Annals of Pure and Applied Logic, 78 (1996), 111–125.

Normal Forms for Second-Order Logic over Finite Structures, and Classification of NP Optimization Problems

Thomas Eiter* Georg Gottlob* Yuri Gurevich†

March 6, 1998

Abstract

We start with a simple proof of Leivant's normal form theorem for Σ^1_1 formulas over finite successor structures. Then we use that normal form to prove the following: (i) over all finite structures, every Σ^1_2 formula is equivalent to a Σ^1_2 formula whose first-order part is a boolean combination of existential formulas, and (ii) over finite successor structures, the Kolaitis-Thakur hierarchy of minimization problems collapses completely and the Kolaitis-Thakur hierarchy of maximization problems collapses partially. The normal form theorem for Σ^1_2 fails if Σ^1_2 is replaced with Σ^1_1 or if infinite structures are allowed.

1 Introduction

We consider second-order logic with equality (unless otherwise stated explicitly) and without function symbols of positive arity. Predicates are denoted by capitals and individual variables by lower case letters; a bold face version of a letter denotes a tuple of corresponding symbols. For brevity, we say that a formula Φ *reduces* to a formula Ψ over a class K of structures if the two formulas have the same vocabulary σ and the same free variables and if the two formulas are equivalent at each σ-structure in K.

We recall the definition of Σ^1_k and Π^1_k formulas, $k \geq 1$, on the example when $k = 3$. A Σ^1_3 (respectively, Π^1_3) formula is a second-order formula of the form

$$(\exists \mathbf{S}_1)(\forall \mathbf{S}_2)(\exists \mathbf{S}_3)\psi \quad \text{(respectively } (\forall \mathbf{S}_1)(\exists \mathbf{S}_2)(\forall \mathbf{S}_3)\psi)$$

where ψ is first-order. Classes Σ^0_k and Π^0_k, $k \geq 1$, of first-order formulas are defined similarly. In particular, a Σ^0_2 formula is a first-order formula of the form $(\exists \mathbf{x})(\forall \mathbf{y})\psi$ where ψ is quantifier free.

*Christian Doppler Laboratory for Expert Systems, Information Systems Department, TU Vienna, Paniglgasse 16, A-1040 Wien, Austria. Email: (eiter|gottlob)@dbai.tuwien.ac.at

†Electrical Engineering and CS Department, University of Michigan, Ann Arbor, MI 48109–2122, USA. Email: gurevich@umich.edu. The work of this author is partially supported by NSF and ONR.

It is well-known that every Σ^1_1 formula reduces to a Σ^1_1 formula with first-order part of the form $(\forall \mathbf{x})(\exists \mathbf{y})\varphi(\mathbf{x},\mathbf{y})$ where φ is quantifier-free. The reduction is a simple skolemization [13, Sec. 2.5.2]. It follows that every Σ^1_k (respectively Π^1_k) formula reduces to a Σ^1_k (respectively Π^1_k) formula with only one quantifier alternation in the first-order part.

Leivant found a simpler normal form for Σ^1_1 formulas over finite successor structures: every such formula reduces to a Σ^1_1 formula with universal first-order part [8]. In Section 2, we give a shorter, simpler and more direct proof of this normal form theorem. Leivant's theorem fails in the case of all finite structures. Moreover, let $\Sigma^1_k(bool)$ (respectively $\Pi^1_k(bool)$) be the collection of Σ^1_k (respectively Π^1_k) formulas whose first-order parts are Boolean combinations of existential formulas. We exhibit a Σ^1_1 formula without individual or predicate variables that does not reduce to any $\Sigma^1_1(bool)$ formula.

In Section 3, we use Leivant's to prove our main result, announced in [3]: Over arbitrary finite structures, every Σ^1_2 formula reduces to a $\Sigma^1_2(bool)$ formula. It follows that every Σ^1_k formula (respectively Π^1_k formula), $k \geq 2$, reduces to a $\Sigma^1_k(bool)$ formula (respectively $\Pi^1_k(bool)$ formula).

In Section 4, we exhibit a Σ^1_2 formula which does not reduce over infinite structures to any $\Sigma^1_2(bool)$.

The final Section 5 is devoted to the classification of NP optimization problems. We recall the definition of NP optimization problems and the Kolaitis-Thakur hierarchies of polynomially bounded minimization and maximization problems [6]. In the context of optimization problems, first-order structures serve as inputs to algorithms. In fact, genuine inputs are representations of structures, e.g, as strings. Such representations order the given structure on one way or another. Thus it is most natural to ask what happens to the Kolaitis-Thakur hierarchies in the case of successor (with or without order) structures. We show that in the case of successor structures the minimization hierarchy collapses completely and the maximization hierarchy collapses partially. The case of ordered successor structures does not differ from the case of successor structures for these purposes.

2 Leivant's Normal Form

Fix a binary predicate *Succ* and unary predicates F and L. A *successor structure* is a structure $\mathcal{A}$ such that

- the vocabulary of $\mathcal{A}$ includes the three fixed predicates,
- $\mathcal{A}$ is finite, and
- there exists a linear order $<$ on the universe of $\mathcal{A}$ such that on the expanded structure $(\mathcal{A}, <)$
  - *Succ* is the successor relation of $<$,
  - $F(x)$ is satisfied by and only by the first element, and
  - $L(x)$ is satisfied by and only by the last element.

The expanded structure $(\mathcal{A}, <)$ is an *ordered successor structure.*

Theorem 2.1 (Leivant) *Over successor structures, every Σ_1^1 formula (possibly with free predicate or individual variables) reduces to a Σ_1^1 formula of the form $(\exists \mathbf{T})(\forall \mathbf{x})\psi$ where ψ is quantifier-free.*

Proof. Without loss of generality, the given formula Φ has the form

$$(\exists \mathbf{S})(\forall \mathbf{x})(\exists \mathbf{y})\varphi(\mathbf{x}, \mathbf{y}), \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_n)$ and φ is quantifier free. We prove that the formula

$$\alpha = (\forall \mathbf{x})(\exists \mathbf{y})\varphi(\mathbf{x}, \mathbf{y})$$

is equivalent to a formula

$$\beta = (\exists G)(\forall \mathbf{x}, \mathbf{y}, \mathbf{y}')\psi$$

where G is a $2n$-ary predicate and ψ is quantifier free. The idea is this: β asserts that (i) $G(\mathbf{x}, \mathbf{y})$ holds if and only if $\varphi(\mathbf{x}, \mathbf{z})$ holds for some $\mathbf{z} \leq \mathbf{y}$, and (ii) $G(\mathbf{x}, \mathbf{y})$ holds for the last $\mathbf{y}$.

