Complexity of First Order ID-Logic

John S. Schlipf
Dept. of CS, Univ. of Cincinnati,
45221-0030, USA.
john.schlipf@uc.edu

Marc Denecker
Dept. of CS, K.U.Leuven, Belgium.
marcd@cs.kuleuven.be

Dedicated to Victor Marek
on his 65th birthday

Abstract

First Order ID-Logic interprets general first order, nonmonotone, inductive definability by generalizing the well-founded semantics for logic programs. We show that, for general (thus perhaps infinite) structures, inference in First Order ID-Logic is complete Π^1_2 over the natural numbers. We also prove a Skolem Theorem for the logic: every consistent formula of First Order ID-Logic has a countable model.

1 Introduction

Many formalisms have been proposed for non-monotonic reasoning. In search of semantics for logic programming, a key intuition of some researchers has been that of inductive definitions, as this view obviously applies for many prototypical Horn programs such as transitive closure, member, append, etc. In the last few years, the second author and others have suggested taking this seriously, replacing formalisms where induction is implicit with one where explicit inductive definability is in the heart of the formalism. A key step was the realization that, as argued in (Denecker, Bruynooghe, & Marek 2001), the well-founded model semantics of logic programming (Van Gelder, Ross, & Schlipf 1991) correctly formalizes the intuitions underlying different types of inductive definitions, not only monotone but also non-monotone inductive definitions over a well-founded order, and transfinite and iterated induction. This led (Denecker 2000; Denecker & Ternovska 2004; 2007) to propose ID-logic, an extension of classical logic with generalized nonmonotone inductive definitions. Similar extensions of classical logic with inductive definitions, in particular fixpoints logics, are used in other areas of Computer Science, but ID-logic differs from these by allowing a uniform representation of a broader class of inductive definitions.

To understand a logical formalism, we want to identify what can be expressed in it, not only in the most natural circumstances, but also in the most extreme. Here, in particular, we study the expressivity of First Order ID-Logic (FO(ID)). In particular, we study ID-logic over arbitrary (and thus potentially infinite) structures; the expressive power of FO(ID) over finite structures has been studied elsewhere (Mitchell & Ternovska 2005). By comparison, we just note that, in the context of finite structures, the expressive power of FO(ID) is similar to that of first order logic. For example, for any fixed FO(ID) formula ϕ, deciding whether a finite structure for a given vocabulary τ is a model of a FO(ID) formula is polynomial in the size of the structure. (That is, the "data complexity" of FO(ID) is polynomial. And, as in first order logic, the so-called "expression complexity" is EXPTIME.) For a fixed FO(ID) formula ϕ in a vocabulary τ, and for $\tau' \subseteq \tau$, determining whether a finite τ' structure is the projection (*i.e.,* reduct) of a model of ϕ is in NP and is NP-complete for certain FO(ID) formulas. And determining whether a formula is satisfiable in *some* finite structure is complete r.e. — just as for first order logic. For any fixed FO(ID) formula ϕ, there is a first order formula ϕ' in a larger language such that, for any finite structure $\mathfrak{A}$ for the language of ϕ, $\mathfrak{A} \models \phi$ if and only if $\mathfrak{A}$ can be expanded to a model of ϕ' (Mitchell & Ternovska 2005). So, in that sense, FO(ID) is no more expressive than first order logic over finite structures; FO(ID) provides only convenience in modeling. Of course, from a practical point of view, this modeling convenience may be considerable, since expressing an inductive definition in FO may amount to explicitly encoding fixpoint computations in FO. But, this is not a concern in a study of expressive power.

In this paper we consider the expressive power of FO(ID) over arbitrary structures. In particular, we determine how undecidable it is to ask whether a FO(ID) theory is satisfiable, or whether a FO(ID) theory logically implies a first order formula. Fairly typically for nonmonotonic logics (compare the survey in (Schlipf 1995)), the answers turn out to come from higher order logics and generalized recursion theory. We shall show that these decision problems are complete-Σ^1_2 and complete-Π^1_2 over the integers, respectively.

Here, in particular, we consider First Order ID-Logic (FO(ID)). For completeness, we repeat the formal definitions, with some notational additions allowing condensed discussion, but we do not repeat any motivating examples for ID-Logic.

1.1 Standard Logical Formalism

We use standard first-order-logic notation for vocabularies (languages) and structures. Of particular interest is the structure of arithmetic, $\mathfrak{N} = \langle \mathbb{N}; 0, \mathit{Succ}, +, \cdot \rangle$, where $\mathbb{N}$ is the set $\{0, 1, 2, \ldots\}$ of natural numbers. (To simplify notation, here we write the same symbol for the vocabulary element and its interpretation, rather than writing $0^{\mathfrak{N}}$, *etc.*) For each natural number n, let $\ulcorner n \urcorner$ be the term $\mathit{Succ}^n(0)$, which names n.

Let $\mathfrak{A} = \langle A; \ldots \rangle$ be a τ-structure. Let relation symbols $R_1, \ldots, R_m \notin \tau$, where each R_i is k_i-ary. For $1 \leq i \leq m$, let $S_i \subseteq A^{k_i}$. Then $\mathfrak{A}[S_1/R_1, \ldots, S_m/R_m]$ is the $\tau \cup \{R_1, \ldots, R_m\}$-structure with universe A that interprets all symbols in τ exactly as $\mathfrak{A}$ does and interprets each R_i by S_i. We use a similar notation on environments (assignments of global variables) on formulas: for $\vec{a}$ an n-tuple of elements of A and $\vec{x}$ an $n-tuple$ of variables, $\mathfrak{A} \models \phi(\vec{a}/\vec{x})$ says that ϕ is satisfied in $\mathfrak{A}$ in the environment where each x_i is interpreted by a_i.

For m-tuples $\vec{R} = \langle R_1, \ldots, R_m \rangle$, $\vec{S} = \langle S_1, \ldots, S_m \rangle$ of sets, we write $\vec{R} \subseteq \vec{S}$ to mean that each $R_i \subseteq S_i$, and $\vec{R} \subset \vec{S}$ to mean in addition that one inclusion is proper. For S a set of such m-tuples, $\bigcup S$ is the the coordinate-by-coordinate union.

Second Order (SO) Logic has two varieties of variable symbols: *first order variables*, written $x, y, z, x_1, \ldots$, vary over elements of (the universe of) the structure being discussed, while *second order variables*, written $R, S, X, Y, Z, R_1, \ldots$, vary over relations of appropriate arity on the structure. Second order variables may be quantified but otherwise have the same syntax as relation symbols of the vocabulary. It is well-known that second order logic is “grossly undecidable.”

Second order formulas are classified based on the number of alternations of second order quantifiers. Formulas with no second order quantifiers — thus only first order variables — are defined to be both Σ^1_0 and Π^1_0. A formula is Σ^1_{k+1} if it is of the form $\exists X_1 \ldots \exists X_n \phi$ where ϕ is Π^1_k; it is Π^1_{k+1} if it is of the form $\forall X_1 \ldots \forall X_n \phi$ where ϕ is Σ^1_k. This is the basis for the *analytical hierarchy* over the natural numbers and, with a change at the 0th level, for the *polynomial time hierarchy* over finite structures.

