Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence

Local-to-Global Consistency Implies Tractability of Abduction

Michał Wrona
Department of Computer and Information Science, Linköping University
SE-581 83 Linköping, Sweden
Email: michal.wrona@liu.se

Abstract

Abduction is a form of nonmonotonic reasoning that looks for an explanation, built from a given set of hypotheses, for an observed manifestation according to some knowledge base. Following the concept behind the Schaefer's parametrization $\mathrm{CSP}(\Gamma)$ of the Constraint Satisfaction Problem (CSP), we study here the complexity of the abduction problem $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ parametrized by certain (ω-categorical) infinite relational structures $\Gamma, \mathcal{HYP}$, and $\mathcal{M}$ from which a knowledge base, hypotheses and a manifestation are built, respectively.

We say that Γ has local-to-global consistency if there is k such that establishing strong k-consistency on an instance of $\mathrm{CSP}(\Gamma)$ yields a globally consistent (whose every solution may be obtained straightforwardly from partial solutions) set of constraints. In this case $\mathrm{CSP}(\Gamma)$ is solvable in polynomial time. Our main contribution is an algorithm that under some natural conditions decides $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ in P when Γ has local-to-global consistency.

As we show in the number of examples, our approach offers an opportunity to consider abduction in the context of spatial and temporal reasoning (qualitative calculi such as Allen's interval algebra or RCC-5) and that our procedure solves some related abduction problems in polynomial time.

Introduction

Abduction is a form of logical inference that aims at finding explanations for observed manifestations, starting from some knowledge base. Abduction found many different applications in computer science and in artificial intelligence (Pople 1973), in particular to explanation-based diagnosis (e.g. medical diagnosis (Bylander et al. 1991)), text interpretation (Hobbs et al. 1993), and in planning (Herzig, Lang, and Marquis 2001). In this paper we study the complexity of abduction in the so-called Schaefer's framework originally used by Schaefer to study the CSP (Schaefer 1978).

The CSP is a computational decision problem whose instance is a list of variables and a set of constraints each of which is imposed locally on some subset of the variables. The question is whether there is an assignment to variables that satisfies simultaneously all present constraints. This problem is known to be NP-hard. Therefore, in order to look for easier subproblems, different parametrizations of the CSP are considered. The one we are interested in here is the problem $\mathrm{CSP}(\Gamma)$ parametrized by a relational ω-categorical structure, called also a *constraint language*, Γ which restricts the instances of the CSP to those set of constraints that can be built upon relations in Γ. In that framework one can express *network satisfaction problems* for many different qualitative calculi of crucial importance to spatial and temporal reasoning such as point algebra (Vilain, Kautz, and van Beek 1989), Allen's interval algebra (Allen 1983), left-linear point algebra (Duentsch 2005; Hirsch 1997) or RCC-5 (Duentsch 2005; Bennett 1994). In fact this approach, explained in details in (Bodirsky 2012), broadens the perspective and allows to obtain complexity results on constraint satisfaction in spatial and temporal reasoning not achievable before, e.g., (Bodirsky and Kára 2009; 2010; Bodirsky and Hils 2012).

Following this path, we study here the abduction problem $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ whose instance consists of three sets of constraints built upon relations in $\Gamma, \mathcal{HYP}$, and $\mathcal{M}$, respectively. The first one defines a *knowledge base (KB)*, the second one offers a set of *hypotheses (HYP)* whereas the third one, a *manifestation (M)*, consists of one constraint only. The question is whether there exists a subset H of HYP, called an *explanation*, consistent with the knowledge base such that the conjunction of KB and H entails M. When $\Gamma, \mathcal{HYP}$, and $\mathcal{M}$ are over the two-element domain, then $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ may be seen as the very-well studied propositional abduction problem (Eiter and Gottlob 1995; Marquis 2000; Creignou and Zanuttini 2006; Nordh and Zanuttini 2008). In this paper, however, we are interested in the problem $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ where $\Gamma, \mathcal{HYP}$, and $\mathcal{M}$ are ω-categorical structures. It offers the possibility to consider abduction in the context of spatial and temporal reasoning (Fisher, Gabbay, and Vila 2005; Renz and Nebel 2007; Duentsch 2005), in the context of qualitative calculi mentioned above. As related papers, we mention (Brusoni et al. 1998; Console, Terenziani, and Dupré 2002; Brusoni et al. 1997) where temporal abduction based on point algebra, and Allen's interval algebra has been studied. All of this motivates the study of the complexity of abduction for ω-categorical structures, which was initiated in (Schmidt and Wrona 2013) and is continued in this paper.

To make things more concrete, we now present an example where $\Gamma, \mathcal{HYP}$, and $\mathcal{M}$ are structures with first-order definitions in the ordered rationals, that is, in the structure $(\mathbb{Q}; <)$. In this case we can speak about abductive reasoning for events related by temporal point-based relations.

Example 1 *Consider the following relational structures:*

- $\Gamma := (\mathbb{Q}; \{(x, y, z) \in \mathbb{Q}^3 \mid (x > y \vee x > z)\})$*;*
- $\mathcal{HYP} := (\mathbb{Q}; \{(x, y) \in \mathbb{Q}^2 \mid (x < y)\})$*;*
- $\mathcal{M} := (\mathbb{Q}; \{(x, y) \in \mathbb{Q}^2 \mid (x < y)\})$.

We look at an instance of $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ *built upon variables* V *(which may be seen as* events*) and three sets of constraints:*

- knowledge base (KB)*, e.g.,* $(x > y \vee x > z) \wedge (y > v \vee y > w)$ *that describes temporal dependencies between events in* V *using the relation in* Γ*;*
- hypotheses (HYP) *containing all constraints of the form* $(x < y)$ *for all* $x, y \in V_H$ *where* $V_H \subseteq V$*;*
- manifestation (M) *which is one constraint:* $(v < w)$ *with* $v, w \in V$.