The order corresponding to *Succ*, F, and L gives rise to the lexicographical order on n-tuples of elements. Obvious quantifier free formulas $Succ^n(\mathbf{y}, \mathbf{y}')$, $F^n(\mathbf{y})$ and $L^n(\mathbf{y})$ describe the successor relation on n-tuples, the first n-tuple and the last n-tuple respectively. The desired ψ is the conjunction of the following formulas:

$$F(\mathbf{y}) \rightarrow [G(\mathbf{x}, \mathbf{y}) \leftrightarrow \varphi(\mathbf{x}, \mathbf{y})],$$
$$Succ(\mathbf{y}, \mathbf{y}') \rightarrow [G(\mathbf{x}, \mathbf{y}') \leftrightarrow (G(\mathbf{x}, \mathbf{y}) \vee \varphi(\mathbf{x}, \mathbf{y}')],$$
$$L(\mathbf{y}) \rightarrow G(\mathbf{x}, \mathbf{y}).$$

We check that α and β are equivalent. Treat free variables of Φ (individual as well as predicate variables) as constants. Suppose that α holds in some successor structure $\mathcal{A}$. For each $\mathbf{x}$, let $M(\mathbf{x})$ be the least $\mathbf{y}$ such that $\varphi(\mathbf{x}, \mathbf{y})$ holds in $\mathcal{A}$. Choose $G(\mathbf{x}, \mathbf{y}) \leftrightarrow \mathbf{y} \geq M(\mathbf{x})$. Clearly $(\mathcal{A}, G) \models (\forall \mathbf{x}, \mathbf{y}, \mathbf{y}')\psi$. Hence, $\mathcal{A} \models \beta$.

Conversely, suppose that β holds in some structure $\mathcal{A}$, and G is a witness to that fact, and $\mathbf{x}$ is an m-tuple of elements of $\mathcal{A}$. Let *end* be the last $\mathbf{y}$; by the third conjunct of ψ, $G(\mathbf{x}, end)$ holds for all $\mathbf{x}$. Let $M(\mathbf{x})$ be the least $\mathbf{y}$ such that $G(\mathbf{x}, \mathbf{y})$ holds. If $M(\mathbf{x})$ is the very first n-tuple then, by the first conjunct of ψ, $\varphi(\mathbf{x}, M(\mathbf{x}))$ holds in $\mathcal{A}$. Otherwise, use the second conjunct of ψ, to establish that $\varphi(\mathbf{x}, M(\mathbf{x}))$ holds in $\mathcal{A}$. Thus, α holds in $\mathcal{A}$. ∎

Remark. Theorem 2.1 can also be derived from Stewart's result that graph 3-colorability is complete for NP via quantifier-free translations with successor [12] and the fact that graph 3-colorability is expressible by a Σ_1^1 formula with universal first-order part.

Theorem 2.2 *There is a Σ^1_1 sentence Φ which does not reduce over finite structures to any $\Sigma^1_1(bool)$ sentence or even to any Σ^1_1 sentence with the first-order part in Σ^0_2.*

Proof. The desired Φ expresses that the universe has an even number of elements. For example, Φ may assert the existence of an equivalence relation such that every equivalence class contains precisely two elements.

By contradiction, suppose that Φ reduces over finite structures to a sentence $\Psi = (\exists \mathbf{X})(\exists \mathbf{x})(\forall \mathbf{y})\psi$ where ψ is quantifier free. Let k be the number of existential individual quantifiers in φ and U be a set of even cardinality with $\|U\| > k$. Clearly, $U \models \Psi$. Therefore, for some tuple $\mathbf{X}_0$ of appropriate relations, $\langle U, \mathbf{X}_0\rangle \models (\exists \mathbf{x})(\forall \mathbf{y})\varphi$.

Choose k appropriate witnesses and let $V \subseteq U$ contain all k witnesses and be of odd cardinality. It is easy to see that $V \models \Psi$ which is impossible. ∎

Remark. It is shown in [7] that the fragment of Σ^1_1 with first-order parts in Σ^0_2 has a 0-1 law on finite structures. This gives another proof of Theorem 2.2.

3 A Σ^1_2 Normal Form

Now we consider arbitrary finite structures. Let $\Sigma^0_1(bool)$ be the collection of Boolean combinations of first-order existential formulas.

Lemma 3.1 *Let $<$ and Succ be binary predicates, and let F, L, Z be unary predicates. There exists a $\Sigma^0_1(bool)$ formula $SUCCORD(<, Succ, F, L, Z)$ such that the formula $(\forall Z)SUCCORD(<, Succ, F, L, Z)$ asserts that $<$, Succ, F, and L give an ordered successor structure.*

Proof. Define the following formulas:

- $$LINORD(<) = (\forall x,y,z)((x<y \wedge y<z) \rightarrow x<z) \wedge (\forall x)\neg(x<x) \wedge (\forall x,y)(x=y \vee x<y \vee y<x),$$

asserting that $<$ is a linear order;

- $FIRST(<, F) = (\forall x,y)(F(x) \rightarrow \neg(y<x)) \wedge (\exists x)(F(x))$,

asserting that F describes the smallest element according to $<$;

- $LAST(<, L) = (\forall x,y)(L(x) \rightarrow \neg(x<y)) \wedge (\exists x)(L(x))$,

asserting that L describes the greatest element according to $<$;

- $SUCCESSOR_1(<, Succ) =$

$$(\forall x,y,z)[Succ(x,y) \rightarrow (x<y \wedge \neg(x<z \wedge z<y))];$$

and

- $SUCCESSOR_2(Succ, L, Z)$ =

$$[(\exists u)(Z(u) \wedge \neg L(u)) \rightarrow (\exists u, v)(Z(u) \wedge Succ(u, v))] .$$

Note that the formula $(\forall Z) SUCCESSOR_2$ implies that every element, except the greatest, has a successor. To see that, consider the case when Z is of cardinality one. The second-order quantification allows us to avoid alternating quantifiers in the first-order formula.