Each formula (first or second order) φ in τ with free first order variables, say $\vec{x}$, defines an n-ary relation in the context of a τ-structure $\mathfrak{M}$:

$$\{\vec{a} \in A^n : \mathfrak{M} \models \varphi(\vec{a}/\vec{x})\}.$$

A relation X in the domain of structure $\mathfrak{M}$ is said to be Σ^1_k *definable* in $\mathfrak{M}$ if it is definable over $\mathfrak{M}$ by a Σ^1_k formula. A set X of integers is Σ^1_k hard if, for every Σ^1_k set Y, there is a recursive function f_Y so that, for all integers y,

$$y \in Y \text{ if and only if } f_Y(y) \in X.$$

A set is Σ^1_k complete if it is both Σ^1_k definable and Σ^1_k hard. The definitions of Π^1_k definability, hardness and completeness are analogous. In the context of natural numbers, where we have recursive bijections from $\mathbb{N}^n$ to $\mathbb{N}$ (*e.g.*, Gödel numbering), an n-ary relation is Σ^1_k iff its mapping under this bijection is a Σ^1_k set. It therefore suffices to study Σ^1_k sets $(n = 1)$.

2 Induction in ID-Logic

Definition 2.1 *An* (inductive) definition Δ, *in a vocabulary (language)* τ, *for relations* $R_1, \ldots, R_m \notin \tau$,[1] *is a set of formulas* $\Delta =$

$$\begin{array}{c}\{\forall\vec{x}[R_1(\vec{x}) \;\leftarrow\; \phi_{1,1}(\vec{x})], \cdots, \forall\vec{x}[R_1(\vec{x}) \;\leftarrow\; \phi_{1,n_1}(\vec{x})],\\ \forall\vec{x}[R_2(\vec{x}) \;\leftarrow\; \phi_{2,1}(\vec{x})], \cdots, \forall\vec{x}[R_2(\vec{x}) \;\leftarrow\; \phi_{2,n_2}(\vec{x})],\\ \vdots\\ \forall\vec{x}[R_m(\vec{x}) \;\leftarrow\; \phi_{m,1}(\vec{x})], \cdots, \forall\vec{x}[R_m(\vec{x}) \;\leftarrow\; \phi_{k,n_m}(\vec{x})]\}\end{array}$$

where:

1. *formulas* $\phi_{i,j}$ *are in first order formulas of* $\tau \cup \{R_1, \ldots, R_m\}$ (we shall sometimes write $\phi(\vec{R}, \vec{x})$ to emphasize the extra relation symbols);
2. *symbol* $\leftarrow$ *is a new binary symbol* — called *definitional implication* in ID-logic;
3. *for* k_i *the arity of* R_i, *in each formula* $\forall\vec{x}[R_i(\vec{x}) \;\leftarrow\; \phi_{i,j}(\vec{x})]$, $\vec{x}$ *is* $x_1 x_2 \ldots x_{k_i}$*; and*
4. *we write* $\phi_{i,j}(\vec{x})$ *to indicate that the free variables of* $\phi_{i,j}$ *are among* $\vec{x}$.[2]

The intuition is that these rules inductively define the relations $R_1, \ldots, R_m$ by (simultaneous) induction. Thus, the individual formulas $\forall\vec{x}[R_i(\vec{x}) \;\leftarrow\; \phi_{i,j}(\vec{x})]$ are treated as closure rules. For any τ-structure $\mathfrak{A}$ and all $\vec{x}$ in structure $\mathfrak{A}$, whatever the interpretation of the R_hs are in $\mathfrak{A}$, if $\phi_{i,j}(\vec{x})$ is true, then $\vec{x}$ must be in (the interpretation of) R_i, and if no $\phi_{i,j}(\vec{x})$ is true, then $\vec{x}$ must not be in R_i. But this is not a full specification of what an inductive definition means. The next section defines the formal semantics of inductive definitions.

In ID-logic, the formulas $\forall\vec{x}[R_i(\vec{x}) \;\leftarrow\; \phi_{i,j}(\vec{x})]$ are called *definitional rules*, with *head* $R_i(\vec{x})$ and *body* $\phi_{i,j}(\vec{x})$. The predicates $R_1, \ldots, R_m$ are called the defined predicates of Δ, and all other symbols are called *open* symbols of Δ.

2.1 Positive Inductive Definability

When all occurrences of all defined relation symbols R_i in all the $\phi_{h,j}$'s are *positive*,[3] the inductive definition Δ is also referred to as *positive*. In this case, the semantics of inductive definability is standard (see, *e.g.*, (Moschovakis 1974a; Aczel 1977; Barwise 1975; Immerman 1986; Vardi 1982)). We quickly summarize it here.

[1] In (Denecker & Ternovska 2007), the vocabulary of Δ is considered to be our $\tau \cup \{R_1, \ldots, R_m\}$. Defining the vocabulary not to include $\{R_1, \ldots, R_m\}$ makes the discussion in this section a bit more straightforward and also stays closer to the language of some standard sources on inductive definitions, such as (Moschovakis 1974a; Aczel 1977; Barwise 1975). There is no significant difference.

[2] (Denecker & Ternovska 2007) does allow other free variables to appear in the $\phi_{i,j}$'s. But in FO(ID) these free variables cannot be quantified, so, for the questions asked in this paper, they can be treated as constant symbols.

[3] An occurrence if a relation symbol R in a first order formula ϕ is *positive* if, when ϕ is converted to negation normal form, that occurrence of R is not in the scope of a $\neg$. Otherwise, the occurrence of R is *negative*.

Given a τ-structure $\mathfrak{A}$ and an inductive definition Δ, define an operator Γ_Δ on m-tuples of relations: For relations $S_1, \ldots, S_m$ on $\mathfrak{A}$ (each relation S_i of the same arity as symbol R_i), let, for $1 \le i \le m$

$$S'_i = \{\vec{x} \in \mathfrak{A} : \mathfrak{A}[S_1/R_1, \ldots, S_m/R_m] \models \bigvee_{1 \le j \le n_i} \phi_{i,j}(\vec{x})\}. \quad (1)$$

$$\text{And let } \Gamma_\Delta(S_1, \ldots, S_m) = (S'_1, \ldots, S'_m) \quad (2)$$

Since the ϕ_i's are all positive in all the R_j's, operator Γ_Δ is monotone in its arguments. So by the Tarski-Knaster theorem, it has a least fixed point, which can be constructed by induction:

$$\vec{S}_0 = (\emptyset, \ldots, \emptyset), \qquad \vec{S}_{\alpha+1} = \Gamma_\Delta(\vec{S}_\alpha) \text{ for } \alpha \text{ any ordinal,}$$

$$\text{and } \vec{S}_\lambda = \bigcup_{\alpha<\lambda} \vec{S}_\alpha \text{ for } \lambda \text{ a limit ordinal.}$$

Denote the least fixed point by $\mathfrak{A}^\Delta = (R_1^{\mathfrak{A},\Delta}, \ldots, R_m^{\mathfrak{A},\Delta})$. Whenever $\mathfrak{A}$ is obvious, write, simply, R_i^Δ. Each relation $R_i^{\mathfrak{A},\Delta}$ is said to be *positively inductively definable* over $\mathfrak{A}$.