The question is whether there exists H $\subseteq$ *HYP (a partial order on events in* V_H*) such that constraints in KB*$\cup$*H entails* $(v < w)$ *(if constraints in KB*$\cup$*H are satisfied, then the event* v *has to take place before the event* w*).*

A primary algorithmic technique for solving $\mathrm{CSP}(\Gamma)$, is the process of establishing k-consistency (Mackworth 1977). According to the definitions in (Freuder 1982; Dechter 1992), the process of establishing k-consistency converts an instance P of the CSP into an instance P' that has the same set of solutions as P and is k-consistent, that is, every partial solution to $(k-1)$ variables may be extended to any other variable. An instance of the CSP with n variables is *strongly* k*-consistent* if it is i-consistent for any $(i \leq k)$ and is *globally consistent* if it is strongly n-consistent.

It is very well known that for many constraint languages Γ there is k such that the problem $\mathrm{CSP}(\Gamma)$ can be solved by establishing (strong) k-consistency. In this case $\mathrm{CSP}(\Gamma)$ is solvable in polynomial time. Sometimes it is enough to obtain global consistency. We say that Γ has *local-to-global consistency* if there is k such that establishing strong k-consistency on every instance P of $\mathrm{CSP}(\Gamma)$ yields a globally consistent variant P' of P.

In this article we investigate a natural question of whether under any natural conditions the problem $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ can be solved by the algorithm based on establishing k-consistency. To answer this question affirmatively, we consider a natural restriction $\mathrm{OE\text{-}ABD}_\Delta(\Gamma, \mathcal{M})$ of $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$. Our main contribution is a procedure `LtG-OEAbd` that solves the problem $\mathrm{OE\text{-}ABD}_\Delta(\Gamma, \mathcal{M})$ in polynomial time when the expansion of Γ with complements of relations in $\mathcal{M}$ has local-to-global consistency.

To show the strength of our algorithm, we investigate its applicability to temporal and spatial reasoning (point algebra, Allen's interval algebra, RCC-5). In particular we show that it solves $\mathrm{OE\text{-}ABD}_\Delta(\Gamma, \mathcal{M})$ when Γ is a basic Ord-Horn language (Nebel and Bürckert 1995; Chen and Wrona 2012; Wrona 2012) and $\mathcal{M}$ that contains all binary relations definable in $(\mathbb{Q}; <)$. We also show that our procedure solves all tractable equality abduction problems studied in (Schmidt and Wrona 2013).

We finally observe that there are structures Γ such that $\mathrm{CSP}(\Gamma)$ is solvable in P by establishing k-consistency for which $\mathrm{OE\text{-}ABD}_\Delta(\Gamma, \mathcal{M})$ is NP-hard. As an example consider $\Gamma = (\mathbb{N}; \{(x, y, z) \in \mathbb{N}^3 \mid (x \neq y \vee x = z)\})$ and $\mathcal{M} = (\mathbb{N}; =, \neq)$. It is easy to show that $\mathrm{CSP}(\Gamma)$ is solvable in P by establishing k-consistency (one can write an appropriate Datalog program, see (Bodirsky and Dalmau 2013)). On the other hand, it follows from (Schmidt and Wrona 2013) that $\mathrm{OE\text{-}ABD}_\Delta(\Gamma, \mathcal{M})$ is NP-complete. In (Wrona 2012) it was proved that Γ does not have local-to-global consistency.

Constraint Languages under Consideration

We now provide a few definitions from model theory that are of concern in this paper. For a more comprehensive reading in this topic we refer the reader to (Hodges 1993). We consider here relational structures (always if it is not specified otherwise) over finite signatures denoted by Greek letters such as Γ, or Δ. For the sake of simplicity, we use the same symbols to denote relations and their corresponding relational symbols. If it is not stated otherwise we denote the domain of the structure under consideration by D. A signature is denoted by τ or σ. For a structure Γ over τ and a relational symbol $R \in \tau$ we write R^τ to denote the relation in Γ corresponding to R. Let σ and τ be signatures with $\sigma \subseteq \tau$. When Δ is a σ-structure and Γ is a τ-structure with the same domain such that $R^\Delta = R^\Gamma$ for all $R \in \sigma$, then Γ is called an *expansion* of Δ. By Γ_Δ we denote the expansion of Γ by relations in Δ.

If Γ is a relational structure over a finite signature τ, then we write arity(Γ) or arity(τ) to denote the largest arity of a relational symbol occurring in τ. For an n-ary relation R on D we write $\overline{R}$ to denote the relation $D^n \setminus R$ and for a relational structure Γ the notation $\overline{\Gamma}$ indicates a structure obtained from Γ by replacing every R in Γ by $\overline{R}$.

We say that a relational structure Γ is *first-order (fo-) definable* in Δ (is a *first-order reduct* of Δ) if Γ has the same domain as Δ, and for every relation R of Γ there is a first-order formula ϕ in the signature of Δ such that ϕ holds exactly on those tuples that are contained in R. A *primitive-positive (pp-) formula* over Δ is a first-order formula built exclusively from conjunction, existential quantifiers, atomic formulas over the signature of Δ and equalities: $(x = y)$. A structure Γ is pp-definable in Δ if it is fo-definable by a pp-formula.

We say that a countably infinite structure is ω*-categorical* if all countable models of its first-order theory are isomorphic. All the structures considered in this paper are countably infinite and ω-categorical. A relational structure Γ is *homogeneous* (Macpherson 2011) if every isomorphism between finite substructures may be extended to an automorphism of Γ.