The desired

$$\begin{aligned} SUCCORD(<, Succ, F, L, Z) \;\; = \;\; & LINORD(<) \wedge FIRST(<, F) \wedge LAST(<, L) \wedge \\ & SUCCESSOR_1(<, Succ) \wedge SUCCESSOR_2(Succ, L, Z) \end{aligned}$$

∎

Corollary 3.2 *Over finite structures, every Σ_1^1 formula (possibly with free predicate or individual variables) reduces to a $\Pi_2^1(bool)$ formula.*

Proof. Let $SUCCORD(<, Succ, F, L, Z)$ be as above and Ψ be any Σ_1^1 formula. Since every nonempty finite set supports an ordered successor structure, Ψ is equivalent (we consider only finite structures here) to

$$(\forall <, Succ, F, L)(\, [(\forall Z)(SUCCORD(<, Succ, F, L, Z)] \rightarrow \Psi \,) \tag{2}$$

By Theorem 2.1, over successor structures, Ψ reduces to a formula $(\exists \mathbf{S})(\forall \mathbf{x})\varphi$ where φ is quantifier-free. Clearly, formula (2) is equivalent to

$$(\forall <, Succ, F, L)(\, [(\forall Z)(SUCCORD(<, Succ, F, L, Z)] \rightarrow (\exists \mathbf{S})(\forall \mathbf{x})\varphi \,) \tag{3}$$

which is equivalent to

$$(\forall <, Succ, F, L)(\exists Z, \mathbf{S})[SUCCORD(<, Succ, F, L, Z) \rightarrow (\forall \mathbf{x})\varphi] \tag{4}$$

∎

Theorem 3.3 *For $k \geq 2$, every Σ_k^1 formula (respectively Π_k^1 formula) reduces over finite structures to a $\Sigma_k^1(bool)$ formula (respectively $\Pi_k^1(bool)$ formula).*

Proof. It is sufficient to prove the theorem for Π_2^1 formulas. All other cases follow trivially.

Let $\Psi = (\forall \mathbf{P})(\exists \mathbf{Q})\varphi$, where φ is first-order. By Corollary 3.2, the Σ_1^1 formula $\Phi = (\exists \mathbf{Q})\varphi$ reduces over finite structures to some $\Pi_2^1(bool)$ formula Φ^*. Thus, Ψ reduces over finite structures to $(\forall \mathbf{P})\Phi^*$, which is a $\Pi_2^1(bool)$ formula. ∎

Notice that the theorem remains valid for logic without equality because the fact that a free binary predicate is equality can be expressed by a Π_1^1 formula without alternating first-order quantifiers.

Remark. A Σ^1_k formula does not necessarily reduce over finite structures to a Σ^1_k formula whose first-order part is purely universal or existential. In fact, even a $\Sigma^0_1(bool)$ formula does not necessarily reduce over finite structures to any Σ^1_k formula whose first-order part is purely universal or existential. The reason is that every Σ^1_k sentence whose first-order part is purely universal or existential is preserved under submodels or extensions, respectively. Indeed, each universal first-order sentence is preserved under submodels (see [2]). An easy induction on k shows that each Σ^1_k formula with universal first-order part is preserved under submodels. On the other hand, it is easy to construct a $\Sigma^0_1(bool)$ sentence that is not preserved under submodels or extensions.

4 Limited expressiveness of the $\Sigma^1_2(bool)$ fragment

In the previous section, we considered normal forms for second-order logic over finite structures. Now, let us consider second-order logic over arbitrary structures.

Even if the syntactical form of the formulas in $\Sigma^1_2(bool)$ is rather simple, many interesting properties can be expressed within this fragment. For example, the following properties of sets can be expressed.

Infinity:

$$(\exists <)(\forall X)\ [\mathit{LINORD}(<) \wedge [(\exists x)(X(x)) \rightarrow (\exists x, y)(X(x) \wedge y < x)]\,], \tag{5}$$

where $\mathit{LINORD}(<)$ is the universal first-order formula, defined above, that asserts that $<$ is a linear order.

Countability (that is, finite or infinite countability):

$$(\exists <, \mathit{Succ}, F)(\forall X)\ [\,\Gamma(<, \mathit{Succ}, F) \wedge \mathit{INDUCTION}(F, \mathit{Succ}, X)\,], \tag{6}$$

where $\Gamma(<, \mathit{Succ}, F)$ is the conjunction of the formulas $\mathit{LINORD}(<)$, $\mathit{FIRST}(<, F)$, and $\mathit{SUCCESSOR}_1(<, \mathit{Succ})$, defined above and the formula

$$\mathit{INDUCTION}(F, \mathit{Succ}, X) \ =$$

$$[(\forall x)(F(x) \rightarrow X(x)) \wedge (\forall x, y)(\mathit{Succ}(x, y) \wedge X(x) \rightarrow X(y))] \rightarrow (\forall x)X(x)$$

Finiteness:

$$(\exists <, \mathit{Succ}, F, L)(\forall X)\ [\,\Gamma(<, \mathit{Succ}, F) \wedge \mathit{LAST}(<, L)\ \wedge \mathit{INDUCTION}(F, \mathit{Succ}, X)\,] \tag{7}$$

where $\mathit{LAST}(<, L)$, $\Gamma(<, \mathit{Succ}, F)$ and $\mathit{INDUCTION}(F, \mathit{Succ}, X)$ are as above.

Infinite countability ($\aleph_0$) is the conjunction of (5) and (6) and thus can be expressed in $\Sigma^1_2(bool)$.

The question arises whether all Σ^1_2 properties over arbitrary structures can be expressed by $\Sigma^1_2(bool)$ formulas. This is not the case.

Theorem 4.1 *There exists a Σ^1_2 formula that is not equivalent to any $\Sigma^1_2(bool)$ formula or even to any Σ^1_2 formula with first-order part in Π^0_2.*

Proof. It suffices to prove the claim when we allow formulas to use the standard arithmetical operations and restrict attention to structures on the set of natural numbers where the arithmetical operations have their usual interpretations. Indeed, suppose that Φ is a Σ^1_2 formula that is not equivalent to any Σ^1_2 formula with first-order part in Π^0_2 over the standard arithmetical structure $\mathbf{A}$ (that is over the class of structures described above). Let $\mathbf{A}'$ be the relational structure obtained from $\mathbf{A}$ by replacing the arithmetical operations with their graphs, e.g. the successor operation *Succ* is replaced with the binary relation $\{(x,y) : y = Succ(x)\}$. In the obvious way, define the notion that a formula about $\mathbf{A}$ is equivalent to a formula about $\mathbf{A}'$. Transform Φ into an equivalent relational Σ^1_2 formula Φ'. By contradiction suppose that Φ' is logically equivalent to a Σ^1_2 formula Ψ' with first-order part in Π^0_2. Then Φ' and Ψ' are equivalent over $\mathbf{A}'$. Transform Ψ' to an equivalent Σ^1_2 formula Ψ with first-order part in Π^0_2 about $\mathbf{A}$. Then Φ and Ψ are equivalent, which gives the desired contradiction.

