defines $\text{Age}(\mathfrak{A}_i)$. Let $\Phi'(\mathfrak{A}_i)$ be the sentence obtained from $\Phi(\mathfrak{A}_i)$ by replacing each occurrence of an atomic formula of the form $(x = y)$ in $\Phi'(\mathfrak{A}_i)$ by $(x =_i y)$. Furthermore, for each symbol $R \in \tau_i$ of arity n other than $=_i$, let ψ_R be the sentence

$$\forall x_1, \dots, x_n, y_1, \dots, y_n. \left(\bigwedge_{j=1}^n x_j =_i y_j \right) \Rightarrow (R(x_1, \dots, x_n) \Leftrightarrow R(y_1, \dots, y_n)).$$

Now consider the τ -sentence

$$\begin{aligned}
 \Phi(\mathfrak{A}) := & \bigwedge_{i=1}^k \left(\forall x, y, z. (x =_i x) \wedge (x =_i y \Leftrightarrow y =_i x) \wedge (x =_i y \wedge y =_i z \Rightarrow x =_i z) \right) \\
 & \wedge \left(\forall x, y. (x = y) \Leftrightarrow \bigwedge_{i=1}^k (x =_i y) \right) \wedge \bigwedge_{i=1}^k \Phi'(\mathfrak{A}_i) \wedge \bigwedge_{R \in \tau \setminus \{=1, \dots, =k\}} \psi_R.
 \end{aligned}$$

We claim that $\Phi(\mathfrak{A})$ defines $\text{Age}(\mathfrak{A})$.

For the forward direction, let \mathfrak{B} be a finite substructure of $\mathfrak{A} := \mathfrak{A}_1 \boxtimes \dots \boxtimes \mathfrak{A}_k$. By the definition of \boxtimes , the relation $=_i^{\mathfrak{A}}$ is an equivalence relation for each $i \in [k]$, because $=_i^{\mathfrak{A}_i}$ is an equivalence relation. Since \mathfrak{B} is a substructure of \mathfrak{A} , $=_i^{\mathfrak{B}}$ is an equivalence relation for each $i \in [k]$ as well. Thus \mathfrak{B} satisfies the first line in $\Phi(\mathfrak{A})$. For all $\bar{t}, \bar{s} \in B$ we have $\bar{t} = \bar{s}$ iff $\bar{t} =_i^{\mathfrak{B}} \bar{s}$ for each $i \in [k]$, because $=_i^{\mathfrak{B}}$ stands for the equality in the i -th coordinate. Thus \mathfrak{B} satisfies the first clause on the second line in $\Phi(\mathfrak{A})$. Let \mathfrak{B}_i be the substructure of \mathfrak{A}_i on $\text{proj}_i(B)$. As a substructure of \mathfrak{A}_i , \mathfrak{B}_i satisfies $\Phi(\mathfrak{A}_i)$ because $\Phi(\mathfrak{A}_i)$ defines $\text{Age}(\mathfrak{A}_i)$. But then \mathfrak{B}_i must also satisfy $\Phi'(\mathfrak{A}_i)$ because $=_i$ interprets as the binary equality predicate in \mathfrak{B}_i . We claim that \mathfrak{B} satisfies $\Phi'(\mathfrak{A}_i)$ for each $i \in [k]$. Let $\bar{t}_1, \dots, \bar{t}_m \in B$ be any tuples to be substituted for the universally quantified variables x_1, \dots, x_m of $\Phi'(\mathfrak{A}_i)$.