The set of automorphisms of a structure Γ is denoted by $Aut(\Gamma)$. It is closed under composition and hence it can be also viewed as the group of automorphisms of Γ. Let n be a natural number. An orbit of an n-tuple $(a_1, \ldots, a_n)$, with elements in D, under $Aut(\Gamma)$ is a set of the form $\mathrm{Orb} = \{(\alpha(a_1), \ldots, \alpha(a_n)) \in D^n \mid \alpha \in Aut(\Gamma)\}$. We say that $Aut(\Gamma)$ is *oligomorphic* if there are only finitely many orbits of n-tuples under $Aut(\Gamma)$ for every $n \in \mathbb{N}$.

We now present a classical result on ω-categoricity.

Theorem 2 *(Engeler, Ryll-Nardzewski, Svenonius) For a countably infinite structure Γ with countable signature, the following are equivalent:*

1. *Γ is ω-categorical;*
2. *$Aut(\Gamma)$ is oligomorphic.*

Representation of Relations

In this paper whenever we consider an instance of the CSP or the abduction problem we always assume that all involved relations have a first-order definition in some countably infinite ω-categorical structure Δ. Since these structures are subject to computation a natural question of a finite representation of these relations arises. The solution to this problem is suggested by Theorem 2. Indeed, every n-ary relation R with a first-order definition in Δ is the union of orbits of n-tuples under $Aut(\Delta)$. By Theorem 2, it follows that R is the union of the finite number of orbits under $Aut(\Delta)$. Thus R can be represented in the finite way by the set S of representatives of orbits under $Aut(\Delta)$ consisting of one representative for each orbit contained in R. In this case we say that S is a *representation* of R wrt. Δ and that R is *represented* wrt. Δ. See also an example below.

Example 3 *A relation $R = \{(x, y, z) \in \mathbb{Q}^3 \mid (x < y < z) \vee (x > y > z)\}$ with a first-order definition in $(\mathbb{Q}; <)$ consists exactly of two orbits of triples under $\mathrm{Aut}(\mathbb{Q}; <)$, namely: $Orb_1 = \{(x, y, z) \in \mathbb{Q}^3 \mid (x < y < z)\}$ and $Orb_2 = \{(x, y, z) \in \mathbb{Q}^3 \mid (x > y > z)\}$. Now, since $(0, 1, 2) \in Orb_1$ and $(2, 1, 0) \in Orb_2$, it follows that the set $S = \{(0, 1, 2), (2, 1, 0)\}$ is a representation of R wrt. $(\mathbb{Q}; <)$.*

Abduction Problems

The usual definition of the *Constraint Satisfaction Problem (CSP)* goes as follows.

Definition 4 *An instance of the CSP is a triple $P = (V, D, C)$ with*

- *$V = \{v_1, \ldots, v_n\}$, set of variables, for some $n \in \mathbb{N}$,*
- *D, a non-empty set (domain),*
- *C, a set of constraints $C_1, \ldots, C_m$ over V and D, where each C_i is a pair (s_i, R_i), with*
  - *s_i, a tuple of variables from V of lenght m_i, the scope of C_i, and*
  - *R_i an m_i-ary relation over D, called the constraint relation of C_i.*

  Given P the question is as follows.
- *Is there a solution to P, i.e., is there a function $f : V \to D$ such that for each $i \leq m$, the m_i-tuple $f(s_i) \in R_i$?*

For an instance P we write $\mathrm{Sol}(P)$ to denote the relation consisting of tuples $(f(v_1), \ldots, f(v_n))$ such that f is a solution to P.

Due to intractability of CSP, which is in general NP-hard, one considers the following parametrization of the problem.

Definition 5 *Let Γ be a relational structure. Then $\mathrm{CSP}(\Gamma)$ is the restriction of the CSP from Definition 4 to instances where all constraints are built from relations in Γ.*

In this paper we study the abduction problem where a knowledge base as well as a set of hypotheses are given by sets of constraints; and a manifestation is a single constraint. We look for an explanation, that is, a subset of hypotheses consistent with the knowledge base (Condition 1 in the definition below) which together with the knowledge base entails the manifestation (Condition 2). Now comes the formal definition.

For a constraint $C = (s, R)$, the notation $\overline{C}$ indicates $(s, \overline{R})$.

Definition 6 *An instance of the* abduction problem *is a tuple $P = (V, V_H, D, KB, Hyp, M)$ with*

- *$V = \{v_1, \ldots, v_n\}$, set of variables, for some $n \in \mathbb{N}$,*
- *a subset V_H of V upon which explanations are built,*
- *D, a non-empty set (domain),*
- *and two sets of constraints given as in Definition 4:*
  - *KB over D and V, (knowledge base),*
  - *Hyp over D and V_H, (hypotheses),*
- *and a single constraint M over D and V, (manifestation).*

Given P the question is whether there is an explanation *for P, i.e., a set of constraints $H \subseteq Hyp$ such that:*

1. *$\mathrm{CSP}(V, D, KB \cup H)$ is satisfiable, and*
2. *$\mathrm{CSP}(V, D, KB \cup H \cup \overline{M})$ is not satisfiable.*

The abduction problem is in general Σ_2^P-hard. Thus, it makes sense to consider parametrizations of this problem. The most natural one is in our opinion the following.

Definition 7 *We define $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ to be the restriction of the abduction problem from Definition 6 to instances such that*

- *KB is any set of constraints built upon relations in Γ and variables in V,*
- *Hyp is any set of constraints built upon relations in $\mathcal{HYP}$ and variables in V_H, and*
- *M is a single constraint built upon a relation in $\mathcal{M}$ and variables in V.*

In this paper we consider a more restrictive version of the abduction problem defined in the following.