It is well known (see e.g. [13, Sec. 3.2]) that there exists a Σ^1_2 formula Ψ without free individual variables that is not equivalent to any Σ^1_1 formula over $\mathbf{A}$. Thus it suffices to prove that every arithmetical Σ^1_2 formula with first-order part in Π^0_2 is equivalent to an arithmetical Σ^1_1 formula. This follows from the following lemma.

Lemma 4.2 *Over $\mathbf{A}$, every formula*

$$\Phi = (\forall \mathbf{T})(\exists \mathbf{y})\alpha(\mathbf{T}, \mathbf{y}),$$

where α is quantifier free, is equivalent to a first-order formula ψ.

We note that α may have free predicate and individual variables. Lemma 4.2 is not new to experts. In fact, it remains true if bounded universal quantification is allowed in the first-order part of Φ and if it is required that all universal quantors in ψ are bounded. Barwise attributes the stronger result to Kreisel and proves a generalization of it to countable admissible sets in [1]. For reader's convenience, we give a direct proof of our lemma.

Proof of Lemma 4.2. Since every recursively enumerable relation can be expressed with a first-order existential formula [9] over $\mathbf{A}$, there are first-order existential formulas $G_k(x, y_1, \ldots, y_k)$ such that, for every x, there exists a unique sequence $(y_1, \ldots, y_k)$ with $\mathbf{A} \models G_k(x, y_1, \ldots, y_k)$, and, for every sequence $(y_1, \ldots, y_k)$, there is a unique x with $\mathbf{A} \models G_k(x, y_1, \ldots, y_k)$ [5]. Fix appropriate formulas G_k. To make our intentions clearer we write $x = Code(y_1, \ldots, y_k)$ instead of $G_k(x, y_1, \ldots, y_k)$.

Without loss of generality $\mathbf{T}$ is a single unary predicate T. The reason is that the sequence $\mathbf{T}$ of predicates $(T_1, \ldots, T_m)$ can be appropriately encoded by a single unary predicate T. We assume without loss of generality that all predicates T_i have the same arity and illustrate the coding procedure on an example. Suppose that α contains only two atomic formulas involving $\mathbf{T}$, namely $\beta = T_1(u_1, u_2)$ and $\gamma = T_2(v_1, v_2)$. Let z_β and z_γ be fresh variables and α^* be the result of replacing β with $T(z_\beta)$ and γ with $T(z_\gamma)$ in $\alpha(\mathbf{T}, \mathbf{y})$. Given any $\mathbf{T}$, let the desired T contain a number x if and only if either there are u_1, u_2 such that $x = Code(1, u_1, u_2)$ and $T_1(u_1, u_2)$ holds or else there are v_1, v_2 such that $x = Code(2, v_1, v_2)$ and $T_2(v_1, v_2)$ holds. It is easy to see that $(\exists \mathbf{y})\alpha(\mathbf{T}, \mathbf{y})$ is equivalent to

$$(\exists \mathbf{y})(\exists z_\beta, z_\gamma)[\, z_\beta = Code(1, u_1, u_2) \wedge z_\gamma = Code(2, v_1, v_2) \wedge \alpha^* \,]$$

and therefore Φ is equivalent to

$$(\forall T)(\exists \mathbf{y})(\exists z_\beta, z_\gamma)[\, z_\beta = Code(1, u_1, u_2) \wedge z_\gamma = Code(2, v_1, v_2) \wedge \alpha^* \,].$$

Now consider the tree B of binary strings where the empty string is the root and each node ℓ has two children $\ell 0$ and $\ell 1$. For each T, let *Branch*(T) be the infinite branch of nodes $l_0 l_1 \cdots l_{d-1}$ such that the restriction $T|d = T \cap \{0, \ldots, d-1\}$ is equal to $\{i : l_i = 1\}$.

For brevity we write $x > \mathbf{y}$ (or $\mathbf{y} < x$) to mean that $x > y_i$ for every component y_i of $\mathbf{y}$. Fix all free variables of Φ. We construct a subset P of B such that, for every unary relation T, $(\exists \mathbf{y})\alpha(T, \mathbf{y})$ holds if and only if *Branch*(T) intersects P. For each $\mathbf{y}$, let $c(\mathbf{y})$ be the maximum among the arguments of predicate T that occur in the quantifier-free formula $\alpha(T, \mathbf{y})$. Clearly,

(*) for every $d > c(\mathbf{y})$, $\alpha(T, \mathbf{y})$ is equivalent to $\alpha(T|d, \mathbf{y})$.

Put a string $l_0 \cdots l_{d-1}$ into P if and only if there exists $\mathbf{y}$ such that $c(\mathbf{y}) < d$ and the set $T' = \{i : l_i = 1\}$ satisfies $\alpha(T', \mathbf{y})$. We fix a unary relation T and check that $(\exists \mathbf{y})\alpha(T, \mathbf{y})$ holds if and only if *Branch*(T) intersects P. First suppose that $\alpha(T, \mathbf{y})$ holds for some $\mathbf{y}$ and set $d = c(\mathbf{y}) + 1$, $T' = T|d$. By (*), $\alpha(T', \mathbf{y})$ holds. By the definition of P, the unique string ℓ of length d in *Branch*(T') belongs to P. Of course $\ell \in$ *Branch*(T) as well. Thus, *Branch*(T) intersects P. Second suppose that a string $\ell = l_0 \cdots l_{d-1}$ belongs to *Branch*(T) and P, and let $T' = T|d = \{i : l_i = 1\}$. Since $\ell \in P$, we have that, for some $\mathbf{y}$ with $c(\mathbf{y}) < d$, $\alpha(T', \mathbf{y})$ holds and therefore $\alpha(T|d, \mathbf{y})$ holds. By (*), $(\exists \mathbf{y})\alpha(T, \mathbf{y})$ holds.

Let B' be the subtree of B obtained by removing all strings ℓ such that a proper prefix of ℓ belongs to P. Notice that every maximal string ℓ in B' (the maximality of ℓ means that neither $\ell 0$ nor $\ell 1$ belongs to B') belongs to P. We show that B' has no infinite branches if and only if

(**) $(\exists k)(\forall T \subseteq \{0, \ldots, k-1\})(\exists \mathbf{y})[\, c(\mathbf{y}) < k \wedge \alpha(T, \mathbf{y}) \,]$.