Let $\psi'(x_1, \dots, x_m)$ be a formula in DNF equivalent to the quantifier-free part of $\Phi'(\mathfrak{A}_i)$. Let ψ^* be a disjunct in ψ' such that $\mathfrak{B}_i \models \psi^*(\bar{t}_1[i], \dots, \bar{t}_m[i])$. Recall that $\Phi'(\mathfrak{A}_i)$ contains no atomic formulas of the form $(x = y)$. Also recall that, for every n -ary symbol $R \in \tau_i$, we have $(\bar{t}_1, \dots, \bar{t}_n) \in R^{\mathfrak{B}}$ iff $(\bar{t}_1[i], \dots, \bar{t}_n[i]) \in R^{\mathfrak{B}_i}$ by the definition of \boxtimes . This means that, if ψ^* contains an atomic formula of the form $R(x_{i_1}, \dots, x_{i_n})$ for some n -ary symbol $R \in \tau_i$, then we have $\mathfrak{B}_i \models R(\bar{t}_1[i], \dots, \bar{t}_n[i])$ iff $\mathfrak{B} \models R(\bar{t}_1, \dots, \bar{t}_n)$. Likewise we have $\mathfrak{B}_i \models \neg R(\bar{t}_1[i], \dots, \bar{t}_n[i])$ iff $\mathfrak{B} \models \neg R(\bar{t}_1, \dots, \bar{t}_n)$. Since $\mathfrak{B} \models \psi^*(\bar{t}_1, \dots, \bar{t}_m)$ and $\bar{t}_1, \dots, \bar{t}_n$ were chosen arbitrarily, we conclude that $\mathfrak{B} \models \Phi'(\mathfrak{A}_i)$. It follows directly from the argumentation above and the fact that $=_i$ interprets as the binary equality predicate in \mathfrak{A}_i that $\mathfrak{B} \models \psi_R$ for each $R \in \tau \setminus \{=, \dots, =_k\}$. Hence $\mathfrak{B} \models \Phi(\mathfrak{A})$.

For the backward direction, let \mathfrak{B} be a finite τ -structure that satisfies $\Phi(\mathfrak{A})$. Then $=_i^{\mathfrak{B}}$ is an equivalence relation for each $i \in [k]$. For each $i \in [k]$, consider the following τ_i -structure \mathfrak{B}_i . The domain of \mathfrak{B}_i consists of the equivalence classes w.r.t. $=_i^{\mathfrak{B}}$. Moreover, for each n -ary symbol $R \in \tau_i$, we have $(X_1, \dots, X_n) \in R^{\mathfrak{B}_i}$ iff $(b_1, \dots, b_n) \in R^{\mathfrak{B}}$ for some representatives $b_i \in X_i$. The relations of \mathfrak{B}_i are well defined because $\mathfrak{B} \models \psi_R$ for each $R \in \tau \setminus \{=, \dots, =_k\}$. We claim that $\mathfrak{B}_i \models \Phi'(\mathfrak{A}_i)$ for each $i \in [k]$. Recall that $\Phi'(\mathfrak{A}_i)$ contains no atomic formulas of the form $(x = y)$. Let X_1, \dots, X_m be any equivalence classes of elements from B w.r.t. $=_i^{\mathfrak{B}}$ to be substituted for the universally quantified variables x_1, \dots, x_m of $\Phi'(\mathfrak{A}_i)$, and b_1, \dots, b_m any representatives of these equivalence classes, respectively. Let $\psi'(x_1, \dots, x_m)$ be a formula in DNF equivalent to the quantifier-free part of $\Phi'(\mathfrak{A}_i)$. Since $\mathfrak{B} \models \Phi'(\mathfrak{A}_i)$, we have that $\mathfrak{B} \models \psi'(b_1, \dots, b_m)$. Let ψ^* be a disjunct in ψ' such that $\mathfrak{B} \models \psi^*(b_1, \dots, b_m)$.

If ψ^* contains an atomic formula of the form $(x_{i_1} =_i x_{i_2})$, then we have $\mathfrak{B} \models (b_{i_1} =_i b_{i_2})$. This means that b_{i_1} and b_{i_2} are contained in the same equivalence class w.r.t. $=_i^{\mathfrak{B}}$, that is, $X_{i_1} = X_{i_2}$. We conclude that $\mathfrak{B}_i \models (X_{i_1} =_i X_{i_2})$ because the symbol $=_i$ interprets in \mathfrak{B}_i as the binary equality predicate. If ψ^* contains the negation of an atomic formula of the form $(x_{i_1} =_i x_{i_2})$, then we have $\mathfrak{B} \models \neg(b_{i_1} =_i b_{i_2})$ which means that b_{i_1} and b_{i_2} are contained in distinct equivalence classes. Then clearly $\mathfrak{B}_i \models \neg(X_{i_1} =_i X_{i_2})$.