Definition 8 *Let Δ be an ω-categorical and homogeneous relational structure over a finite signature. By Δ_{Orb} we denote the structure over the domain of Δ having as relations all orbits of tuples under $\mathrm{Aut}(\Delta)$ of arity at most arity(Δ).*

Example 9 *Let $\Delta = (\mathbb{Q}; <)$. Then arity$(\Delta) = 2$ and $\Delta_{Orb} = (\mathbb{Q}; \mathbb{Q}, <, >, =)$, where the first occurrence of $\mathbb{Q}$ states for the domain of Δ_{Orb} and the second one for the trivial unary relation having as elements all rationals.*

We now define the restriction of the problem $\mathrm{Abduction}(\Gamma, \mathcal{HYP}, \mathcal{M})$ that we call $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ and which depends on three relational structures: $\Delta, \Gamma, \mathcal{M}$ over a finite signature such that Δ is ω-categorical, homogeneous, and Γ and $\mathcal{M}$ are first-order reducts of Δ. We will assume that all relations occurring in instances of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ are represented wrt. Δ. In the name of this problem OE stands for 'orbit explanation'. Indeed, an explanation for an instance of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ is built from constraints formed upon relations in Δ_{Orb} and variables in V_H. In fact in every instance of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ all such constraints are also available to build an explanation.

Definition 10 *Let Δ be homogeneous and ω-categorical structure and $\Gamma, \mathcal{M}$ be first-order reducts of Δ. We define $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ to be the abduction problem from Definition 6, where in every instance $P = (V, V_H, D, KB, Hyp, M)$:*

- *constraints KB and M are as in Definition 7, and*
- *Hyp is a set of constraints that for every l, every l-tuple s of variables in V_H, and every l-ary relation R in Δ_{Orb} contains a constraint (s, R).*

Without loss of generality we will always assume that $V_H = \{v_1, \ldots, v_a\}$ for some $a \leq n$.

Recall that Δ, Γ, and $\mathcal{M}$ are over finite signatures. We define the size of an instance P of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ to be the number of constraints in KB.

Example 11 *Let $R_g = \{(x, y, z) \in \mathbb{Q}^3 \mid (x \neq y \vee z > x)\}$ and $R_l = \{(x, y, z) \in \mathbb{Q}^3 \mid (x \neq y \vee z < x)\}$. Consider an instance of the abduction problem $P = (V, V_H, D, KB, Hyp, M)$ where:*

- $V = \{v_1, v_2, v_3, v_4, v_5\}$,
- $V_H = \{v_1, v_2, v_3\}$,
- $D = \mathbb{Q}$,
- $KB = \{((v_1, v_2, v_4), R_g), ((v_1, v_3, v_5), R_l)\}$,
- $Hyp = \{(v_1, \mathbb{Q}), (v_2, \mathbb{Q}), (v_3, \mathbb{Q}),$
  $((v_1, v_2), <), ((v_1, v_3), <), ((v_2, v_3), <),$
  $((v_1, v_2), >), ((v_1, v_3), >), ((v_2, v_3), >),$
  $((v_1, v_2), =), ((v_1, v_3), =), ((v_2, v_3), =)\}$,
- $M = ((v_4, v_5), >)$.

It is easy to verify that $H = \{\{(v_1, v_2), =\}, \{(v_1, v_3), =\}\}$ is an explanation for P.

Observe that P may be also seen as an instance of $\text{OE-ABD}_\Delta(\Gamma, \Delta_{Orb})$ where $\Delta = (\mathbb{Q}; <)$ and Δ_{Orb} is as defined in Example 9.

Observe also that if $\mathcal{HYP}$ contained any constraint less, then P would not be an instance of $\text{OE-ABD}_\Delta(\Gamma, \Delta_{Orb})$.

We now prove that using the relations in Δ_{Orb} we can pp-define all orbits of n-tuples under $Aut(\Delta)$, for every n.

Proposition 12 *Let Δ and Δ_{Orb} be as in Definition 8, n be any natural number and $(a_1, \ldots, a_n)$ a tuple over domain of Δ. Then there is a set of constraints built over variables $\{v_1, \ldots, v_n\}$ and relations in Δ_{Orb} such that $\mathrm{Sol}(V, D, C)$ is equal to an orbit of $(a_1, \ldots, a_n)$ under $Aut(\Delta)$.*

Local-to-Global Consistency

Recall the definition of a solution to an instance P of the CSP given in Definition 4. A *partial solution* to an instance $P = (V, D, C)$ is a mapping h from a subset V' of V to D such that for every constraint $C_i = (s_i, R_i)$ there is an extension $f : V \to D$ of h such that $f(s_i) \in R_i$.

For $P = (V, D, C)$ and $V' \subseteq V$, we write $\mathrm{PartSol}(V', P)$ with $V' = \{v_{i_1}, \ldots, v_{i_l}\}$ such that $i_1 < \cdots < i_l$ to denote the relation consisting of all tuples $(f(v_{i_1}), \ldots, f(v_{i_l}))$ such that $f : V' \to D$ is a partial solution to P.

Definition 13 *Let $P = (V, D, C)$ be an instance of the CSP. We say that an instance P is* k-consistent *if for every partial solution $h : V' \to D$ to P with $V' \subseteq V$ and $|V'| = (k-1)$, and $v \in V \setminus V'$ there is a partial solution $f : V' \cup \{v\} \to D$ to P extending h. An instance $P = (V, D, C)$ of the CSP is* strongly k-consistent *if it is l-consistent for every $l \leq k$ and it is* globally consistent *if it is strongly $|V|$-consistent.*

Let $P = (V, D, C)$ be an instance of the CSP. We say that $P' = (V, D, C')$ is a k-consistent, strongly k-consistent or globally consistent variant of P if P' has the same set of solutions as P and is k-consistent, strongly k-consistent or globally consistent, respectively. In (Bodirsky and Dalmau 2013), it was proved that if Γ is ω-categorical and over a finite relational signature, then for every k there is a polynomial time algorithm that computes a strongly k-consistent variant of a given instance P of $\mathrm{CSP}(\Gamma)$.

