Solving first-order constraints over the monadic class

Dmitri Chubarov and Andrei Voronkov

Department of Computer Science,
University of Manchester, Manchester, M13 9PL
{dchubarov,voronkov}@cs.man.ac.uk

Abstract. First-order constraints over arbitrary theories or structures can be formalised as the *formula instantiation problem* as defined in [9]. Several results have been previously obtained for the formula instantiation problem in the case of quantifier-free formulas of first-order logic. In this paper we prove the first general result on formula instantiation for quantified formulas, namely that formula instantiation is decidable for the monadic class without equality.

1 Introduction

Consider a first order sentence φ which contains an existentially quantified variable x. If the sentence is valid, it can be interpreted as a definition of an element (not necessarily unique) which can be found in any model of the corresponding signature. It is not always possible to define such an element explicitly as a term t of the signature. We call t a *witness* for x in φ. Given a sentence and a designated existential variable in it, we study the question if it has a witness. This question gives rise to a decision problem which is called *formula instantiation*, following [9, 11].

Definition 1 (formula instantiation). Formula instantiation *is the following decision problem: given a formula $\varphi(\bar{x})$ does there exist a sequence of terms $\bar{t}$ in the signature Σ_φ of the formula such that $\varphi(\bar{t})$ is valid. Every such tuple of terms $\bar{t}$ is called a* witness *for $\varphi(\bar{x})$.*

We can consider a broader definition when $\varphi(\bar{x})$ contains other free variables. We can define validity of $\varphi(\bar{t})$ if we bind the other variables with quantifiers. Given the quantification of free variables, this problem can be reduced to the problem as defined above.

In other words, formula instantiation holds for $\varphi(\bar{x})$ if and only if $\varphi(\bar{x})$ has a witness. In fact, formula instantiation is a family of problems parametrized by a class of formulas (possible values for $\varphi(\bar{x})$) and a theory with respect to which validity is defined. Naturally, we can also consider the dual problem of finding witness terms which make $\varphi(\bar{t})$ satisfiable.

Formula instantiation has applications in tableau-based automatic reasoning in first-order logic. It naturally arises when one wants to close a tableau (or a tableau branch) by an appropriate substitution. The tableau, or a tableau branch can be considered as a formula $\varphi(\bar{x})$. Then formula instantiation holds for the formula $\neg\varphi(\bar{x})$ if and only if the branch (or the whole tableau) is closed. This search for a substitution that closes a quantifier-free tableau was formalised as the simultaneous rigid E-unification problem in [2]. A number of results about simultaneous rigid E-unification were proved (see [1] for an overview). The problem of formula instantiation was formulated in [9], where several results about formula instantiation for quantifier-free formulas were also proved and several open problems related to formula instantiation were formulated.

Some of the results known so far are given in Table 1.

It is easy to see that the validity problem is a special case of formula instantiation, where the formula φ has no free variables. Therefore, it makes sense only to consider formula instantiation for classes with the

Class	Complexity
no equality	Σ_2^p-complete
equality, binary symbols	undecidable
equality, one variable	decidable and EXPTIME-hard
Monadic signature:	
equality, general	open
equality, at most one unary symbol	decidable
one variable	PSPACE-complete

Table 1. Complexity of formula instantiation for quantifier-free formulas

decidable validity problem. The decidability of the validity problem does not, however, guarantee that formula instantiation is decidable. Two extreme examples of hardness of formula instantiation are the following.

1. The validity problem for quantifier-free formulas with equality is coNP-complete, but formula instantiation for quantifier-free formulas is undecidable already for formulas with two variables [7].
2. The validity problem for arbitrary ground-negative formulas with equality is Π_p^2-complete [10], but formula instantiation is undecidable even for quantifier-free ground-negative formulas with two variables [7], see also [8].

Despite a large number of results on formula instantiation for quantifier-free formulas, no nontrivial general results are known for quantified formulas. In this paper we will consider formula instantiation for the monadic class.

Definition 2 (Monadic Class). *A* monadic signature *is a signature which contains no function or relation symbols of arity greater than 1. The* monadic class *is the class of all formulas whose signature is monadic.*

Note that formulas in the monadic class may have arbitrary quantifiers.

The main result of this paper is the following.

Theorem 1. *Formula instantiation problem for the monadic class is decidable.*

To prove the decidability of formula instantiation for the monadic class, for each instance φ we first skolemize out all universally quantified variables in φ to obtain a formula φ' in an extended signature such that φ and φ' have the same witnesses in the signature of φ. We skolemize in such a way that only new constants are introduced, but no function symbols, and hence the signature of φ' remains monadic. Then we note that the truth of φ' on a tuple of terms $\bar{t}$ can be expressed using a second-order formula of the term algebra in a monadic signature. Likewise, the existence of a witness in the original signature of φ can also be expressed by such a formula. then we use the decidability of the monadic second-order theory of any term algebra in a monadic signature, which follows from the decidability of the monadic second-order theory of two successors [6].