For any positive inductive definition Δ, and over any structure $\mathfrak{A}$ and for any sequence $\vec{a}$ of values for the free variables of $\mathfrak{A}$, the least ordinal β where $\vec{S}_\beta = \Gamma_\Delta(\vec{S}_\beta)$ is called the *closure ordinal of* Δ (over $\mathfrak{A}, \vec{a}$); we denote it by $|\Delta|_{\mathfrak{A},\vec{a}}$, or simply by $|\Delta|$ where context makes $\mathfrak{A}$ and $\vec{a}$ clear. Note that if κ is an infinite ordinal with more than $|A|$ predecessors (A the domain of $\mathfrak{A}$), then $|\Delta| < \kappa$, simply because the sequence of $\vec{S}_\alpha$'s is increasing and there are less than κ possible tuples to add into the relations.

A key result for this paper is the following:

Theorem 2.1 (Kleene-Spector Theorem) *A relation R on the natural numbers is positively inductively definable over $\mathfrak{N}$ if and only if it is Π^1_1 definable over $\mathfrak{N}$*[4].

2.2 Nonpositive Definitions

More complex forms of induction used in mathematics are inherently non-monotonic. For instance, the standard definition of the satisfaction relation $\models$ contains the non-monotonic rule

$$I \models \neg\varphi \text{ if } I \not\models \varphi.$$

As observed in (Denecker, Bruynooghe, & Marek 2001), the well-founded model semantics of logic programming (Van Gelder, Ross, & Schlipf 1991) uniformally formalizes the intuitions underlying different types of inductive definitions, not only monotone but also non-monotone inductive definitions over a well-founded order, and transfinite and iterated induction.[5] Thus, ID-logic uses the well-founded semantics for inductive definitions Δ in which the ϕ_i's need not be positive in all their arguments.

[4]This result has been generalized to broad classes of countably infinite structures; see, *e.g.*, (Barwise 1975; Moschovakis 1974a).

[5]A sort of non-monotonic induction not formalized by the well-founded model semantics is the "inflationary" induction of (Moschovakis 1974b).

There are many ways of formalizing the well-founded semantics. For the purposes of this paper, the alternating fixed point construction of the well-founded semantics for logic programs (Van Gelder 1993) is well-suited — extended in the obvious way to the broader class of inductive definitions defined above. For completeness, we present it here; for more discussion and many examples, again see (Denecker & Ternovska 2007).

Definition 2.2 *Let Δ be an inductive definition in vocabulary τ. For $i = 1, \ldots, m$, let R_i^- be a new*[6]*relation symbol of the same arity as R_i. For each definitional rule $\forall\vec{x}[R_i(\vec{x}) \leftarrow \phi_{i,j}(\vec{x})]$, form rule $\forall\vec{x}[R_i(\vec{x}) \leftarrow \hat{\phi}_{i,j}(\vec{x})]$ by replacing each* negative occurrence *of each R_h in $\phi_{i,j}$ with R_h^-. Let $\hat{\Delta}$ be the set of these new definitional rules.*

Observe that $\hat{\Delta}$ is a positive inductive definition of $R_1, \ldots, R_m$ over $\mathfrak{A}[S_1/R_1^-, \ldots, S_m/R_m^-]$. Call its least fixed point, $\mathfrak{A}[S_1/R_1^-, \ldots, S_m/R_m^-]^{\hat{\Delta}}$, simply $\mathcal{S}_\Delta(S_1, \ldots, S_m)$.

The above definition thus defines an operator of tuples of relations, mapping $(S_1, \ldots, S_m)$ to $\mathcal{S}_\Delta(S_1, \ldots, S_m)$. It extends the well-known stable operator of logic programming (Gelfond & Lifschitz 1991).

Van Gelder's Intuition for the Alternating Fixed Point Construction (phrased in the vocabulary of ID-Logic)**:** Suppose there is a "correct" interpretation $R_i^{\mathfrak{A}}$ of each R_i.

1. An inductively definable set should have only elements that are somehow "forced" to be in the set, so $\mathcal{S}_\Delta(R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$ should be $(R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$.
2. If $(S_1, \ldots, S_m) \subset (R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$, then negative literals $\neg R_i^-$ in rule bodies will be true "too often," so $\mathcal{S}_\Delta(S_1, \ldots, S_m) \supseteq (R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$.
3. Similarly, if $(S_1, \ldots, S_m) \supset (R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$, the negative subgoals will be true "too seldom," so $\mathcal{S}_\Delta(S_1, \ldots, S_m) \subseteq (R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$.

Formally, since the R_i^-'s occur only negatively in $\hat{\Delta}$, operator $\mathcal{S}_\Delta$ is anti-monotone, and hence $(\mathcal{S}_\Delta)^2$ is monotone. So the intuition gives us, for the "correct" relations $\vec{R^{\mathfrak{A}}} = (R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$ and $\vec{\emptyset} = (\emptyset, \ldots, \emptyset)$:

$$\begin{aligned}\vec{\emptyset} \subseteq (\mathcal{S}_\Delta)^2(\vec{\emptyset}) \subseteq (\mathcal{S}_\Delta)^4(\vec{\emptyset}) \subseteq \cdots \subseteq \vec{R^{\mathfrak{A}}} = \mathcal{S}_\Delta(\vec{R^{\mathfrak{A}}}) \\ \subseteq \cdots (\mathcal{S}_\Delta)^5(\vec{\emptyset}) \subseteq (\mathcal{S}_\Delta)^3(\vec{\emptyset}) \subseteq \mathcal{S}_\Delta(\vec{\emptyset}).\end{aligned} \quad (3)$$

And $(\mathcal{S}_\Delta)^2$ has a least fixed point, $\vec{S}_\infty$, constructible by transfinite induction:

$$\begin{aligned}&\vec{S}_0 = \vec{\emptyset},\\ &\vec{S}_{\alpha+1} = (\mathcal{S}_\Delta)^2(\vec{S}_\alpha) \ \ (\alpha \text{ any ordinal}),\\ &\vec{S}_\lambda = \textstyle\bigcup_{\alpha<\lambda} \vec{S}_\alpha \ \ (\lambda \text{ a limit ordinal}), \text{and}\\ &\vec{S}_\infty = \text{the least fixed point.}\end{aligned}$$

[6]By "new" we imply that the symbols R_i^- are all distinct and that none are in $\tau \cup \{R_1, \ldots, R_m\}$.