Definition 14 *We say that a relational structure Γ has* local-to-global consistency wrt. k *if every strongly k-consistent variant P' of every instance P of $\mathrm{CSP}(\Gamma)$ is also globally consistent. We say that Γ has* local-to-global consistency *if it has local-to-global consistency wrt. some k.*

Local-to-Global consistency for ω-categorical structures has also two more known characterizations. One is in terms of so-called polymorphisms, see (Bodirsky and Dalmau 2013). The other one, more suitable for our purposes is stated in terms of decomposability.

Definition 15 *We say that a constraint language Γ is* k-decomposable *if every primitive positive formula in Γ is equivalent to a conjunction of at most k-ary primitive positive formulas.*

The following result follows from Corollary 3 in (Bodirsky and Dalmau 2013) and Theorem 19 in (Bodirsky and Chen 2007).

Theorem 16 *Let Γ be an ω-categorical relational structure. Then Γ has local-to-global consistency wrt. k if and only if Γ is $(k-1)$-decomposable.*

We now translate Theorem 16 to the language of constraints.

Proposition 17 *Let Γ be an ω-categorical structure with local-to-global consistency wrt. k and $P = (\{v_1, \ldots, v_n\}, D, C)$ be an instance of $\mathrm{CSP}(\Gamma)$. Then $t \in \mathrm{Sol}(P)$ if and only if for every subset $\{v_{i_1}, \ldots, v_{i_l}\}$ of V with $l < k$ it holds that $\mathrm{proj}_{\{i_1, \ldots, i_l\}}(t) \in \mathrm{PartSol}(\{v_{i_1}, \ldots, v_{i_l}\}, P)$.*

The above characterization of local-to-global consistency suggest the following representation, called the k-*decomposition*, of $\mathrm{Sol}(P)$ for an instance P of the CSP.

Definition 18 *Let $P = (V, D, C)$ with $V = \{v_1, \ldots, v_n\}$ be an instance of the CSP. Then the k-*decomposition *of $\mathrm{Sol}(P)$ is the instance $P' = (V, D, C')$ of the CSP such that for every $V' \subseteq V$ with $|V'| < k$ the set C' contains a constraint of the form*

$$((v_{i_1}, \ldots, v_{i_l}), \mathrm{PartSol}(V', P)),$$

where $V' = \{v_{i_1}, \ldots, v_{i_l}\}$.

The k-decomposition of $\mathrm{Sol}(P)$ for an instance P of $\mathrm{CSP}(\Gamma)$ is of polynomial size and can be obtained in polynomial time wrt. size of P.

Proposition 19 *Let Δ be an ω-categorical structure and k a fixed natural number. Then there exists an algorithm that for an instance P of the CSP such that all involved relations have a first-order definition in Δ computes the k-decomposition of $\mathrm{Sol}(P)$ in which all relations are represented wrt. Δ. The algorithm works in time polynomial wrt. the size of P.*

Algorithm

This section is devoted to present our main tractability result which simply follows from Theorem 23 proved below.

Theorem 20 *Let Δ be an ω-categorical and homogeneous structure and $\Gamma, \mathcal{M}$ first-order reducts of Δ such that $\Gamma_{\overline{\mathcal{M}}}$ has local-to-global consistency. Then $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ is decidable in polynomial time.*

Let $P = (V, V_H, D, \mathrm{KB}, \mathrm{Hyp}, M)$ be an instance of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$. As we will show, the question of whether there is an explanation for P is equivalent to the question of whether there is a partial assignment $a_p : V_H \to D$ that can be extended to $a : V \to D$ so that a satisifies KB but cannot be extended to any $a' : V \to D$ that satifies $\mathrm{KB} \cup \overline{M}$. This is exactly Condition (1) in Proposition 21. This proposition is inspired by Proposition 1 in (Zanuttini 2003) where a similar observation has been made in the case of propositional abduction. We refer to this condition as to *Zanuttini's condition*.

Proposition 21 *Let Δ be an ω-categorical and homogeneous structure and $\Gamma, \mathcal{M}$ be first-order reducts of Δ. Let $P = (V, V_H, D, KB, Hyp, M)$ be an instance of $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$. Let $P^{proj}_{KB} = \mathrm{PartSol}(V_H, (V, D, KB))$ and $P^{proj}_{\overline{M}} = \mathrm{PartSol}(V_H, (V, D, KB \cup \overline{M}))$. Then $P \in \text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ if and only if*

$$P^{proj}_{KB} \not\subseteq P^{proj}_{\overline{M}}. \qquad (1)$$

As an Algorithm 1, we present the procedure `LtG-OEAbd` that solves the problem $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$ from Theorem 20. The algorithm calls the following sub procedures.

- `(k+1)-DecompAlg`$_\Delta(P)$ that for a given instance P of the CSP such that all involved relations have a first-order definition in Δ returns the $(k+1)$-decomposition of $\mathrm{Sol}(P)$ such that all relations in P are represented wrt. Δ. By Proposition 19, we have that the procedure `(k+1)-DecompAlg`$_\Delta$ exists and that we can assume that it works in polynomial time wrt. the size of P.
- `ProjConstr`$(V', (V, D, C))$ returns (V', D, C') such that C' is a subset of C containing all constraints whose scope is completely included in V'.

We will need the following auxiliary result.

