Generalized Implicit Definitions on Finite Structures

Stéphane Grumbach*¹ and Zoé Lacroix**² and Steven Lindell***³

¹ University of Toronto and INRIA
² Université de Paris-Sud and INRIA
³ Haverford College

Abstract. We propose a natural generalization of the concept of implicit definitions over finite structures, allowing non-determinism at an intermediate level of a (deterministic) definition. These generalized implicit definitions offer more expressive power than classical implicit definitions. Moreover, their expressive power can be characterized over unordered finite structures in terms of the complexity class NP ∩ co-NP. Finally, we investigate a subclass of these where the non-determinism is restricted to the choice of a unique relation with respect to an implicit linear order, and prove that it captures UP ∩ co-UP also over the class of all finite structures. These results shed some light on the expressive power of non-deterministic primitives.

1 Introduction

Let I be a structure over a vocabulary σ, with R a relation symbol not in σ. A first-order sentence φ(R) over σ ∪ {R} implicitly defines a relation on I if there is a unique relation R satisfying φ(R) on I. On the other hand, a relation R is explicitly definable on I if there is a first-order formula φ(ā) over σ, which defines it in the ordinary sense: R = {ā ∈ I^k : I ⊨ φ(ā)}. Explicit definitions express a property of the tuples in the relation, whereas implicit definitions express a property of entire relation.

Beth's definability theorem [Bet53] states that if a global relation is implicitly definable over the class of all σ-structures, then it is explicitly definable. However, it is well known that this result fails when restricted to specific classes of models. In particular, this failure was illustrated by Gurevich for the class of finite structures [Gur84]. Implicit definitions over finite structures were then further investigated by Kolaitis, where their expressive power as definitions of

* I.N.R.I.A. Rocquencourt, Domaine de Voluceau BP 105, 78153 Le Chesnay, France – E-mail: stephane.grumbach@inria.fr – Work supported in part by an EC-NSF grant DEUS EX MACHINA, and an NSERC fellowship in Canada.
** L.R.I., Université de Paris Sud, 91405 Orsay, France – E-mail: zoe@lri.fr – Work done in the Verso Project at INRIA.
*** Department of Computer Science, Haverford College, Haverford, PA 19041-1392 – E-mail: slindell@haverford.edu – Work partially supported by NSF grant CCR-9403447 and the John C. Whitehead faculty research fund at Haverford College.
queries was studied [Kol90]. A sentence $\varphi(R)$ over $\sigma \cup \{R\}$ implicitly defines a query on a class of $\sigma$-structures $\mathcal{K}$, if for every structure $I$ in $\mathcal{K}$, there is a unique relation $R^I$, such that $I \models \varphi(R^I)$.

More generally, Kolaitis introduced the class IMP of queries which use a vector of implicit relations. A $k$-ary query $Q$ is definable in IMP on $\mathcal{K}$ if there exists a sentence $\varphi(R_1, \ldots, R_n)$ over $\sigma \cup \{R_1, \ldots, R_n\}$ where $R_1, \ldots, R_n$ are relation symbols not in $\sigma$ and $R_1$ is of arity $k$, such that on any $\sigma$-structure $I \in \mathcal{K}$, there is a unique sequence of relations $(R_1^I, \ldots, R_n^I)$, such that $I \models \varphi(R_1^I, \ldots, R_n^I)$ and $Q(I) = R_1^I$. It was shown in [Kol90] that fixpoint queries [CH82] can be expressed implicitly, as soon as the definition is based on two implicitly defined relations. Implicit definitions based on only one relation symbol provide less expressive power. On the other hand, it is easy to show that every implicit query can be expressed by only two implicitly defined relations, the output, $R$, and some other relation, $S$, by a first-order sentence $\varphi(R, S)$ over $\sigma \cup \{R, S\}$, such that $\varphi(R, S)$ implicitly defines $R$ and $S$. That is, for each $I \in \mathcal{K}$,

$$I \models \exists! R \exists! S \varphi(R, S).$$

The relation $S$ can be seen as an intermediate relation used as the "working area", necessary to get the full expressive power of IMP. The assumption that the intermediate relation is itself unique seems rather restrictive and unjustified. So we propose a generalization of the concept of an implicit query, by allowing non-determinism at the level of the intermediate relation $S$, provided the output relation $R$ is still deterministic. We do this by means of a sentence $\varphi(S)$ satisfied by all intermediate relations, and a formula $\psi(S, \bar{x})$ which is $S$-invariant with respect to $\varphi$, and defines the (unique) output relation explicitly. This relaxation of the strict uniqueness of classical implicit relations was first considered in [Lin87], where it was shown to be always equivalent to $\Delta^1_1$ definability. We call this class of generalized implicit queries, GIMP.

This results in a strict increase in expressive power, i.e. IMP $\subset$ GIMP. As an example, the query true if the cardinality of the domain is even can be expressed in GIMP, while it is not definable in IMP. In fact, it is easy to verify directly that every PTIME query is definable in GIMP. Using the inductive definition of PTIME [Imm86], it suffices to first guess an order on the domain, and then implicitly specify the fixed-point (and its negation). For further details consult [Lin87].

The use of non-determinism to compute deterministic queries has also been studied by Abiteboul, Simon and Vianu [ASV90]. For a non-deterministic language $\mathcal{L}$, they introduce two different deterministic semantics, the possibility semantics, and the certainty semantics. In the possibility semantics, each formula of $\mathcal{L}$ defines the set of tuples satisfying the formula for at least one of the non-deterministic choices. In the certainty semantics, each formula of $\mathcal{L}$ defines the set of tuples satisfying the formula for all choices. Generalized implicit definitions are related to these semantics in the following way. Queries in GIMP are those such that on every finite $\sigma$-structure the possibility and certainty semantics coincide.