Using the monotonicity and anti-monotonicity of respectively $(\mathcal{S}_\Delta)^2$ and $\mathcal{S}_\Delta$, we can prove a result slightly weaker than (3):

$$\vec{\emptyset} \subseteq (\mathcal{S}_\Delta)^2(\vec{\emptyset}) \subseteq (\mathcal{S}_\Delta)^4(\vec{\emptyset}) \subseteq \cdots \subseteq \vec{S}_\infty \subseteq \mathcal{S}_\Delta(\vec{S}_\infty) \\ \subseteq \cdots (\mathcal{S}_\Delta)^5(\vec{\emptyset}) \subseteq (\mathcal{S}_\Delta)^3(\vec{\emptyset}) \subseteq \mathcal{S}_\Delta(\vec{\emptyset}).$$

Thus, in case $\vec{S}_\infty = \mathcal{S}_\Delta(\vec{S}_\infty)$, we have fully captured the intuition above.

Definition 2.3 *For an inductive definition Δ in τ, any τ structure $\mathfrak{A}$, and $\vec{S}_\infty$ as above, if $\mathcal{S}_\Delta(\vec{S}_\infty) = \vec{S}_\infty$, we say definition Δ defines $\mathfrak{A}^\Delta = \vec{S}_\infty = (R_1^{\mathfrak{A},\Delta}, \ldots, R_m^{\mathfrak{A},\Delta})$ inductively; we also say Δ defines each $R_i^{\mathfrak{A},\Delta}$ inductively. Otherwise, we say that $\mathfrak{A}^\Delta$ and all $R_i^{\mathfrak{A},\Delta}$'s are undefined.*

The least ordinal α where $\vec{S}_\alpha = \vec{S}_{\alpha+2}$ is called the closure ordinal of Δ *(over $\mathfrak{A}, \vec{a}$); it is denoted $|\Delta|_{\mathfrak{A},\vec{a}}$ — or simply $|\Delta|$ when $\mathfrak{A}, \vec{a}$ are clear.*

As with positive induction, if κ is an infinite ordinal with more than $|A|$ predecessors, then $|\Delta| < \kappa$.

Proposition 2.2 *Let $\mathfrak{A}$ be a τ-structure.*

1. *If Δ is a positive inductive definition in τ, then Δ defines the same relations over $\mathfrak{A}$ in ID-logic than it does in first order positive inductive definability.*
2. *If relations $S_1, \ldots, S_m$ are inductively definable in ID-logic over $\mathfrak{A}$, and relation S is inductively definable in ID-logic over structure $\mathfrak{A}[S_1/R_1, \ldots, S_m/R_m]$, then S is inductively definable in ID-logic over $\mathfrak{A}$.*
3. *The set of relations inductively definable over $\mathfrak{A}$ in ID-logic is closed under boolean combinations and projections.*

Proof:

1. Since the inductively defined relations do not occur negatively in Δ, the tuple $\mathcal{S}_\Delta(S_1, \ldots, S_m)$ does not depend upon $S_1, \ldots, S_m$. Hence $(\mathcal{S}_\Delta)^2(\vec{\emptyset})$ is a fixed point and equals $(\mathcal{S}_\Delta)^3(\vec{\emptyset})$.
2. This is a standard consequence of the monotonicity of $(\mathcal{S}_\Delta)^2$ (sometimes referred to as showing that iterated induction is the same as simultaneous induction); compare the Transitivity Theorem 1.C.3 of (Moschovakis 1974a). Alternatively, it is also a simple consequence of the abstract stratification theorem 3.11 for non-monotone operators (Vennekens, Gilis, & Denecker 2006).
3. By part (2), if Δ defines $R_1, \ldots, R_4$ (and R_1, R_2 have the same arity, and R_4 has arity > 1), then $\Delta \cup \{\ R_\cup(\vec{x}) \leftarrow R_1(\vec{x}) \vee R_2(\vec{x}),\ \ R_\cap(\vec{x}) \leftarrow R_1(\vec{x}) \wedge R_2(\vec{x}),\ \ R_\neg(\vec{x}) \leftarrow \neg R_3(\vec{x}),\ \ R_\pi(\vec{x}) \leftarrow \exists y R_4(\vec{x}, y)\ \}$ defines the desired union, intersection, complement, and projection. ∎

Observation 2.3 *The above proposition shows a key difference between positive inductive definability and the inductive definability in ID-logic: over many infinite structures, e.g. the structure $\mathfrak{N}$ for arithmetic, the class of first order positively inductively definable sets is not closed under complementation. Thus, over such structures, the inductive definitions of ID-logic are very different than in first order positive inductive definability.*

On the other hand, Immerman and Vardi proved (Immerman 1986; Vardi 1982) proved that, over finite structures, the class of relations that are (uniformly) positive inductively definable is closed under complementation.

3 FO(ID)

First Order Inductive Definition Logic (FO(ID)) extends first order logic with inductive definitions. We merely give the formal definition here; see (Denecker & Ternovska 2007) for many motivating examples.

An FO(ID) formula over is defined by adding an additional base case to the standard inductive rules defining first order formulas over a vocabulary τ:

- A definition Δ of predicate symbols $R_1, \ldots, R_m$ in $\tau \setminus \{R_1, \ldots, R_m\}$ is an FO(ID) formula in τ. Δ may contain free variables.

Thus, the construction units of FO(ID) are the atoms and the definitions, and the logic is closed under conjunction, disjunction, negation, existential and universal quantification. Since rule bodies of definitions are FO, nested definitions are not allowed in FO(ID), contrary to, e.g., the logic LFP (Libkin 2004).

Definition 3.1 *A FO(ID)* theory *T in a vocabulary τ is a set of FO(ID) sentences in τ.*

Definition 3.2 *The satisfaction relation — denoted $\models_{[ID]}$ — of FO(ID) is defined by the same structural rules defining satisfaction $\models$ of FO, augmented with one extra base rule:*

- *for a structure $\mathfrak{A}$ interpreting τ and all free variables of Δ, let $\mathfrak{A}'$ be the reduct of $\mathfrak{A}$ to $\tau' = \tau \setminus \{R_1, \ldots, R_n\}$. We define $\mathfrak{A} \models \Delta$ if $(\mathfrak{A}')^\Delta$ exists and each $R_i^{\mathfrak{A}',\Delta}$ is equal to $R_i^{\mathfrak{A}}$ (the interpretation of R_i in $\mathfrak{A}$).*

Example 3.1 *Let Δ_1, Δ_2 be two inductive definitions of the same set $\{R_1, \ldots, R_m\}$ of relations. A structure $\mathfrak{A} = \langle A; \ldots, R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}} \rangle$ is a model of $\Delta_1 \wedge \Delta_2$ if and only if $R_1^{\mathfrak{A}} \ldots R_m^{\mathfrak{A}}$ are the relations defined by Δ_1 and also are the relations defined by Δ_2 — so, in particular, only if it turns out that the relations defined by Δ_1 and Δ_2 over $\mathfrak{A}' = \langle A; \ldots \rangle$ are the same.*

For T a theory of FO(ID) and ϕ an FO(ID) sentence, T *logically implies* ϕ — written $T \models_{[ID]} \phi$ — if ϕ is true in every model of T. Note that, just as in first order logic, for ϕ a sentence of FO(ID), $T \models_{[ID]} \phi$ if and only if $T \cup \{\neg\phi\}$ is unsatisfiable.[7]

4 Inference in FO(ID) is Π^1_2 over Arithmetic

Our proof below is *heavily* dependent upon (i) formalization of model theory in the universe V of sets (with relation $\in$), and (ii) results in ordinal recursion theory.