Lemma 22 *Let $P = (V, D, C)$ be an instance of the CSP, P' the k-decomposition of $\mathrm{Sol}(P)$, and $V_H = \{v_1, \ldots, v_a\}$ for some $a \leq n$. Then `ProjConstr`(V_H, P') is the k-decomposition of $\mathrm{PartSol}(V_H, P')$.*

Algorithm 1: Algorithm `LtG-OEAbd`

// Δ - an ω-categorical homogeneous structure and
// $\Gamma, \mathcal{M}$ - fo-reducts of Δ such that $\Gamma_{\overline{\mathcal{M}}}$ has local-to-global consistency wrt. k.
Data: An instance $P = (V, V_H, D, \mathrm{KB}, \mathrm{Hyp}, M)$ of the problem $\text{OE-ABD}_\Delta(\Gamma, \mathcal{M})$.
Result: `True` if there is an explanation for P, and `False` otherwise.
// Computation of $(k+1)$-decompositions
1 $P^{\text{kdec}}_{\text{KB}}$:= `(k+1)-DecompAlg`$_\Delta(V, D, \mathrm{KB})$;
2 $P^{\text{kdec proj}}_{\text{KB}}$:= `ProjConstr`$(V_H, P^{\text{kdec}}_{\text{KB}})$;
3 $P^{\text{kdec}}_{\overline{M}}$:= `(k+1)-DecompAlg`$_\Delta(V, D, \mathrm{KB} \cup \overline{M})$;
4 $P^{\text{kdec proj}}_{\overline{M}}$:= `ProjConstr`$(V_H, P^{\text{kdec}}_{\overline{M}})$;
// Checking Zanuttini's condition
5 **forall the** $\{v_{i_1}, \ldots, v_{i_l}\} \subseteq V_H$ *with* $l \leq k$ **do**
6   Let $((v_{i_1}, \ldots, v_{i_l}), R_1) \in P^{\text{kdec proj}}_{\text{KB}}$
7   and $((v_{i_1}, \ldots, v_{i_l}), R_2) \in P^{\text{kdec proj}}_{\overline{M}}$;
8   **if** $R_1 \not\subseteq R_2$ **then**
9   **return** `True`
10   **end**
11 **end**
12 **return** `False`

In a nutshell, the algorithm first stores in the variable $P^{\text{kdec}}_{\text{KB}}$ in Line 1 the $(k+1)$-decomposition of $\mathrm{Sol}((V, D, \mathrm{KB}))$. By Lemma 22, the variable $P^{\text{kdec proj}}_{\text{KB}}$ in Line 2 stores the $(k+1)$-decomposition of $\mathrm{PartSol}(V_H, (V, D, \mathrm{KB}))$. In the same way, we argue that $P^{\text{kdec proj}}_{\overline{M}}$ in Line 4 stores the $(k+1)$-decomposition of $\mathrm{PartSol}(V_H, (V, D, \mathrm{KB} \cup \overline{M}))$. By Proposition 21, it is now enough to argue that the loop checks Condition (1). On one hand, if for all $\{v_{i_1}, \ldots, v_{i_l}\} \subseteq V_H$ with $l \leq k$ we have that $R_1 \subseteq R_2$, then by Proposition 17, it follows that $P^{\text{proj}}_{\text{KB}} \subseteq P^{\text{proj}}_{\overline{M}}$. On the other hand, if there is $\{v_{i_1}, \ldots, v_{i_l}\} \subseteq V_H$ with $l \leq k$ and t' in R_1 but not in R_2, then t' can be extended to $t \in (\mathrm{PartSol}(V_H, (V, D, \mathrm{KB})) \setminus \mathrm{PartSol}(V_H, (V, D, \mathrm{KB} \cup \overline{M})))$. Indeed, it follows since: (1) $\mathrm{Sol}(P^{\text{kdec}}_{\text{KB}}) = \mathrm{Sol}((V, D, \mathrm{KB}))$, (2) $P^{\text{kdec}}_{\text{KB}}$ is strongly k-consistent, and (3) $\Gamma_{\overline{\mathcal{M}}}$ has local-to-global consistency wrt. k.

The polynomiality of the algorithm follows by the observations that calls to procedures `(k+1)-DecompAlg`$_\Delta$ and `ProjConstr` works in polynomial time and that the loop is launched $O(n^k)$ times, each of which takes constant time.

Theorem 23 *Let Δ be an ω-categorical homogeneous structure and $\Gamma, \mathcal{M}$ be first-order reducts of Δ such that $\Gamma_{\overline{\mathcal{M}}}$ has local to global consistency wrt. k. Then $P = (V, V_H, D, KB, Hyp, M)$ is a positive instance of* OE-ABD$_\Delta(\Gamma, \mathcal{M})$ *if and only if the algorithm* `LtG-OEAbd` *returns* `True`*. The algorithm* `LtG-OEAbd` *works in time polynomial wrt. the size of P.*

Applicability of the Algorithm LtG-OEAbd

Temporal Point-Based Abduction

We first apply our algorithm to temporal point-based abduction which may be seen as abduction in the context of point algebra. In this case, we fix $\Delta = (\mathbb{Q}; <)$, which is a very well-known ω-categorical and homogeneous structure. Recall from Example 9 that $\Delta_{\text{Orb}} = (\mathbb{Q}; \mathbb{Q}, <, >, =)$. We consider now the problem OE-ABD$_\Delta(\Gamma, \mathcal{M})$ where Γ is so-called *ORD-Horn (OH) language* (Nebel and Bürckert 1995), a special kind of structure with a first-order definition in $(\mathbb{Q}; <)$ and $\mathcal{M} = (\mathbb{Q}; <, >, \leq, \geq, =, \neq)$.

We say that Γ is an OH language if every relation in Γ can be defined as a conjunction of clauses of one of the following forms: (1) $((x_1 \neq y_1 \vee \cdots \vee x_k \neq y_k) \vee (x \lhd y))$, (2) $(x_1 \neq y_1 \vee \cdots \vee x_k \neq y_k)$, or (3) $(x \lhd y)$, where $\lhd \in \{<, \leq, =\}$.