If ψ^* contains an atomic formula of the form $R(x_{i_1}, \dots, x_{i_n})$ for some n -ary symbol $R \in \tau_i \setminus \{=_i\}$, then we have $\mathfrak{B} \models R(b_{i_1}, \dots, b_{i_n})$. It follows directly from the definition of \mathfrak{B}_i that $\mathfrak{B}_i \models R(X_{i_1}, \dots, X_{i_n})$. If ψ^* contains the negation of an atomic formula of the form $R(x_{i_1}, \dots, x_{i_n})$ for some n -ary symbol $R \in \tau_i$, then we have $\mathfrak{B} \models \neg R(b_{i_1}, \dots, b_{i_n})$. Suppose that $(X_{i_1}, \dots, X_{i_n}) \in R^{\mathfrak{B}_i}$. Then $(b'_{i_1}, \dots, b'_{i_n}) \in R^{\mathfrak{B}}$ for some representatives b'_{i_ℓ} of X_{i_ℓ} . But then $(b_{i_1}, \dots, b_{i_n}) \in R^{\mathfrak{B}}$ because $\mathfrak{B} \models \psi_R$, a contradiction. Thus $\mathfrak{B}_i \models \neg R(X_{i_1}, \dots, X_{i_n})$.

Since $\mathfrak{B}_i \models \psi'(X_1, \dots, X_m)$ and X_1, \dots, X_m were chosen arbitrarily, we conclude that $\mathfrak{B}_i \models \Phi'(\mathfrak{A}_i)$. Since the symbol $=_i$ interprets in \mathfrak{B}_i as the binary equality predicate, we have that $\mathfrak{B}_i \models \Phi(\mathfrak{A}_i)$. Thus $\mathfrak{B}_i \in \text{Age}(\mathfrak{A}_i)$ for each $i \in [k]$. For each $i \in [k]$, let e_i be an embedding from \mathfrak{B}_i into \mathfrak{A}_i . For each $b \in B$ and each $i \in [k]$, we denote by $[b]_{=_i^{\mathfrak{B}}}$ the equivalence class of $b \in B$ w.r.t. $=_i^{\mathfrak{B}}$. Now consider the map

$$e: B \rightarrow A_1 \times \dots \times A_k, \quad b \mapsto (e_1([b]_{=_1^{\mathfrak{B}}}), \dots, e_k([b]_{=_k^{\mathfrak{B}}}))$$

Note that e is well defined because we map from elements to their equivalence classes and not the other way around. By the first clause on the second line in $\Phi(\mathfrak{A})$, for all $x, y \in B$, we have $x = y$ iff $x =_i^{\mathfrak{B}} y$ for each $i \in [k]$. This means that e is injective. For every $i \in [k]$ and every n -ary symbol $R \in \tau_i$, we have

$$\begin{aligned}
(b_1, \dots, b_n) \in R^{\mathfrak{B}} &\text{ iff } ([b_1]_{=i}^{\mathfrak{B}}, \dots, [b_n]_{=i}^{\mathfrak{B}}) \in R^{\mathfrak{B}_i} \\
&\text{ iff } (e_i([b_1]_{=i}^{\mathfrak{B}}), \dots, e_i([b_n]_{=i}^{\mathfrak{B}})) \in R^{\mathfrak{A}_i} \\
&\text{ iff } (e(b_1)[i], \dots, e(b_n)[i]) \in R^{\mathfrak{A}_i} \\
&\text{ iff } (e(b_1), \dots, e(b_n)) \in R^{\mathfrak{A}}.
\end{aligned}$$