[7] Note that a definitional rule ϕ is not an FO(ID) sentence, such that $T \models_{[ID]} \phi$ is not defined.

4.1 Background in Set Theory

In this paper we use several results about definability in set theory. We summarize them here and in the next section. (Proofs of theorems can be found in, for example, (Barwise 1975).)

The vocabulary of set theory is $\{\in\}$; the usual axioms are the Zermelo-Fraenkel (ZF) axioms plus the Axiom of Choice (AC). It is assumed that there is a real universe of sets, called V, and that it satisfies ZF+AC. When we talk about definability in set theory, we talk about definability in V, not in arbitrary models of ZF+AC.

An $\{\in\}$-formula is Δ_0 if it is built up from atomic formulas using only boolean connectives and *bounded quantification*: $\exists x \in y\, \phi$ (defined to be $\exists x(x \in y \wedge \phi)$) and $\forall x \in y\, \phi$ (defined to be $\forall x(x \in y \rightarrow \phi)$). A formula is Σ_1 if it is of the form $\exists x_1, \ldots, x_k \phi$ where ϕ is Δ_0. An important result (provable even in a fairly weak set theory such as KPU) is that if ψ_1, ψ_2 are Σ_1 formulas and x, y are any variable symbols, then $(\psi_1 \vee \psi_2)$, $(\psi_1 \wedge \psi_2)$, $\exists x \in y\, \psi_1$, and $\forall x \in y\, \psi_1$ are equivalent to Σ_1 formulas.

Recall that, for a set theorist, objects such as structures, formulas and tuples are all special kinds of sets. For example, a structure $\mathfrak{A} = \langle A, R_1^{\mathfrak{A}}, \ldots, R_k^{\mathfrak{A}}\rangle$ for a vocabulary $\tau = \{R_1, \ldots, R_m\}$ can be represented as a tuple $\langle A, \{\langle n_1, R_1^{\mathfrak{A}}\rangle, \ldots, \langle n_k, R_k^{\mathfrak{A}}\rangle\}\rangle$ with n_i the Gödel-number of symbol R_i. And τ-formulas may be represented, for example, by their Gödel numbers. The standard inductive definition of satisfaction is by an induction that can be captured by a straightforward Σ_1 formula:

Theorem 4.1 *There is a Σ_1 formula $\chi_{Sat}(\ell, s, f, x, v)$ of set theory such that,*

$$V \models \chi_{Sat}(\tau/\ell, \mathfrak{A}/s, \phi/f, \vec{x}/x, \vec{a}/v) \qquad \textit{if and only if}$$

- *τ is a finite language,*
- *$\mathfrak{A}$ is a τ-structure,*
- *ϕ is (the Gödel number of) a first order formula of τ,*
- *for some natural number n, $\vec{x}$ is the n-tuple of all free variable symbols of ϕ, and $\vec{a}$ is an n-tuple of elements of A, and*
- *$\mathfrak{A} \models \phi(\vec{a}/\vec{x})$.* ∎

A set or class S of sets is *transitive* if whenever any $x \in S$ and $y \in x$, $y \in S$. A (von Neumann) *ordinal* is a set that is transitive and linearly ordered by relation $\in$. The class of ordinals is also linearly ordered by $\in$, and the axioms of ZF prove that this is a well-ordering in V. On ordinals α, β, $\alpha < \beta$ means $\alpha \in \beta$, and $\alpha = \beta + 1$ means $\alpha = \beta \cup \{\beta\}$. (By compactness, there are "non-standard" models of ZF+AC, where the "ordinals" are not well-ordered, but that is not a concern here; we are working over V.)

It is obvious than when a Σ_1^1 formula $\exists X_1 \ldots \exists X_k \psi$ (with free variables $\vec{x}$) uniformly defines relations $S_{\mathfrak{A}}$ in τ-structures $\mathfrak{A}$, then these relations $S_{\mathfrak{A}}$ can be defined uniformly in V by a Σ_1 formula:

$$S_{\mathfrak{A}} = \{\vec{a} \in A^n : V \models \exists Y_1 \ldots \exists Y_k \\ (\chi_{Sat}(\tau'/l, \mathfrak{A}[Y_1/X_1, \ldots, Y_k/X_k]/s, \psi/f, \vec{x}/x, \vec{a}/v))\}$$

where $\tau' = \tau \cup \{X_1, \ldots, X_k\}$, i.e., the symbols $X_1, \ldots, X_n$ are now treated as predicate constants.

But the converse fails:

Example 4.1 (Well-known) *Suppose τ contains binary relation G.*

1. *There is no Σ_1^1 formula $\exists\vec{X}\phi(x, y)$ uniformally defining the transitive closure of G in all τ-structures $\mathfrak{A}$.* This follows directly from compactness for first order logic. Indeed, assume towards contradiction that Σ_1^1 formula $\exists\vec{X}\phi$ expresses the transitive closure of G. Consider the infinite theory $\Psi = \{\neg G(a, b), \neg\exists x_0 \ldots \exists x_n(G(a, x_0) \wedge \ldots \wedge G(x_n, b)) : n \in \mathbb{N}\}$, where a, b are constants not appearing in $\exists\vec{X}\phi$. Clearly, $\Psi \cup \{\exists\vec{X}\phi(a, b)\}$ is unsatisfiable, and so is $\Psi \cup \{\phi(a, b)\}$. By compactness of FO, the latter theory should have a finite unsatisfiable subset, and this is clearly not the case.
2. *But there is a Σ_1 formula $\psi(s, x, y)$* of set theory *where $V \models \psi(\mathfrak{A}/s, d/x, d'/y)$ if and only if (d, d') is in the transitive closure of $G^{\mathfrak{A}}$.* The formula is — in mixed formal notation and English —

$$\begin{aligned} \exists f, \alpha(\quad & (\alpha \text{ is a natural number}) \wedge (f : \alpha + 1 \rightarrow A) \wedge \\ & f(0) = x \wedge f(\alpha) = y \wedge \\ & \forall n \in \alpha(G(f(n), f(n+1))) \\ ) \quad & \end{aligned}$$

Theorem 4.2 *There is a Σ_1 formula $\chi_{posInd}(\ell, s, d, x, v, r)$ of set theory where*

$$V \models \chi_{posInd}(\tau/\ell, \mathfrak{A}/s, \Delta/d, \vec{v}/x, \vec{a}/v, \langle S_1, \ldots, S_m\rangle/r)$$

if and only if

1. *τ is a finite language,*
2. *$\mathfrak{A}$ is a τ structure,*
3. *Δ is a positive inductive definition of some relations $R_1, \ldots, R_m \notin \tau$, with m the length of $\langle S_1, \ldots, S_m\rangle$, and with free variables $\vec{x} = x_1, \ldots, x_n$,*
4. *the arity of each S_i is the same as of R_i and $\vec{a}$ is an n-tuple of elements of $\mathfrak{A}$ (for n as above), and*
5. *$\langle S_1, \ldots, S_m\rangle = \mathfrak{A}^{\Delta}$, where the environment binds each v_i to a_i.* ∎

For a proof of the above, again see, *e.g.*, (Barwise 1975). (And the same approach is used in our proof of Lemma 4.6.) There is a Σ_1 formula that identifies the list of symbols R_i and their arities from the syntactic form of Δ.