In Theorem 1 in (Wrona 2012) it was shown that an OH language Γ over a finite signature has local-to-global consistency if and only if Γ is *basic OH* that is every relation in Γ can be defined as a conjunction of clauses of one of the following forms: (1) $((x_1 \neq x_2 \vee \cdots \vee x_1 \neq x_k) \vee (y_1 \neq y_2 \vee \cdots \vee y_1 \neq y_l) \vee (x_1 < y_1))$, (2) $(x_1 \neq y_1 \vee \cdots \vee x_k \neq y_k)$, or (3) $(x \lhd y)$, where $\lhd \in \{<, \leq, =\}$.

Since for every basic OH language Γ we have that $\Gamma_{\overline{\mathcal{M}}}$ is also basic OH, by Theorem 1 in (Wrona 2012) and Theorem 20 above we obtain the following non-trivial result on the complexity of temporal abduction.

Theorem 24 *Let $\Delta = (\mathbb{Q}; <)$, Γ be any basic OH language, and $\mathcal{M} = (\mathbb{Q}; <, >, \leq, \geq, =, \neq)$. Then* OE-ABD$_\Delta(\Gamma, \mathcal{M})$ *is decidable in polynomial time by algorithm* `LtG-OEAbd`*.*

Equality Abduction Problem

To the very best of our knowledge, the only complete complexity classification of the problem OE-ABD$_\Delta(\Gamma, \mathcal{M})$ has been obtained in (Schmidt and Wrona 2013) for so-called equality languages Γ, that is, first-order reducts of $(\mathbb{N}; =)$ where $\mathbb{N}$ is the set of natural numbers. In this case $\Delta = (\mathbb{N}; =)$ and $\mathcal{M} = (\mathbb{N}; =, \neq)$. Depending on Γ it was shown that the problem OE-ABD$_\Delta(\Gamma, \mathcal{M})$ is in P, it is NP-complete, or Σ_2^P-complete.

The algorithm for the polynomial case presented along with the classification is unfortunately somewhat barehanded and specialized to equality languages only. But this problem is in P when every relation in Γ can be defined as a conjunction of clauses of the form $(x_1 \neq y_1 \vee \cdots \vee x_k \neq y_k)$, that is, when Γ is basic OH. By Theorem 11 in (Schmidt and Wrona 2013) and Theorem 24 it follows that the polynomial case in the classification of the equality abduction problem may be handled by procedure `LtG-OEAbd`.

Theorem 25 *Let $\Delta = (\mathbb{N}; =)$, Γ be an equality language, and $\mathcal{M} = (\mathbb{N}; =, \neq)$. Then, if Γ is basic OH, then* OE-ABD$_\Delta(\Gamma, \mathcal{M})$ *may be solved in P by the procedure* `LtG-OEAbd`*. Otherwise, it is NP-hard.*

Abduction and Allen's Interval Algebra

Allen's interval algebra is a formalism introduced for temporal reasoning in Artificial Intelligence (Allen 1983). Consider now a structure $\mathbb{B}$ whose domain $\mathbb{I}$ are the pairs $(u, v) \in \mathbb{Q}^2$ with $(u < v)$ and relations are so-called *basic relations* of Allen's interval algebra: P, M, O, S, F, D, E whose names stands for precedes, meets, overlaps with, starts, finishes, is during, and equals and are defined in a natural way (Bodirsky 2012). For instance, $P = \{((a, b), (c, d)) \mid b < c\}$ and $M = \{((a, b), (c, d)) \mid b = c\}$. Now the network satisfaction problem for Allen's interval algebra can be viewed as CSP$(\mathbb{A})$ where $\mathbb{A}$ is a first-order reduct of $\mathbb{B}$ that contains all binary relations first-order definable in $\mathbb{B}$. This problem and hence also OE-ABD$_{\mathbb{B}}(\mathbb{A}, \mathbb{A})$ is NP-hard. Consider now a *pointizable fragment* Γ^{PIA} of Allen's interval algebra (van Beek and Cohen 1990) which consists of all binary relations over intervals that can be defined as a 4-ary relation over $(\mathbb{Q}; <)$ by conjunctions of atomic formulas of the form $(x_1 < x_2), (x_1 = x_2), (x_1 \neq x_2)$ and $(x_1 \leq x_2)$. Observe that Γ^{PIA} contains all relations in $\mathbb{B}$. Theorem 40 in (Bodirsky and Chen 2009) shows that Γ^{PIA} has local-to-global consistency wrt. 5. By this result and Theorem 20 we have what follows.

Theorem 26 *Let $\mathbb{B}$ and Γ^{PIA} be as specified above. Then* OE-ABD$_{\mathbb{B}}(\Gamma^{PIA}, \overline{\Gamma^{PIA}})$ *is decidable in polynomial time by the procedure* `LtG-OEAbd`*.*

Abduction in Spatial Reasoning

The last example considers spatial reasoning in the context of the very well known formalism RCC-5 which operates on sets and the following basic binary relations: DR, PO, PP, PPI, and EQ stands for disjointness, proper overlap, proper containment, its inverse and equality, respectively. It was shown in (Bodirsky and Chen 2009) that the constraint satisfaction problem involving all these five relations may be defined as CSP$(\mathbb{B}_0)$ for certain ω-categorical and homogeneous structure $\mathbb{B}_0$. By Theorem 33 in that paper we have that a structure Γ which is an expansion of $\mathbb{B}_0$ by a finite number of relations each of which is definable as a conjunction of clauses of the form $(x_1 \neq y_1 \vee \cdots \vee x_k \neq y_k)$ has local-to-global consistency.

By Theorem 20, we obtain the following result.