Hence e is an embedding from \mathfrak{B} into \mathfrak{A} . This completes the proof. \square

References

1. Allen, J.F.: Maintaining knowledge about temporal intervals. *Commun. ACM* **26**(11), 832–843 (1983)
2. Baader, F., Hanschke, P.: A scheme for integrating concrete domains into concept languages. In: Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI'91), pp. 452–457 (1991). Long version available as [6]
3. Baader, F., Hanschke, P.: A scheme for integrating concrete domains into concept languages. Tech. Rep. RR-91-10, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) (1991). <https://lat.inf.tu-dresden.de/research/reports/1991/DFKI-RR-91-10.pdf>
4. Baader, F., Hanschke, P.: Extensions of concept languages for a mechanical engineering application. In: Proceedings of the 16th German Workshop on Artificial Intelligence (GWAI'92), Lecture Notes in Computer Science, vol. 671, pp. 132–143. Springer, Berlin (1992)
5. Baader, F., Rydval, J.: An algebraic view on p-admissible concrete domains for lightweight description logics (extended version). LTCS-Report 20-10, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2020). <https://tu-dresden.de/inf/lat/reports#BaRy-LTCS-20-10>
6. Baader, F., Rydval, J.: Description logics with concrete domains and general concept inclusions revisited. In: Stokkermans, V.S., Peltier, N. (eds.) Proceedings of the International Joint Conference on Automated Reasoning (IJCAR 2020), Lecture Notes in Computer Science, vol. 12166, pp. 413–431. Springer, New York (2020)
7. Baader, F., Rydval, J.: Using model-theory to find ω -admissible concrete domains. LTCS-Report 20-01, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2020). <https://tu-dresden.de/inf/lat/reports#BaRy-LTCS-20-01>
8. Baader, F., Rydval, J.: An algebraic view on p-admissible concrete domains for lightweight description logics. In: Faber, W., Friedrich, D., Gebser, M. (eds.) 17th European Conference on Logics in Artificial Intelligence (JELIA 2021), Lecture Notes in Artificial Intelligence, vol. 12678, pp. 194–209. Springer, New York (2021)
9. Baader, F., Bürckert, H.J., Hollunder, B., Nutt, W., Siekmann, J.H.: Concept logics. In: Lloyd, J.W. (ed.) Computational Logics, Symposium Proceedings, pp. 177–201. Springer, Berlin (1990)
10. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, Cambridge (2003)
11. Baader, F., Brandt, S., Lutz, C.: Pushing the \mathcal{EL} envelope. In: Kaelbling, L.P., Saffiotti, A. (eds.) Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005), pp. 364–369. Morgan Kaufmann, Los Altos, Edinburgh (UK) (2005)
12. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press, Cambridge (2017)
13. Barto, L., Kompatscher, M., Olšák, M., Van Pham, T., Pinsker, M.: Equations in oligomorphic clones and the Constraint Satisfaction Problem for ω -categorical structures. *J. Math. Logic* **19**(2) (2019)
14. Bell, P., Potapov, I.: On undecidability bounds for matrix decision problems. *Theor. Comput. Sci.* **391**(1–2), 3–13 (2008)
15. Bodirsky, M.: The core of a countably categorical structure. In: Diekert, V., Durand, B. (eds.) Proceedings of the 22nd Annual Symposium on Theoretical Aspects of Computer Science (STACS 2005), Lecture Notes in Computer Science, vol. 3404, pp. 110–120. Springer, Berlin (2005)
16. Bodirsky, M.: Complexity classification in infinite-domain constraint satisfaction (2012). Mémoire d'Habilitation à Diriger des Recherches, Université Diderot—Paris 7. <https://arxiv.org/pdf/1201.0856.pdf>
17. Bodirsky, M., Dalmau, V.: Datalog and constraint satisfaction with infinite templates. *J. Comput. Syst. Sci.* **79**(1), 79–100 (2013)