First assume that B' has no infinite branch. By König's lemma, B' is finite. Let k be one plus the maximal length of a string in B'. Every string ℓ of length k has a proper prefix in P; otherwise ℓ would belong to B' which is impossible. Now let T be an arbitrary subset of $\{0, \ldots, k-1\}$ and ℓ be the string $l_0 \cdots l_{k-1}$ such that $T = \{i : l_i = 1\}$. Since $|\ell| = k$, there is $d < k$ such that the string $l_0 \ldots l_{d-1}$ belongs to P. By the definition of P, there exists $\mathbf{y}$ such that $c(\mathbf{y}) < d$ and the set $T' = \{i : i < d \text{ and } l_i = 1\}$ satisfies $\alpha(T', \mathbf{y})$. Clearly, $T' = T|d$. Now use (*) to establish (**).

Second assume (**) and fix an appropriate k. We prove that B' has no infinite branch. By contradiction suppose that B' has an infinite branch

$l_0,\ l_0 l_1,\ l_0 l_1 l_2,\ \ldots$

and define $T = \{i : i < k \text{ and } l_i = 1\}$. By (**), there exists $\mathbf{y}$ such that $c(\mathbf{y}) < k$ and $\alpha(T, \mathbf{y})$. By the definition of P, the string $l_0 \cdots l_{k-1}$ belongs to P. This contradicts the fact that $l_0 \cdots l_k$ belongs to B'.

We have that Φ holds if and only if every *Branch*(T) intersects P if and only if B' has no infinite branches if and only (**) holds. It remains to notice that (**) is equivalent to a first-order formula of the form.

$$(\exists k)(\forall t)(\exists \mathbf{y})\alpha^*(k, t, \mathbf{y}).$$

α^* can be obtained from $[c(\mathbf{y}) < k \wedge \alpha(T, \mathbf{y})]$ by replacing each $T(z)$ with a formula saying that $z < k$ and there exist p and m such that p is prime, m is not divisible by p and $t = p^z m$.

Lemma 4.2 is proved and thus Theorem 4.1 is proved. ▮

5 On Classification of NP Optimization Problems

For brevity, we write Σ_k for Σ_k^0 and the same for Π. An *NP minimization problem* [10, 6] is given by a tuple $\mathcal{M} = (\mathcal{I}_\mathcal{M}, \mathcal{F}_\mathcal{M}, f_\mathcal{M})$ such that

- $\mathcal{I}_\mathcal{M}$ is a set of input instances, which is assumed to be recognizable in polynomial time.
- $\mathcal{F}_\mathcal{M}(I)$ is a set of so-called feasible solutions for the input I.
- $f_\mathcal{M}$ is a polynomial time computable function, called the *objective function*, which takes positive integer values and is defined on pairs I, T, where I is an input instance and T is a feasible solution of I.
- The following decision problem is in NP: Given $I \in \mathcal{I}_\mathcal{M}$ and an integer k, does there exist a feasible solution $T \in \mathcal{F}_\mathcal{M}(I)$ such that $f_\mathcal{M}(I, T) \le k$?

Define $opt_\mathcal{M}(I) = \min_T f_\mathcal{M}(I, T)$.

NP maximization problems are defined similarly. In particular, we have:

- The following decision problem is in NP: Given $I \in \mathcal{I}_\mathcal{M}$ and an integer k, does there exist a feasible solution $T \in \mathcal{F}_\mathcal{M}(I)$ such that $f_\mathcal{M}(I, T) \ge k$?
- $opt_\mathcal{M}(I) = \max_T f_\mathcal{M}(I, T)$.

A well-known minimization problem is MIN CHROMATIC NUMBER, where the instances $\mathcal{I}_\mathcal{M}$ are finite graphs, feasible solutions for graph G are colorings of the vertices of G such that no two adjacent vertices have the same color, and the function $f_\mathcal{M}(G, T)$ is the number of colors used in T. An example of a maximization problem is MAX CONNECTED COMPONENT: find the size of a largest connected component of an undirected finite graph.

In the spirit of Fagin's logical characterization of NP [4], characterizations of NP optimization problems in terms of logical definability have been given e.g. in [11, 10, 6]. An NP optimization problem is said to be *polynomially bounded* if there is a polynomial p such that

$$opt_{\mathcal{M}}(I) \leq p(|I|), \quad \text{for all } I \in \mathcal{I}_{\mathcal{M}},$$

where $|I|$ is the length of the input I. MIN $\mathcal{PB}$ (respectively MAX $\mathcal{PB}$) denotes the class of all polynomially bounded NP minimization (respectively maximization) problems.

We restrict attention to optimization problems whose inputs are finite structures of a fixed vocabulary. As shown in [6], an NP optimization problem $\mathcal{M}$ with finite structures $\mathcal{A}$ over a vocabulary σ is polynomially bounded if and only if there is a first-order formula $\varphi(\mathbf{w}, \mathbf{S})$ with predicates among those of σ and $\mathbf{S}$ such that for every instance $\mathcal{A}$ of $\mathcal{M}$,

$$opt_{\mathcal{M}}(\mathcal{A}) = |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S}) \,\}|.$$

(Here and thereafter it is assumed that the universe of $\mathcal{A}$ contains at least two elements.) Moreover, in [6] hierarchies of classes of NP optimization problems have been analyzed.

We start with minimization problems. Denote by MIN Σ_k (respectively MIN Π_k) the class of NP minimization problems definable by a Σ_k (respectively Π_k) formulas, $k \geq 0$. According to [6],

$$\text{MIN}\,\Sigma_0 = \text{MIN}\,\Sigma_1 \subset \text{MIN}\,\Pi_1 = \text{MIN}\,\Sigma_2 = \text{MIN}\,\mathcal{PB}. \tag{8}$$

We show that for minimization problems over successor structures, MIN $\mathcal{PB}$ is contained in MIN Σ_1, so that all polynomially bounded minimization problems can be defined with a quantifier-free first-order formula. This fits Kolaitis and Thakur's observation that "... the pattern of the quantifier prefix does not impact on the approximability of minimization problems" [6, p.348].

Theorem 2.1 allows one to strengthen Fagin's theorem [4], that a class of finite structures closed under isomorphisms is recognizable in NP if and only if it is definable in existential second-order logic, as follows.