Theorem 27 *Let Γ be an expansion of $\mathbb{B}_0$ by a finite number of relations each of which is definable by a conjunction of disjunction of disequalities. Then the problem* OE-ABD$(\Gamma, \overline{\mathbb{B}_0})$ *is decidable in polynomial time by the procedure* `LtG-OEAbd`*.*

Acknowledgements

The author is supported by the *Swedish Research Council* (VR) under grant 621-2012-3239.

References

Allen, J. F. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM* 26(11):832–843.

Bennett, B. 1994. Spatial reasoning with propositional logics. In *Proceedings of the International Conference on Knowledge Representation and Reasoning*, 51–62. Morgan Kaufmann.

Bodirsky, M., and Chen, H. 2007. Oligomorphic clones. *Algebra Universalis* 57(1):109–125.

Bodirsky, M., and Chen, H. 2009. Qualitative temporal and spatial reasoning revisited. *Journal of Logic and Computation* 19(6):1359–1383.

Bodirsky, M., and Dalmau, V. 2013. Datalog and constraint satisfaction with infinite templates. *J. Comput. Syst. Sci.* 79(1):79–100.

Bodirsky, M., and Hils, M. 2012. Tractable set constraints. *Journal of Artificial Intelligence Research* 45:731–759.

Bodirsky, M., and Kára, J. 2009. The complexity of temporal constraint satisfaction problems. *Journal of the ACM* 57(2). An extended abstract appeared in the proceedings of STOC'08.

Bodirsky, M., and Kára, J. 2010. A fast algorithm and Datalog inexpressibility for temporal reasoning. *ACM Transactions on Computational Logic* 11(3).

Bodirsky, M. 2012. Complexity classification in infinite-domain constraint satisfaction. Memoire d'habilitation à diriger des recherches, Université Diderot – Paris 7. Available at arXiv:1201.0856.

Brusoni, V.; Console, L.; Terenziani, P.; and Pernici, B. 1997. Later: Managing temporal information efficiently. *IEEE Expert* 12(4):56–64.

Brusoni, V.; Console, L.; Terenziani, P.; and Dupré, D. T. 1998. A spectrum of definitions for temporal model-based diagnosis. *Artif. Intell.* 102(1):39–79.

Bylander, T.; Allemang, D.; Tanner, M. C.; and Josephson, J. R. 1991. The computational complexity of abduction. *Artif. Intell.* 49(1-3):25–60.

Chen, H., and Wrona, M. 2012. Guarded ord-horn: A tractable fragment of quantified constraint satisfaction. In *TIME*, 99–106.

Console, L.; Terenziani, P.; and Dupré, D. T. 2002. Local reasoning and knowledge compilation for efficient temporal abduction. *IEEE Trans. Knowl. Data Eng.* 14(6):1230–1248.

Creignou, N., and Zanuttini, B. 2006. A complete classification of the complexity of propositional abduction. *SIAM J. Comput.* 36(1):207–229.

Dechter, R. 1992. From local to global consistency. *Artificial Intelligence* 55(1):87–108.

Duentsch, I. 2005. Relation algebras and their application in temporal and spatial reasoning. *Artificial Intelligence Review* 23:315–357.

Eiter, T., and Gottlob, G. 1995. The complexity of logic-based abduction. *J. ACM* 42(1):3–42.

Fisher, M.; Gabbay, D.; and Vila, L., eds. 2005. *Handbook of Temporal Reasoning in Artificial Intelligence*. Elsevier.

Freuder, E. 1982. A sufficient condition for backtrack-free search. *Journal of the ACM* 29(1):24–32.

Herzig, A.; Lang, J.; and Marquis, P. 2001. Planning as abduction. In *IJCAI-01 Workshop on Planning under Uncertainty and Incomplete Information*.

Hirsch, R. 1997. Expressive power and complexity in algebraic logic. *Journal of Logic and Computation* 7(3):309 – 351.

Hobbs, J. R.; Stickel, M. E.; Appelt, D. E.; and Martin, P. A. 1993. Interpretation as abduction. *Artif. Intell.* 63(1-2):69–142.

Hodges, W. 1993. *Model theory*. Cambridge University Press.

Mackworth, A. K. 1977. Consistency in networks of relations. *AI* 8:99–118.

Macpherson, D. 2011. A survey of homogeneous structures. *Discrete Mathematics* 311(15):1599–1634.

Marquis, P. 2000. Consequence finding algorithms. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems (DRUMS), Vol. 5*, 41–145. Kluwer Academic.

Nebel, B., and Bürckert, H.-J. 1995. Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra. *Journal of the ACM* 42(1):43–66.

Nordh, G., and Zanuttini, B. 2008. What makes propositional abduction tractable. *Artif. Intell.* 172(10):1245–1284.

Pople, H. E. 1973. On the mechanization of abductive logic. In *IJCAI*, 147–152.

Renz, J., and Nebel, B. 2007. Qualitative spatial reasoning using constraint calculi. In Aiello, M.; Pratt-Hartmann, I.; and van Benthem, J., eds., *Handbook of Spatial Logics*. Springer Verlag, Berlin. 161–215.

Schaefer, T. J. 1978. The complexity of satisfiability problems. In *Proceedings of the Symposium on Theory of Computing (STOC)*, 216–226.

Schmidt, J., and Wrona, M. 2013. The complexity of abduction for equality constraint languages. In *CSL*, 615–633.

van Beek, P., and Cohen, R. 1990. Exact and approximate reasoning about temporal relations. *Computational Intelligence* 6:132–144.

Vilain, M.; Kautz, H.; and van Beek, P. 1989. Constraint propagation algorithms for temporal reasoning: A revised report. *Reading in Qualitative Reasoning about Physical Systems* 373–381.

Wrona, M. 2012. Syntactically characterizing local-to-global consistency in ord-horn. In *CP*, 704–719.

Zanuttini, B. 2003. New polynomial classes for logic-based abduction. *J. Artif. Intell. Res. (JAIR)* 19:1–10.