18. Bodirsky, M., Kára, J.: The complexity of temporal constraint satisfaction problems. *J. ACM (JACM)* **57**(2), 1–41 (2010)
19. Bodirsky, M., Mottet, A.: Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction. In: *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2016)*, pp. 623–632. ACM/IEEE (2016). <https://arxiv.org/pdf/1601.04520v1.pdf>
20. Bodirsky, M., Nešetřil, J.: Constraint satisfaction with countable homogeneous templates. *J. Log. Comput.* **16**(3), 359–373 (2006)
21. Bodirsky, M., Wöfl, S.: RCC8 is polynomial on networks of bounded treewidth. In: Walsh, T. (ed.) *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pp. 756–761. IJCAI/AAAI (2011)
22. Bodirsky, M., Chen, H., Kára, J., von Oertzen, T.: Maximal infinite-valued constraint languages. *Theor. Comput. Sci.* **410**(18), 1684–1693 (2009)
23. Bodirsky, M., Hils, M., Martin, B.: On the scope of the universal-algebraic approach to constraint satisfaction. In: *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS 2010)*, pp. 90–99. IEEE (2010)
24. Bodirsky, M., Chen, H., Feder, T.: On the complexity of MMSNP. *SIAM J. Discret. Math.* **26**(1), 404–414 (2012)
25. Bodirsky, M., Jonsson, P., Pham, T.V.: The complexity of phylogeny constraint satisfaction problems. *ACM Trans. Comput. Logic (TOCL)* **18**(3), 1–42 (2017)
26. Bodirsky, M., Madelaine, F., Mottet, A.: A universal-algebraic proof of the complexity dichotomy for monotone monadic SNP. In: *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2018)*, pp. 105–114 (2018)
27. Bodirsky, M., Martin, B., Mottet, A.: Discrete temporal constraint satisfaction problems. *J. ACM (JACM)* **65**(2), 1–41 (2018)
28. Bodirsky, M., Martin, B., Pinsker, M., Pongrácz, A.: Constraint satisfaction problems for reducts of homogeneous graphs. *SIAM J. Comput.* **48**(4), 1224–1264 (2019)
29. Bodirsky, M., Knäuper, S., Starke, F.: ASNP: A tame fragment of existential second-order logic. In: Anselmo, M., Vedova, G.D., Manea, F., Pauly, A. (eds.) *Beyond the Horizon of Computability—Proceedings of the 16th Conference on Computability in Europe (CiE 2020)*, *Lecture Notes in Computer Science*, vol. 12098, pp. 149–162. Springer (2020)
30. Bodirsky, M., Rydval, J., Schrottenloher, A.: Universal Horn sentences and the joint embedding property (2021). <https://arxiv.org/pdf/2104.11123.pdf>
31. Bojańczyk, M., Segoufin, L., Toruńczyk, S.: Verification of database-driven systems via amalgamation. In: *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pp. 63–74 (2013)
32. Brandt, S.: Polynomial time reasoning in a description logic with existential restrictions, GCI axioms, and—what else? In: de Mántaras, R.L., Saitta, L. (eds.) *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, pp. 298–302 (2004)
33. Braunfeld, S.: The undecidability of joint embedding and joint homomorphism for hereditary graph classes. *Discret. Math. Theor. Comput. Sci.* **21**(2) (2019)
34. Bulatov, A.A.: A dichotomy theorem for nonuniform CSPs. In: Umans, C. (ed.) *58th IEEE Annual Symposium on Foundations of Computer Science (FOCS 2017)*, pp. 319–330. IEEE Computer Society (2017)
35. Carapelle, C., Turhan, A.: Description logics reasoning w.r.t. general boxes is decidable for concrete domains with the EHD-property. In: Kaminka, G.A., Fox, M., Bouquet, P., Hüllermeier, E., Dignum, V., Dignum, F., van Harmelen, F. (eds.) *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI 2016)*, *Frontiers in Artificial Intelligence and Applications*, vol. 285, pp. 1440–1448. IOS Press (2016)
36. Carapelle, C., Feng, S., Kartzow, A., Lohrey, M.: Satisfiability of ECTL* with local tree constraints. *Theory Comput. Syst.* **61**(2), 689–720 (2017)
37. Chandra, A.K., Merlin, P.M.: Optimal implementation of conjunctive queries in relational data bases. In: Hopcroft, J.E., Friedman, E.P., Harrison, M.A. (eds.) *Proceedings of the 9th Annual ACM Symposium on Theory of Computing (STOC 1977)*, pp. 77–90. ACM (1977)
38. Cherlin, G., Shelah, S., Shi, N.: Universal graphs with forbidden subgraphs and algebraic closure. *Adv. Appl. Math.* **22**(4), 454–491 (1999)
39. Dabrowski, K.K., Jonsson, P., Ordyniak, S., Osipov, G.: Solving infinite-domain CSPs using the patchwork property. In: *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-21)*. AAAI Press/The MIT Press (2021)