4.2 Background in Constructibility and Ordinal Recursion Theory

Gödel proved the relative consistency of the axiom of choice and the continuum hypothesis using a smaller class of sets, the class of constructible sets, called L. The constructible sets are constructed by transfinite induction over all ordinals $\alpha \in V$; $L(\alpha)$ is the set of sets thus constructed before stage α. Each $L(\alpha)$, as well as L, is transitive.

Theorem 4.3 (absoluteness of Δ_0 and persistence of Σ_1) *Let S be a transitive set or class of sets. For any Δ_0 formula*

$\phi(\vec{x})$*, and for any vector* $\vec{s}$ *of appropriate length and of elements of* S*,*

$$S \models \phi(\vec{s}/\vec{x}) \text{ if and only if } V \models \phi(\vec{s}/\vec{x}).$$

For any Σ_1 *formula* $\phi(\vec{x})$*, and for any vector* $\vec{s}$ *of appropriate length and of elements of* S*,*

$$\text{if } S \models \phi(\vec{s}/\vec{x}) \text{ then } V \models \phi(\vec{s}/\vec{x}).$$

An ordinal σ is *stable* if, for every Σ_1 formula ϕ with free variables among $\vec{x}$, and for every $\vec{a} \in L(\sigma)$, $L \models \phi[\vec{a}/\vec{x}]$ if and only if $L(\sigma) \models \phi[\vec{a}/\vec{x}]$ — *i.e.*, in standard notation, $L(\sigma) \prec_1 L$. There are countable stable ordinals. The least one is called σ_0. It is not difficult to show that the structure $\mathfrak{N} \in L(\sigma_0)$, which makes $L(\sigma_0)$ suitable to study expressive power in the context of $\mathfrak{N}$.

Theorem 4.4 *A relation* R *on the natural numbers is* Σ^1_2 *definable on* $\mathfrak{N}$ *if and only if it is* Σ_1*-definable on* $L(\sigma_0)$*.* ∎

Theorem 4.5 (Schoenfield Absoluteness Theorem) *Every* Σ_1 *sentence (i.e.,formula with no free variables) of set theory true in* V *is also true in* L*.* ∎

It is an easy generalization to show that, if $\theta(x)$ is a Σ_1 formula whose only free variable is x, and if n is an integer, if $V \models \theta(n/x)$ then $L \models \theta(n/x)$.

4.3 Inference in FO(ID) is Π^1_2 over $\mathfrak{N}$

Lemma 4.6 *There is a* Σ_1 *formula*

$$\exists \alpha, \delta, F, \pi, \hat{d}, w(\theta_{wfInd}(\alpha, \delta, F, \pi, \hat{d}, \ell, s, d, x, v, r, w))$$

of set theory where θ_{wfInd} *is* Δ_0 *and*

$$V \models \theta_{wfInd}(\tau/\ell, \mathfrak{A}/s, \Delta/d, \vec{x}/x, \vec{a}/v, \langle S_1, \ldots, S_m\rangle/r, u/w)$$

if and only if

1. τ *is a finite language,*
2. $\mathfrak{A}$ *is a* τ *structure,*
3. Δ *is a definition of some relations* $R_1, \ldots, R_m \notin \tau$*, with* m *the length of* $\langle S_1, \ldots, S_m\rangle$*, and with free variables* $\vec{x}$ *equal to some* $x_1, \ldots, x_n$*,*
4. *the arities of each* S_i *is the same as of* R_i *and* $\vec{a}$ *is an* n*-tuple of elements of* $\mathfrak{A}$ *(for* n *as above), and,*
5. *in the notation of formula 2.2,* $\langle S_1, \ldots, S_m\rangle$ *is the least fixed point of* $(\mathcal{S}_\Delta)^2$*.*

Proof:

The constructions for parts (1-4) are fairly standard, so we omit them; see, *e.g.*, the proof of our Theorem 4.1 in (Barwise 1975).

Theorem 4.2 shows that positive induction is Σ_1 definable, say by formula $\exists \vec{y}\theta(\ell, s, d, x, v, r, \vec{y})$. The alternating fixed point construction is by induction, where each stage of the induction uses positive inductive definition. To capture the inner induction we use χ_{posInd}; to capture the outer induction, we use a function F with domain an ordinal α (where, in light of earlier remarks, we could choose α to be any infinite ordinal with $> |A|$ predecessors).

For each ordinal β in its domain, $F(\beta)$ wil be an ordered pair

$$\langle \mathfrak{A}[S_1/R^-_1, \ldots, S_m/R^-_m], \vec{y}\rangle, \tag{4}$$

where each inductively defined predicate R_i has been split into positive occurrences R_i and negative occurrences R^-_i as in Definition 2.2 — and $\vec{y}$ will be witnesses for existensial quantifiers, as noted below. Below, let

- $F(\beta)_\mathfrak{A}$ be $\mathfrak{A}[S_1/s^-_1, \ldots, S_m/s^-_m]$,
- $F(\beta)_{(i)}$ be the S_i in $F(\beta)$, and
- $F(\beta)_{(\vec{y})}$ be the $\vec{y}$ of $F(\beta)$.

Formally, we can replace mention of them below with Δ_0 formulas involving $F(\beta)$.

Our formula is

$\exists \alpha, \delta, F, \pi, \hat{d}$
(π *is a parsing function witnessing that* ϕ *defines* $\vec{r}$
  $\wedge$ $\hat{d}$ *is constructed from* d *as in Definition 2.2*
  $\wedge$ α, δ *are ordinals* $\wedge$ $(\delta + 2 < \alpha)$
  $\wedge$ F *is a function with domain* α
  $\wedge$ $\forall \beta < \alpha(F(\beta)$ *is of the form of (4))*
  $\wedge$ $F(\delta) = F(\delta + 2)$ (so a fixpoint has been reached)
  $\wedge$ $\langle F(\delta)_{(1)}, \ldots, F(\delta)_{(m)}\rangle \subseteq \langle F(\delta+1)_{(1)}, \ldots, F(\delta+1)_{(m)}\rangle$
  $\wedge$ $\bigwedge_{1 \le i \le m} F(0)_{(i)} = \emptyset$
  $\wedge$ $\forall \beta < \alpha(\beta$ *is a successor ordinal* $\rightarrow$
    $\theta(\ell, F(\beta-1)_\mathfrak{A}, \hat{d}, x, v, \langle F(\beta)_{(1)}, \ldots, F(\beta)_{(m)}\rangle, F(\beta)_{(\vec{y})})$
  $\wedge$ $\forall \beta < \alpha(\beta$ *is a limit ordinal* $\rightarrow$
    $\bigwedge_{1 \le i \le m} F(\beta)_{(i)} = \bigcup_{\gamma<\beta}(F(\gamma)_{(i)} \cap F(\gamma+1)_{(i)})$
  $\wedge$ $\bigwedge_{1 \le i \le m} \forall \vec{x}(\vec{x} \in r_i \leftrightarrow \vec{x} \in F(\delta)_{(i)})$
)