Lemma 5.1 *Let K be a class of successor structures of vocabulary σ closed under isomorphisms. Then, K is NP if and only if K is definable by a Σ^1_1 σ-formula of the form $(\exists \mathbf{S})(\forall \mathbf{x})\psi$, where ψ is quantifier-free.*

Proof. Use Theorem 2.1. ∎

Theorem 5.2 *Let $\mathcal{M}$ be an NP minimization problem over successor structures of vocabulary σ. Then, $\mathcal{M}$ is in* MIN $\mathcal{PB}$ *if and only if there is an existential first-order formula $\varphi(\mathbf{w}, \mathbf{S})$ with predicates among those in σ and $\mathbf{S}$ such that for every instance $\mathcal{A}$ of $\mathcal{M}$,*

$$opt_{\mathcal{M}}(\mathcal{A}) = \min_{\mathbf{S}} |\{\mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S})\}|.$$

Thus,

$$\text{MIN}\,\mathcal{PB} = \text{MIN}\,\Sigma_1 = \text{MIN}\,\Sigma_k = \text{MIN}\,\Pi_k, \quad \textit{for all } k \geq 1.$$

Proof. The proof is analogous to the proof of Theorem 3 in [6], but uses Lemma 5.1 instead of Fagin's theorem.

The if direction is obvious. For the only if direction, let m be a positive integer such that for any instance $\mathcal{A}$, we have that $opt_{\mathcal{M}}(\mathcal{A}) \leq \|\mathcal{A}\|^m$, where $\|\mathcal{A}\|$ is the size of the structure $\mathcal{A}$ (i.e., the size of the universe of $\mathcal{A}$).

Consider the following NP problem Q: Given a finite σ-structure $\mathcal{A}$ and an m-ary relation W on the universe A of $\mathcal{A}$, is there a feasible solution T for $\mathcal{A}$ such that $f_{\mathcal{M}}(\mathcal{A}, T) \leq |W|$? Here, $f_{\mathcal{M}}$ is the objective function of $\mathcal{M}$ and $|W|$ is the cardinality of W. By Lemma 5.1, there is an existential second-order formula $(\exists \mathbf{T})(\forall \mathbf{x})\psi(\mathbf{T}, W, \mathbf{x})$, where ψ is quantifier-free, such that the expanded structure $(\mathcal{A}, W)$ is a YES instance of Q if and only if $(\mathcal{A}, W) \models (\exists \mathbf{T})(\forall \mathbf{x})\psi$. Since the minimization problem $\mathcal{M}$ is bounded by $\|\mathcal{A}\|^m$, we have that

$$opt_{\mathcal{M}}(\mathcal{A}) = \min_{\mathbf{T}, W}\{\, |W| : (\mathcal{A}, \mathbf{T}, W) \models (\forall \mathbf{x})\psi(\mathbf{T}, W, \mathbf{x})\,\}.$$

It follows that

$$\begin{aligned} opt_{\mathcal{M}}(\mathcal{A}) &= \min_{\mathbf{T}, W} |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{T}, W) \models (\forall \mathbf{x})(\psi(\mathbf{T}, W, \mathbf{x})) \rightarrow W(\mathbf{w})\,\}| \\ &= \min_{\mathbf{T}, W} |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{T}, W) \models (\exists \mathbf{x})(\neg\psi(\mathbf{T}, W, \mathbf{x})) \vee W(\mathbf{w})\,\}| \end{aligned}$$

Let $\mathbf{S}$ denote the sequence $(\mathbf{T}, W)$ and let $\varphi(\mathbf{w}, \mathbf{S})$ be a Σ_1 formula logically equivalent to $(\exists \mathbf{x})(\neg\psi(\mathbf{T}, W, \mathbf{x})) \vee W(\mathbf{w})$. It follows that

$$opt_{\mathcal{M}}(\mathcal{A}) = \min_{\mathbf{S}} |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S})\,\}|.$$

This proves the theorem. ∎

Combining Theorem 5.2 with Kolaitis and Thakur's result that $\text{MIN}\,\Sigma_0 = \text{MIN}\,\Sigma_1$, we get:

Corollary 5.3 *In the case of successor structures, a minimization problem is polynomially bounded if and only if it is definable with a quantifier-free first order formula, that is,*

$$\text{MIN}\,\mathcal{PB} = \text{MIN}\,\Sigma_0 = \text{MIN}\,\Sigma_k, \quad \textit{for all } k.$$

Remark. The classes $\text{MIN}\,\Sigma_1$ and $\text{MIN}\,\Sigma_2$ are separated in [6] by showing that the problem MIN CHROMATIC NUMBER is in $\text{MIN}\,\Sigma_2$ but not in $\text{MIN}\,\Sigma_1$ (Theorem 4, Part B). The proof in [6] uses the fact that a graph G obtained by taking the direct sum of graphs H_1 and H_2 without common vertices is an extension of both H_1 and H_2. In our case, each graph has to have, in addition to the edge relation, a successor relation and the corresponding First and Last relations. In fact, G can extend neither H_1 nor H_2. Indeed, assume G extends, say, H_1. Check by induction on k, that kth element of H_1 is the kth element of G. Further, the last element of H_1 is the last element of G. Thus G contains no elements of H_2 which is impossible. Thus, the proof in [6] fails for successor structures.

We turn our attention to maximization problems. Denote by $\mathrm{MAX}\,\Sigma_k$ (respectively $\mathrm{MAX}\,\Pi_k$) the class of NP maximization problems definable by Σ_k (respectively Π_k) formulas, $k \geq 0$. According to [6],

$$\mathrm{MAX}\,\Sigma_0 \subset \mathrm{MAX}\,\Sigma_1 \subset \mathrm{MAX}\,\Sigma_2 = \mathrm{MAX}\,\Pi_1 \subset \mathrm{MAX}\,\Pi_2 = \mathrm{MAX}\,\mathcal{PB}. \tag{9}$$

Over successor structures, $\mathrm{MAX}\,\Pi_2$ collapses to $\mathrm{MAX}\,\Pi_1$.