40. Demri, S., D'Souza, D.: An automata-theoretic approach to constraint LTL. *Inf. Comput.* **205**(3), 380–415 (2007)
41. Dowling, W.F., Gallier, J.H.: Linear-time algorithms for testing the satisfiability of propositional horn formulae. *J. Log. Program.* **1**(3), 267–284 (1984)
42. Fagin, R.: Generalized first-order spectra and polynomial-time recognizable sets. In: Karp, R. (ed.) *SIAM-AMS Proceedings on Complexity of Computation*, vol. 7, pp. 43–73 (1974)
43. Feder, T., Vardi, M.Y.: Homomorphism closed vs. existential positive. In: *Proceedings of the 18th Annual IEEE Symposium of Logic in Computer Science (LICS 2003)*, pp. 311–320. IEEE (2003)
44. Gillibert, P., Jonušaš, J., Kompatscher, M., Mottet, A., Pinsker, M.: Hrushovski's encoding and ω -categorical CSP monsters. In: *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik (2020)
45. Haarslev, V., Möller, R., Wessel, M.: The description logic $\mathcal{ALCN}^{\mathcal{H}}_{R^+}$ extended with concrete domains: a practically motivated approach. In: *Proceedings of the International Joint Conference on Automated Reasoning (IJCAR 2001)*, *Lecture Notes in Artificial Intelligence*, vol. 2083, pp. 29–44. Springer, Berlin (2001)
46. Henson, C.W.: A family of countable homogeneous graphs. *Pac. J. Math.* **38**(1), 69–83 (1971)
47. Henson, C.W.: Countable homogeneous relational structures and \aleph_0 -categorical theories. *J. Symbol. Logic* **37**(3), 494–500 (1972)
48. Hirsch, R.: Relation algebras of intervals. *Artif. Intell.* **83**(2), 267–295 (1996)
49. Hodges, W.: *Model Theory*. Cambridge University Press, Cambridge (1993)
50. Hodges, W.: *A Shorter Model Theory*. Cambridge University Press, Cambridge (1997)
51. Hoehndorf, R., Schofield, P.N., Gkoutos, G.V.: The role of ontologies in biological and biomedical research: a functional perspective. *Brief. Bioinform.* **16**(6), 1069–1080 (2015)
52. Horrocks, I., Patel-Schneider, P.F., van Harmelen, F.: From SHIQ and RDF to OWL: the making of a web ontology language. *J. Web Semant.* **1**(1), 7–26 (2003)
53. Hubička, J., Nešetřil, J.: Homomorphism and embedding universal structures for restricted classes. *J. Multiple-Valued Logic Soft Comput.* **27**, 229–253 (2016)
54. Immerman, N.: *Descriptive Complexity*. Graduate Texts in Computer Science, Springer, Berlin (1999)
55. Kegel, O.H., Wehrfritz, B.A.: *Locally Finite Groups*. Elsevier, New York (2000)
56. Klin, B., Lasota, S., Ochremiak, J., Torunczyk, S.: Homomorphism problems for first-order definable structures. In: *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2016)
57. Knight, J., Lachlan, A.: Shrinking, stretching and codes for homogeneous structures. In: *Classification Theory (Chicago, IL, 1985)*, *Lecture Notes in Mathematics*, 1292, pp. 192–228. Springer, Berlin (1987)
58. Kompatscher, M., Van Pham, T.: A complexity dichotomy for poset constraint satisfaction. In: Vollmer, H., Vallée, B. (eds.) *34th Symposium on Theoretical Aspects of Computer Science (STACS 2017)*, vol. 66, pp. 47:1–47:12. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2017)