It is fairly clear that the formula "says" that F encodes the stages of the alternating fixed point construction and that $F(\delta)$ is a fixed point of $(\mathcal{S}_\Delta)^2$. But there are really two fixed points achieved this way, $\vec{S}_\infty$ and $\mathcal{S}_\Delta(\vec{S}_\infty)$. As noted earlier, $\vec{S}_\infty \subseteq \mathcal{S}_\Delta(\vec{S}_\infty)$, so the line following the fixed point condition $F(\delta) = F(\delta + 2)$ states that $F(\delta)$ gives the intended one. For a limit ordinal β, we should express that the $F(\beta)_{(i)}$'s are the limit of earlier lower approximations. For each $\gamma < \beta$, the intersection $F(\gamma)_{(i)} \cap F(\gamma+1)_{(i)}$ is the lower approximation of the pair $S_{\gamma,i}, S_{\gamma+1,i}$.

It can be shown that each of the conjuncts following the initial $\exists \alpha, \delta, F, \vec{r}, z$ is expressible with a Δ_0 formula. It can be shown that replacement/reflection axioms of ZF show that the function F itself exists (indeed, this is basically just the standard proof that in set theory one can do definition by recursion). ∎

Theorem 4.7 *There is a* Σ_1 *formula* $\chi_{FO(ID)}(\ell, s, f, x, v)$ *of set theory where*

$$V \models \chi_{FO(ID)}(\tau/\ell, \mathfrak{A}/s, \phi/f, \vec{x}/x, \vec{a}/v) \quad \text{if and only if}$$

1. τ *is a finite language,*
2. $\mathfrak{A}$ *is a* τ *structure,*
3. ϕ *is a formula of FO(ID) with free variables* $\vec{x}$*,*
4. $\vec{a}$ *is a tuple of elements of* $\mathfrak{A}$*, of the same length as* $\vec{x}$*, and*

5. $\mathfrak{A} \models \phi(\vec{a}/\vec{x})$.

Proof (sketch):

This construction repeats the standard construction used for Theorem 4.1. First use existensial quantification to produce the parse π for ϕ. The construction is now by induction on the nodes the parse tree. The only new step for ID-logic is to handle the base case of definitions. In fact, there are two new base cases. Indeed, we assume that the formula is in negation normal form, so that the only negative occurrences of inductive definitions Δ are in the form $\neg\Delta$. We cannot express that $\mathfrak{A} \models \neg\Delta(\vec{a}/\vec{x})$ through $\neg(\mathfrak{A} \models \Delta(\vec{a}/\vec{x}))$ since then, the Σ_1 formula expressing $\mathfrak{A} \models \Delta(\vec{a}/\vec{x})$ would appear under negation, and Σ_1 formulas are not closed under negation. Fortunately, we can say that $\mathfrak{A} \models \neg\Delta(\vec{a}/\vec{x})$ by a Σ_1 formula, and this is our second base case.

Say $\mathfrak{A} \models \Delta(\vec{a}/\vec{x})$ by writing

$\exists p, l', s', \alpha, \delta, F, \pi, \hat{d}, r$
(p is the tuple of defined predicates of Δ
$\wedge\ \ell' = \ell \setminus p \wedge s' =$ the reduct of $\mathfrak{A}$ to l'
$\wedge\ \theta_{wfInd}(\alpha, \delta, F, \pi, \hat{d}, \ell', s', \Delta/d, \vec{x}/x, \vec{a}/v, r)$
$\wedge\ \bigwedge_{1\leq i\leq m} F(\delta)_{(i)} = F(\delta+1)_{(i)}\ \wedge\ \bigwedge_{1\leq i\leq m} p_i^{\mathfrak{A}} = r_i$
).

Say $\mathfrak{A} \models \neg\Delta(\vec{a}/\vec{x})$ by writing

$\exists p, l', s', \alpha, \delta, F, \pi, \hat{d}, r$
(p is the tuple of defined predicates of Δ
$\wedge\ \ell' = \ell \setminus p \wedge s' =$ the reduct of $\mathfrak{A}$ to l'
$\wedge\ \theta_{wfInd}(\alpha, \delta, F, \pi, \hat{d}, \ell', s', \Delta/d, \vec{x}/x, \vec{a}/v, r)$
$\wedge\ \bigvee_{1\leq i\leq m} F(\delta)_{(i)} \neq F(\delta+1)_{(i)}\ \vee\ \bigvee_{1\leq i\leq m} p_i^{\mathfrak{A}} \neq r_i$
). ∎

Corollary 4.8 *There is a Σ_1 formula $\chi_{SAT_{ID}}(f)$ of set theory so that a FO(ID) sentence ϕ is satisfiable if and only if $V \models \chi_{SAT_{ID}}(\phi/f)$, and ϕ is satisfiable if and only if $L(\sigma_0) \models \chi_{SAT_{ID}}(\phi/f)$.*

Proof: The formula $\chi_{SAT_{ID}}(f) = \exists\ell, s\chi_{FO(ID)}(\ell, s, f, \emptyset, \emptyset)$ clearly satisfies the first property. It is also routine to show that if $L(\sigma_0) \models \chi_{SAT_{ID}}(\phi/f)$ then ϕ is indeed satisfiable.

So now suppose ϕ is satisfiable — so $V \models \chi_{SAT_{ID}(\phi/f)}$. Recall that we represented formulas in set theory with their Gödel numbers. By the generalization of the Schoenfield Absoluteness Theorem, $L \models \chi_{SAT_{ID}(\phi/f)}$. Since σ_0 is a stable ordinal (and thus also infinite), $L(\sigma_0) \models \chi_{SAT_{ID}}(\phi/f)$. ∎

Corollary 4.9 (Skolem Theorem for FO(ID)) *For ϕ a formula of FO(ID), if ϕ has a model, it has a countable* (finite or countably infinite) *model.*

Proof: By Corollary 4.8, if ϕ is satisfiable, ϕ is satisfiable in $L(\sigma_0)$. And all elements of $L(\sigma_0)$ are countable. ∎

Theorem 4.10 (a) *Satisfiability of FO(ID) formulas is Σ^1_2 over $\mathfrak{N}$.* (b) *For FO(ID) formulas ϕ, ψ, determining whether ψ logically implies ψ is Π^1_2 over $\mathfrak{N}$.*

Proof: (a) is a consequence of Theorem 4.4 and Corollary 4.8. (b) follows immediately. ∎

5 Inference in FO(ID) is Π^1_2-hard over Arithmetic

We show here that determining whether $T \models_{[ID]} \phi$, for T an FO(ID) theory and ϕ first order, is Π^1_2-hard (over arithmetic), even in the special case where Δ is a system of *positive* inductive definitions.