Theorem 5.4 *Let $\mathcal{M}$ be a maximization problem over successor structures of vocabulary σ. Then, $\mathcal{M}$ is in $\mathrm{MAX}\,\mathcal{PB}$ if and only if there is a universal first-order formula $\varphi(\mathbf{w}, \mathbf{S})$ with predicate symbols among those in σ and $\mathbf{S}$ such that for every instance $\mathcal{A}$ of $\mathcal{M}$,*

$$opt_{\mathcal{M}}(\mathcal{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S})\}|.$$

Thus

$$\mathrm{MAX}\,\mathcal{PB} = \mathrm{MAX}\,\Pi_1 = \mathrm{MAX}\,\Sigma_k = \mathrm{MAX}\,\Pi_k, \quad \textit{for all } k \geq 2.$$

Proof. The proof is similar to the proof of Theorem 1 in [6]. Follow the same arguments as in Theorem 5.2 to show that if $\mathcal{M}$ is a polynomially bounded NP maximization problem over finite successor structures, then there is a Π_1 formula $\psi(\mathbf{T}, W)$ such that

$$opt_{\mathcal{M}}(\mathcal{A}) = \max_{\mathbf{T},W} \{\, |W| : (\mathcal{A}, \mathbf{T}, W) \models \psi(\mathbf{T}, W)\,\},$$

or equivalently,

$$opt_{\mathcal{M}}(\mathcal{A}) = \max_{\mathbf{T},W} |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{T}, W) \models \psi(\mathbf{T}, W, \mathbf{x}) \wedge W(\mathbf{w})\,\}|.$$

Let $\mathbf{S}$ denote the sequence $(\mathbf{T}, W)$ and let $\varphi(\mathbf{w}, \mathbf{S})$ be a Π_1 formula logically equivalent to $\psi(\mathbf{T}, W) \wedge W(\mathbf{w})$. It follows that

$$opt_{\mathcal{M}}(\mathcal{A}) = \max_{\mathbf{S}} |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S})\,\}|.$$

∎

Remark. The classes $\mathrm{MAX}\,\Pi_1$ and $\mathrm{MAX}\,\Pi_2$ are separated in [6] by showing that the problem MAX CONNECTED COMPONENT is in $\mathrm{MAX}\,\Pi_2$ but not in $\mathrm{MAX}\,\Pi_1$ (Theorem 2, part B). The proof in [6] fails for successor structures, since it uses the fact that certain graphs H_i, obtained by removing vertices a_i from the input graph G, are substructures of G.

Contrary to the case of minimization problems, not every polynomially bounded NP maximization problem can be defined over successor structures with a quantifier-free first-order formula. In fact, the two leftmost containments of the hierarchy (9) are also strict for successor structures and even for ordered successor structures. We show this by exhibiting two NP maximization problems, EVEN and EMPTY, which separate the classes $\mathrm{MAX}\,\Sigma_0$, $\mathrm{MAX}\,\Sigma_1$, and $\mathrm{MAX}\,\Sigma_2$ in the case of ordered successor structures.

- EVEN: The instances are ordered successor structures without any additional predicates (that is of vocabulary $\{<, Succ, F, L\}$), the only feasible solution for a structure $\mathcal{A}$ is the empty set, and the function $f_{\text{EVEN}}(\mathcal{A}, \emptyset)$ equals 1 if $\|\mathcal{A}\|$ is even and equals 2 otherwise.
- EMPTY: The instances are ordered successor structures of vocabulary $\{<, Succ, F, L, P\}$ where P is a unary predicate. The only feasible solution T for a structure $\mathcal{A}$ is the relation P, and $f_{\text{EMPTY}}(\mathcal{A}, P)$ equals 1 if P is empty and equals 2 otherwise.[1]

Note that both problems are easily solvable in polynomial time.

Theorem 5.5 *(i) EVEN is in* $\text{MAX}\,\Sigma_2$ *but not in* $\text{MAX}\,\Sigma_1$*, and (ii) EMPTY is in* $\text{MAX}\,\Sigma_1$ *but not in* $\text{MAX}\,\Sigma_0$*.*

Proof.
Part (i). Clearly, EVEN is polynomially bounded, and hence, by Theorem 5.4, it is in $\text{MAX}\,\Sigma_2$. By contradiction assume that EVEN is in $\text{MAX}\,\Sigma_1$. There exists a formula $(\exists \mathbf{x})\varphi(\mathbf{w}, \mathbf{x}, \mathbf{S})$ where φ is quantifier-free such that, for all $\mathcal{A}$,

$$opt_{\text{EVEN}}(\mathcal{A}) = |\{\, \mathbf{w} : (\mathcal{A}, \mathbf{S}) \models (\exists \mathbf{x})\varphi(\mathbf{w}, \mathbf{x}, \mathbf{S}) \,\}|.$$

Let k be the number of variables in φ and let $\mathcal{A}$ be such that $\|\mathcal{A}\| > 2k$ and $\|\mathcal{A}\|$ is odd. Let $\mathbf{S}^*$ be such that $|\{\mathbf{w} : (\mathcal{A}, \mathbf{S}^*) \models (\exists \mathbf{x})\varphi\,\}|$ is maximal. Let $\mathbf{v}$ denote the sequence $\mathbf{wx}$. Since $opt_{\text{EVEN}}(\mathcal{A}) = 2$, there exist on $\mathcal{A}$ two tuples $\mathbf{v}_1$, $\mathbf{v}_2$ different on their $\mathbf{w}$-parts such that $(\mathcal{A}, \mathbf{S}^*) \models \varphi(\mathbf{v}_i, \mathbf{S}^*)$ for $i = 1, 2$. Since $\|\mathcal{A}\| > 2k$, there exists an element a of $\mathcal{A}$ such that a does not occur in $\mathbf{v}_1$ or $\mathbf{v}_2$. Let b be an element adjacent to a and let $\mathcal{A}'$ be the structure obtained from $\mathcal{A}$ by inserting a new element between a and b. Each $\varphi(\mathbf{v}_i, \mathbf{S}^*)$ is satisfied in $\mathcal{A}'$. This implies $opt_{\text{EVEN}}(\mathcal{A}') \geq 2$. However, since the size of the universe of $\mathcal{A}'$ is even, $opt_{\text{EVEN}}(\mathcal{A}') = 1$. Thus we arrived at the desired contradiction.

Part (ii). Clearly, EMPTY is in $\text{MAX}\,\Sigma_1$; it can be defined e.g. by formula

$$(\exists x)[\,(F(w) \wedge P(x)) \vee L(w)\,]$$

which is equivalent to $[F(w) \wedge (\exists x)P(x)] \vee L(w)$. (Recall that structures $\mathcal{A}$ are supposed to have at least two elements.)