2 Skolemization of monadic formulas

A formula is called *closed*, or a *sentence* if it contains no free variables. We call a *literal* an atomic formula or its negation. A formula is in *negation normal form* if it is built from literals using the connectives $\wedge, \vee$ and quantifiers $\forall, \exists$. A formula is called *prenex* if it has the form $Q\bar{x}A$, where $Q\bar{x}$ is a quantifier prefix and A a quantifier-free formula.

Our first step is to skolemize out universal quantifiers.

Lemma 1. *Let $\varphi(\bar{x})$ be a formula of a monadic signature Σ with free variables among those in $\bar{x}$. There exists a prenex existential formula $\exists\bar{y}\varphi'(\bar{x},\bar{y})$, where $\varphi'(\bar{x},\bar{y})$ is quantifier-free such that (i) $\varphi'(\bar{x},\bar{y})$ uses symbols in Σ and maybe some new constants; (ii) for all tuples of ground terms $\bar{t}$ of the signature Σ the formula $\varphi(\bar{t})$ is valid if and only if $\exists\bar{y}\varphi'(\bar{t},\bar{y})$ is valid.*

Proof. Let φ be a monadic formula. Without loss of generality we assume that φ is in negation normal form. Repeatedly using the following equivalences to rewrite subformulas of φ:

$$\begin{aligned}
\forall x(A \wedge B) &\equiv \forall x A \wedge \forall x B;\\
\exists x(A \vee B) &\equiv \exists x A \vee \exists x B;\\
\exists x(A(x) \wedge C) &\equiv \exists x A(x) \wedge C;\\
\forall x(A(x) \vee C) &\equiv \forall x A(x) \vee C;\\
\forall x C &\equiv C;\\
\exists x C &\equiv C,
\end{aligned}$$

where x does not occur in C, we obtain a formula φ' equivalent to φ in which all quantifiers occur only in subformulas of the form $\exists x \bigwedge_i L_i(x)$

and $\forall x \bigvee_i L_i(x)$, where all of the L_i's are literals in which x occurs. Using the standard skolemization arguments, we can replace each subformula $\forall x \bigvee_i L_i(x)$ by $\bigvee_i L_i(c)$, where c is a new constant. Finally, we can pull the remaining existential quantifiers out to obtain the required prenex existential formula.

This lemma implies the following result.

Lemma 2. *Let φ be a formula of a monadic signature Σ. Using φ one can effectively find a prenex existential formula φ' of the same variables as φ such that*

1. *the signature of φ' extends the signature of φ by zero or more constants;*
2. *for every tuple of terms $\bar{t}$ of the signature Σ, $\bar{t}$ is a witness for φ if and only if it is a witness for φ'.*

3 Term algebras

Let Δ be a signature with at least one constant. We denote by T_Δ the set of all ground terms of the signature Δ. The *term algebra* of the signature Δ is the structure with the universe T_Δ such that every ground term of Δ is interpreted in this structure by itself, for details see, e.g., [3]. We will denote the term algebra of Δ in the same way as its universe T_Δ.

The term algebra is infinite whenever Δ contains at least one function symbol of positive arity. We are interested in monadic second-order theories of term algebras defined as follows.

Extend the first-order language by adding *second-order variables*. The formulas of the *Monadic second-order calculus* extend first-order formulas by adding (i) a new kind of atomic formulas of the form $Z(t)$, where Z is a second-order variable and t is a term, and (ii) *second-order quantifiers* $\forall Z$ and $\exists Z$, where Z is a second-order variable.

Let $\mathfrak{A}$ be a structure with a universe A. Over this structure second-order variables are interpreted as subsets of A and an atom $Z(t)$ is interpreted as membership of t in Z. *Monadic second-order theory* of $\mathfrak{A}$ is the set of all monadic second-order sentences which hold in $\mathfrak{A}$.

The decidability of the monadic second-order theory of two successors [6] implies the following lemma.

Lemma 3. *The monadic second-order theory of any term algebra in a monadic signature is decidable.*

Proof. Note that every formula of monadic signature can be encoded in a signature consisting of one constant a and two unary functions. Suppose that all the function symbols and constants that appear in formula φ are among functions $f_1 \dots f_m$ and constants $c_1, \dots c_m$. Introduce two new symbols h and g and substitute define the encoding as follows:

- substitute each occurrence of function symbol f_i by hg^i,
- substitute each occurrence of c_i by $h^i(a)$.

This encoding results in at most quadratic increase in the length of formula. Thus without loss of generality monadic second-order theory of term algebra of unary functions can be considered as a theory of binary tree instead of finitely branching forest.