Example 5.1 *Let τ be the usual language $\{0, <, Succ, +, \cdot\}$ for arithmetic, plus one unary relation S to be inductively defined. There is a finite FO(ID) theory $T_{\mathfrak{N}}$ whose models (or rather, their reducts to $\{0, <, Succ, +, \cdot\}$) are just the isomorphic copies of $\mathfrak{N}$. And all definitional rules in $T_{\mathfrak{N}}$ are positive. For example, take $T_{\mathfrak{N}}$ to be Peano's theory in which the induction schema is replaced by*

$$\forall x N(x) \wedge \left\{ \begin{array}{l} \forall x(N(x) \leftarrow x = 0), \\ \forall x(N(x) \leftarrow \exists y(x = Succ(y) \wedge N(y))) \end{array} \right\}$$

All models $\mathfrak{A}$ of Peano's axioms (with or without the induction schema) have a "standard part" $\{0^{\mathfrak{A}}, \ulcorner 1 \urcorner^{\mathfrak{A}}, \ulcorner 2 \urcorner^{\mathfrak{A}}, \ulcorner 3 \urcorner^{\mathfrak{A}}, \ldots\}$, which is isomorphic to $\mathfrak{N}$. The positive inductive definition defines N as this set. So $\forall x N(x)$ asserts that every element is in the standard model, as desired. ∎

Observation 5.1 (Relativized Kleene-Spector Theorem) *There are a recursive function f from formulas $\phi(X, Y, x)$ to Y-positive formulas $\phi^f(X, Y, z)$ and a recursive function g from formulas $\phi(X, Y, x)$ to integers ϕ^g, such that, for all $n \in \mathbb{N}$ and all sets $\mathbf{X} \subseteq \mathbb{N}$, for $\Delta = \{\forall z[Y(z) \ \leftarrow\ \phi^f(X, Y, z)]\}$,*

$$\mathfrak{N}[\mathbf{X}/X] \models \forall Y \phi(X, Y, x)[n/x]$$

if and only if

$$\langle n, \phi^g \rangle \in Y^{\mathfrak{N}[\mathbf{X}/X], \Delta}$$

Proof (sketch): The proof is a straightforward modification of the proof of Moschovakis' generalization in §8A of (Moschovakis 1974a). The observations are merely that (i) the proof there constructs a formula explicitly — and thus recursively — from ϕ, and (ii) the extra relation X may be just carried along as a parameter. ∎

Theorem 5.2 (a) *Determining whether a finite FO(ID) theory is satisfiable is Σ^1_2-hard over $\mathfrak{N}$.* (b) *Determining whether $T \models_{[ID]} \psi$, for T a FO(ID) theory and ψ FO(ID) formula, is Π^1_2-hard over $\mathfrak{N}$.*

Proof: As usual, (a) implies (b). We show (a) even in the special case where the FO(ID) formula is of the form $\psi \wedge \Delta$ where ψ is first order and Δ is a single positive (inductive) definition. We continue to use the notations from above. We use that fact that, when an (inductive) definition is positive, the inductively defined relations always exist. Let $\Delta = \{\forall z[Y(z) \ \leftarrow\ \phi^f(X, Y, z)]\}$.

$$N \models \exists X \forall Y \phi(X, Y, n)$$

if and only if

$$\exists \mathbf{X} \subseteq N(\langle n, \phi^g \rangle \in Y^{\mathfrak{N}[\mathbf{X}/X], \Delta}$$

if and only if

$T_{\mathfrak{N}} \cup \Delta \cup \{Y(\langle \ulcorner n \urcorner, \ulcorner \phi^g \urcorner \rangle)\}$ is satisfiable in ID-Logic. ∎

An interesting observation is that, although the expressive power of the well-founded induction of FO(ID) is, in general, greater than that of first order positive induction, the complexity of determining satisfiability for FO[ID] is the same as it is for first order logic plus just positive inductive definability.

6 Open Problem: Expressive Power Well-founded Induction $\mathfrak{N}$

Here we have discussed the expressive power of FO(ID). A related issue is the expressive power of just well-founded inductive definitions — as interpreted in FO(ID) (see Subsection 2.2) — over particular structures. The classical structure to consider is $\mathfrak{N}$. So what relations are definable using well-founded induction over $\mathfrak{N}$? Since FO(ID)'s inductive definitions include all positive inductive definitions, it follows by the Kleene-Spector Theorem that all Π^1_1 relations on $\mathfrak{N}$ are so definable. It follows easily from Theorem 5.2 that all so definable relations are Π^1_2 over $\mathfrak{N}$. But neither the Π^1_1- nor Π^1_2-definable relations are closed under complementation, so the sets definable using just FO(ID)'s nonmonotonic induction are strictly in between the two. We pose this question for further research.

References

Aczel, P. 1977. An introduction to inductive definitions, In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 739–782, North-Holland, New York.

Barwise, J. 1975. *Admissible Sets and Structures.* Springer-Verlag, New York.

Denecker, M., and Ternovska, E. 2004. A logic of nonmonotone inductive definitions and its modularity properties. In Lifschitz, V., and Niemelä, I., eds., *Seventh International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'7)*.

Denecker, M., and Ternovska, E. 2007. A logic of nonmonotone inductive definitions. *Transactions On Computational Logic (TOCL)*. to appear.

Denecker, M.; Bruynooghe, M.; and Marek, V. 2001. Logic programming revisited: Logic programs as inductive definitions. *ACM Transactions on Computational Logic* 2(4):623–654.

Denecker, M. 2000. Extending classical logic with inductive definitions. In Lloyd et al., J., ed., *First International Conference on Computational Logic (CL'2000)*, volume 1861 of *Lecture Notes in Artificial Intelligence*, 703–717. London: Springer.

Gelfond, M., and Lifschitz, V. 1991. Classical negation in logic programs and disjunctive databases. *New Generation Computing* 9:365–387.

Immerman, N. 1986. Relational queries computable in polynomial time. *Information and Control*, 68(1), 86–104.

Libkin, L. 2004. *Elements of Finite Model Theory*. Springer.

Mitchell, D., and Ternovska, E. 2005. A framework for representing and solving np search problems. In *AAAI'05*, 430–435. AAAI Press/MIT Press.

Moschovakis, Y. N. 1974a. *Elementary Induction on Abstract Structures*. North-Holland Publishing Company, Amsterdam- New York.

Moschovakis, Y. N. 1974b. On non-monotone inductive definability. *Fundamenta Mathematicae* 82(39-83).

Schlipf, J. 1995. Complexity and undecidability results in logic programming. *Annals of Mathematics and Artificial Intelligence* 15, 257-288.

Vardi, M. 1982. The complexity of relational query languages. In *14th ACM Symposium on Theory of Computing*, 137-145.

Van Gelder, A.; Ross, K. A.; and Schlipf, J. S. 1991. The well-founded semantics for general logic programs. *Journal of the ACM* 38(3):620–650.

Van Gelder, A. 1993. The alternating fixpoint of logic programs with negation. *Journal of Computer and System Sciences* 47(1):185–221.

Vennekens, J.; Gilis, D.; and Denecker, M. 2006. Splitting an operator: Algebraic modularity results for logics with fixpoint semantics. *ACM Transactions on computational logic (TOCL)* 7(4):765–797.