By contradiction suppose that EMPTY is in $\text{MAX}\,\Sigma_0$, so that some quantifier-free formula $\varphi(\mathbf{w}, \mathbf{S})$ defines EMPTY:

$$opt_{\text{EMPTY}}(\mathcal{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : (\mathcal{A}, \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S})\}|.$$

Let k be the number of distinct variables in φ, and let $\mathcal{A}$ be an instance of EMPTY such that $\|\mathcal{A}\| = n > 20k$ and $P = \{a_{10k}\}$, where $a_1, a_2, \ldots, a_n$ are the elements of $\mathcal{A}$ in the order

[1]It may seem more natural to use 0 rather than 2 as a value for f_{EMPTY} and f_{EVEN}, but the objective functions take only positive values.

of $\mathcal{A}$. Then, $opt_{\text{EMPTY}}(\mathcal{A}) = 2$. The idea is to modify $\mathcal{A}$ and to find some $\mathbf{S}$ on the new structure $\mathcal{A}'$ such that there are more than two tuples w satisfying $\varphi(\mathbf{w}, \mathbf{S})$ in $(\mathcal{A}', \mathbf{S})$.

Let $\mathbf{S}_0$ be such that $|\{\, \mathbf{w} : (\mathcal{A}, \mathbf{S}_0) \models \varphi(\mathbf{w}, \mathbf{S}_0) \,\}|$ is maximal. Then there exist distinct tuples $\mathbf{w}_1, \mathbf{w}_2$ such that $(\mathcal{A}, \mathbf{S}_0) \models \varphi(\mathbf{w}_i, \mathbf{S}_0)$ for $i = 1, 2$. Elements that occur in $\mathbf{w}_1$ or $\mathbf{w}_2$ will be called *red*. It is easy too see that a_{10k} is red. Otherwise let $\mathcal{A}_0$ be the structure obtained from $\mathcal{A}$ by removing a_{10k} from P. We have $(\mathcal{A}_0, \mathbf{S}_0) \models \varphi(\mathbf{w}_i, \mathbf{S}_0)$ for $i = 1, 2$ and thus $opt_{\text{EMPTY}}(\mathcal{A}_0) \geq 2$. It is clear, however, that $opt_{\text{EMPTY}}(\mathcal{A}_0) = 1$.

Obviously there is $i < 10k$ such that neither a_i nor a_{i+1} is red. Let $[a_\ell, a_u]$ be the first contiguous red segment after a_{i+1}. In other words, a_ℓ is the least red element $> a_{i+1}$ and a_{u+1} is the least non-red element $> a_\ell$. Clearly $u < 10k + 2k < n$.

Let $\mathcal{A}'$ be a structure obtained from $\mathcal{A}$ as follows: Add a segment $a'_\ell, a'_{\ell+1}, \ldots, a'_u$ of new elements between a_i and a_{i+1} and put a'_{10k} into P if $l \leq 10k \leq u$. Let $\mathbf{S}_1$ be obtained from $\mathbf{S}_0$ by replacing elements a_j with the corresponding elements a'_j, $\ell \leq j \leq u$. Let $\mathbf{S}_2$ be the union of $\mathbf{S}_0$ and $\mathbf{S}_1$.

It is clear that $(\mathcal{A}', \mathbf{S}_2) \models \varphi(\mathbf{w}_i, \mathbf{S}_2)$. Let $\mathbf{w}'_i$ be the tuple obtained from $\mathbf{w}_i$ by replacing elements a_j with the corresponding elements a'_j, $\ell \leq j \leq u$. We have $(\mathcal{A}', \mathbf{S}_2) \models \varphi(\mathbf{w}'_1, \mathbf{S}_2)$ for $i = 1, 2$. Obviously, $w'_1 \neq w_1$ or $w'_2 \neq w_2$. It follows that

$$opt_{\text{EMPTY}}(\mathcal{A}') = \max_{\mathbf{S}} |\{\, \mathbf{w} : (\mathcal{A}', \mathbf{S}) \models \varphi(\mathbf{w}, \mathbf{S}) \,\}| \geq |\{\, \mathbf{w} : (\mathcal{A}', \mathbf{S}_2) \models \varphi(\mathbf{w}, \mathbf{S}_2) \,\}| \geq 3.$$

However, $opt_{\text{EMPTY}}(\mathcal{A}') = 2$. This gives the desired contradiction. ∎

From Theorems 5.4 and 5.5 we obtain the following.

Corollary 5.6 *In the case of successor or ordered successor structures, the polynomially bounded maximization problems form the following hierarchy:*

$$\text{MAX}\,\Sigma_0 \subset \text{MAX}\,\Sigma_1 \subset \text{MAX}\,\Sigma_2 = \text{MAX}\,\Pi_1 = \text{MAX}\,\mathcal{PB}.$$

Acknowledgment The authors would like to thank Andreas Blass and Daniel Leivant for helpful discussions related to Lemma 4.2.

References

[1] J. Barwise. *Admissible Sets and Structures: An Approach to Definability Theory.* Springer Verlag, 1975.

[2] C. C. Chang and H. J. Keisler. *Model Theory.* North-Holland, 2nd edition, 1973.

[3] T. Eiter, G. Gottlob, and H. Mannila. Adding Disjunction to Datalog. In *Proceedings of the Thirteenth ACM SIGACT SIGMOD-SIGART Symposium on Principles of Database Systems (PODS-94)*, pages 267–278, May 1994.

[4] R. Fagin. Generalized First-Order Spectra and Polynomial-Time Recognizable Sets. In R. M. Karp, editor, *Complexity of Computation*, pages 43–74. AMS, 1974.

[5] S. C. Kleene. *Introduction to Metamathematics*. North-Holland, 1952.

[6] P. Kolaitis and M. N. Thakur. Logical Definability of NP Optimization Problems. *Information and Computation*, 115:321–353, 1994.

[7] P. Kolaitis and M. Y. Vardi. 0-1 Laws and Decision Problems for Fragments of Second-Order Logic. *Information and Computation*, 87:302–338, 1990.

[8] D. Leivant. Descriptive Characterizations of Computational Complexity. *Journal of Computer and System Sciences*, 39:51–83, 1989.

[9] Y. Matijasevich. *Hilbert's Tenth Problem*. MIT Press, 1993.

[10] A. Panconesi and D. Ranjan. Quantifiers and Approximation. *Theoretical Computer Science*, 107:145–163, 1993.

[11] C. H. Papadimitriou and M. Yannakakis. Optimization, Approximation, and Complexity Classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[12] I. Stewart. Comparing the Expressibility of Languages Formed Using NP-complete Operators. *Journal of Logic and Computation*, 1:305–330, 1991.

[13] J. van Benthem and K. Doets. Higher Order Logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Vol.I*, chapter I.4, pages 275–329. D. Reidel Publishing Company, 1983.