59. Kopczyński, E., Toruńczyk, S.: LOIS: an application of SMT solvers. In: King, T., Piskac, R. (eds.) *Proc. of the 14th International Workshop on Satisfiability Modulo Theories (SMT 2016)*, *CEUR Workshop Proceedings*, vol. 1617, pp. 51–60. ceur-ws.org (2016)
60. Labai, N.: Automata-based reasoning for decidable logics with data values. Dissertation, Technische Universität Wien (2021). <https://doi.org/10.34726/hss.2021.94060>
61. Labai, N., Ortiz, M., Simkus, M.: An ExpTime upper bound for \mathcal{ALC} with integers. In: Calvanese, D., Erdem, E., Thielscher, M. (eds.) *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*, pp. 614–623 (2020)
62. Lutz, C.: Interval-based temporal reasoning with general TBoxes. In: *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI 2001)*, pp. 89–94 (2001)
63. Lutz, C.: NExpTime-complete description logics with concrete domains. In: Goré, R., Leitsch, A., Nipkow, T. (eds.) *Proceedings of the International Joint Conference on Automated Reasoning (IJCAR 2001)*, no. 2083 in *Lecture Notes in Artificial Intelligence*, pp. 45–60. Springer, Siena (2001)
64. Lutz, C.: Adding numbers to the \mathcal{SHTQ} description logic—First results. In: *Proceedings of the 8th International Conference on Principles of Knowledge Representation and Reasoning (KR 2002)*, pp. 191–202. Morgan Kaufmann, Los Altos (2002)
65. Lutz, C., Miličić, M.: A tableau algorithm for description logics with concrete domains and general TBoxes. *J. Automated Reason.* **38**(1–3), 227–259 (2007)
66. Macpherson, D.: A survey of homogeneous structures. *Discret. Math.* **311**(15), 1599–1634 (2011)
67. Mottet, A., Bodirsky, M.: A dichotomy for first-order reducts of unary structures. *Logical Methods in Computer Science* **14** (2018)
68. Pach, P.P., Pinsker, M., Pluhár, G., Pongrácz, A., Szabó, C.: Reducts of the random partial order. *Adv. Math.* **267**, 94–120 (2014)

69. Pan, J.Z., Horrocks, I.: Reasoning in the $\mathcal{SHCQ}(D_n)$ description logic. In: Horrocks, I., Tessaris, S. (eds.) Proceedings of the 2002 Description Logic Workshop (DL 2002), CEUR Workshop Proceedings, vol. 53. CEUR-WS.org (2002)
70. Randell, D.A., Cui, Z., Cohn, A.G.: A spatial logic based on regions and connection. In: Proceedings of the 3rd International Conference on the Principles of Knowledge Representation and Reasoning (KR'92), pp. 165–176. Morgan Kaufmann, Los Altos (1992)
71. Schild, K.: A correspondence theory for terminological logics: preliminary report. In: Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI'91), pp. 466–471 (1991)
72. Zhuk, D.: A proof of CSP dichotomy conjecture. In: C. Umans (ed.) 58th IEEE Annual Symposium on Foundations of Computer Science (FOCS 2017), pp. 331–342. IEEE Computer Society (2017)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.