4 Proof of the main theorem

Lemma 4. *For any monadic signature Σ and its extension $\Delta \supseteq \Sigma$ by constants there exists a monadic second-order formula $TERM_\Sigma(x)$ of one variable x, such that for every ground term t of Δ the formula $TERM_\Sigma(t)$ holds in the weak second-order theory of term algebra of Δ if and only if t is a term of the signature Σ.*

Proof. Let U be the set of all non-constant symbols of Σ (and hence also of Δ). Define the formula $TERM_\Sigma$ as follows.

$$TERM_\Sigma(x) = \exists Z \left(Z(x) \wedge \bigwedge_{f \in U} \forall y (Z(f(y)) \Rightarrow Z(y)) \wedge \bigwedge_{d \in \Delta - \Sigma} \neg Z(d) \right).$$

The proof is straightforward.

Evidently, this lemma can be proved for arbitrary signatures Σ and Δ.

Now we can prove the main theorem.

Proof (of Theorem 1). We prove that Formula Instantiation is decidable for the monadic class. Given a formula $\varphi(\bar{x})$ in a monadic signature Σ, we have to effectively decide whether it has a witness in the signature Σ. Let $P_1, \dots, P_n$ be all predicate symbols of Σ and $x_1, \dots, x_m$ be all variables in $\bar{x}$. Applying Lemma 2 we obtain an existential formula $\varphi'(\bar{x})$ of a signature Δ such that (i) the signature of $\varphi'(\bar{x})$ extends the signature of $\varphi(\bar{x})$ by zero or more constants; (ii) for every tuple of terms $\bar{t}$ of the signature Σ, $\bar{t}$ is a witness for $\varphi(\bar{x})$ if and only if it is a witness for $\varphi'(\bar{x})$.

Therefore, formula instantiation holds for $\varphi(\bar{x})$ if and only if $\varphi'(\bar{x})$ has a witness $\bar{t}$ in the signature Σ.

Since $\varphi'(\bar{t})$ is existential, by Herbrand's theorem it is valid if and only if it holds in all Herbrand interpretations of the signature Δ. But this property can be expressed using the following monadic second-order formula of T_Δ

$$\forall P_1 \dots \forall P_n \varphi'(\bar{t}).$$

where $P_1, \dots, P_n$ are considered as second-order variables.

It follows that $\varphi'(\bar{x})$ has a witness in the signature Σ if and only if the following second-order formula holds in T_Δ:

$$\exists x_1 \dots \exists x_m \left(\bigwedge_{i=1 \dots m} \mathit{TERM}_\Sigma(x_i) \wedge \forall P_1 \dots \forall P_n \varphi'(\bar{x}) \right).$$

The validity of this second-order formula can be effectively decided by Lemma 3.

5 Conclusion

Our proof uses the decidability of the monadic second-order theory of term algebras in monadic signatures. The complexity of such theories is not elementary recursive [5]. At the same time the satisfiability problem for formulas in the monadic class had been proved decidable by Löb and Gurevich independently and nondeterministic exponential time complexity was proved by Lewis [4] using the finite model property. It remains an open problem whether the complexity of formula instantiation for the monadic class is elementary.

Another interesting open question is whether existing resolution-based decision procedures can be modified to find witnesses for formula instantiation.

References

1. A. Degtyarev and A. Voronkov. Equality reasoning in sequent-based calculi. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, volume I, chapter 10, pages 609–704. Elsevier Science, 2001.
2. J.H. Gallier, S. Raatz, and W. Snyder. Theorem proving using rigid E-unification: Equational matings. In *Proc. IEEE Conference on Logic in Computer Science (LICS)*, pages 338–346. IEEE Computer Society Press, 1987.
3. W. Hodges. *Model theory.* Cambridge University Press, 1993.

4. H. Lewis. Complexity results for classes of quantificational formulas. *Journal of Computer and System Sciences*, 21:317–353, 1980.
5. A. Meyer. Weak monadic second order theory of successor is not elementary-recursive. In R. Parikh, editor, *Logic Colloquium: Symposium on Logic Held at Boston*, volume 453 of *Lecture Notes in Mathematics*, pages 132–154. Springer Verlag, 1975.
6. M. O. Rabin. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 141(1):1–35, 1969.
7. M. Veanes. The undecidability of simultaneous rigid E-unification with two variables. In G. Gottlob, A. Leitsch, and D. Mundici, editors, *Computational Logic and Proof Theory. 5th Kurt Gödel Colloquium, KGC'97*, volume 1289 of *Lecture Notes in Computer Science*, pages 305–318, Vienna, Austria, 1997.
8. A. Voronkov. Strategies in rigid-variable methods. In M.E. Pollack, editor, *Proc. of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97)*, volume 1, pages 114–119, Nagoya, Japan, August 23-29 1997.
9. A. Voronkov. Herbrand's theorem, automated reasoning and semantic tableaux. In *Proc. IEEE Conference on Logic in Computer Science (LICS)*, pages 252–263. IEEE Computer Society Press, 1998.
10. A. Voronkov. The ground-negative fragment of first-order logic is Π_2^p-complete. *Journal of Symbolic Logic*, 64(3):984–990, 1999.
11. Andrei Voronkov. Simultaneous rigid E-unification and other decision problems related to the Herbrand theorem. *Theoretical Computer Science*, 224:319–352, 1999.