References

[Ber89] P. Berman. Private communication, 1989.

[BJY89] D. Bruschi, D. Joseph, and P. Young. A structural overview of NP optimization problems. Technical Report 861, Department of Computer Science, University of Wisconsin, Madison, 1989.

[Fag74] R. Fagin. Generalized first–order spectra and polynomial time recognizable sets. In R. Karp, editor, *Complexity of Computations*, pages 43–73. SIAM-AMS Proc. 7, 1974.

[GJ79] M. R. Garey and D. Johnson. *Computers and Intractability: A Guide to the theory of NP-completeness.* Freeman, San Francisco, 1979.

[Joh74] D. Johnson. Approximation algorithms for combinatorial problems. *J. Computer and System Sciences*, 9:256–278, 1974.

[KT90] Ph. G. Kolaitis and M. N. Thakur. Logical definability of NP optimization problems. Technical Report UCSC-CRL-90-48, Computer and Information Sciences, University of California, Santa Cruz, 1990.

[LM81] E. Leggett and J. Moore. Optimization problems and the polynomial time hierarchy. *Theoretical Computer Science*, 15:279 – 289, 1981.

[PR90] A. Panconesi and D. Ranjan. Quantifiers and approximation. In *Proceedings 22nd Annual ACM Symp. on Theory of Computing.* ACM, N. Y., 1990.

[PS82] C. Papadimitriou and Steiglitz. *Combinatorial Optimization - Algorithms and Complexity.* Prentice Hall, New Jersey, 1982.

[PY88] C. Papadimitriou and M. Yannakakis. Optimization, approximation and complexity. In *Proceedings 20th Annual ACM Symp. on Theory of Computing*, pages 229–234. ACM, N. Y., 1988.

[Tra50] B. A. Trakhtenbrot. The impossibility of an algorithm for the decision problem for finite models. *Doklady Akademii Nauk SSR*, 70:569 – 572, 1950.

[Yan81a] M. Yannakakis. Edge-deletion problems. *SIAM Journal for Computing*, 10(2):297 – 309, 1981.

[Yan81b] M. Yannakakis. Node-deletion problems on bipartite graphs. *SIAM Journal for Computing*, 10(2):310 – 326, 1981.

- MIN 3NON-TAUTOLOGY.

  Instance: A DNF formula with 3 literals per disjunct. Every instance I of MIN 3NT is identified with a finite structure $\mathbf{A}(I)$ with four ternary predicates D_0, D_1, D_2, D_3, where $D_i(w_1, w_2, w_3)$ is true if and only if the set $\{w_1, w_2, w_3\}$ is a disjunct with $w_1, \cdots, w_i$ appearing as negative literals and $w_{i+1}, \cdots, w_3$ appearing as positive literals, $0 \leq i \leq 3$.

  Solution: The minimum number of disjuncts simultaneously satisfiable under some truth assignment.

  $\mathrm{opt}_{\mathrm{MIN\ 3NT}}(I) = \min_S |\{(w_1, w_2, w_3) : \mathbf{A} \models \phi(w_1, w_2, w_3, S)\}|$,
  where $\phi(w_1, w_2, w_3, S)$ is the following quantifier-free formula:

  $$(D_0(w_1, w_2, w_3) \wedge S(w_1) \wedge S(w_2) \wedge S(w_3))\vee$$
  $$(D_1(w_1, w_2, w_3) \wedge \neg S(w_1) \wedge S(w_2) \wedge S(w_3))\vee$$
  $$(D_2(w_1, w_2, w_3) \wedge \neg S(w_1) \wedge \neg S(w_2) \wedge S(w_3))\vee$$
  $$(D_3(w_1, w_2, w_3) \wedge \neg S(w_1) \wedge \neg S(w_2) \wedge \neg S(w_3)).$$

- MAX SAT.

  Instance: A boolean formula in conjunctive normal form. An instance I of MAX SAT can be identified with a finite structure $\mathbf{A}(I) = (X, C, P, N)$, where X is the set of variables and clauses of I, the predicate $C(x)$ expresses that x is a clause, and $P(c, v)$ and $N(c, v)$ are binary predicates expressing that a variable v occurs positively or negatively in a clause c.

  Solution: The maximum number of clauses simultaneously satisfiable under some some truth assignment.

  $\mathrm{opt}_{\mathrm{MAX\ SAT}}(\mathbf{A}(I)) = \max_S |\{w : \mathbf{A}(I) \models (\exists y)(P(w, y) \wedge S(y)) \vee (N(w, y) \wedge \neg S(y))\}|$.

- MIN SET COVER.

  Instance: A set X and a collection, C, of subsets of X such that $X = \bigcup_{S \in C} S$. It is encoded by a finite structure $\mathbf{A} = (A, C, M)$, where $A = X \cup C$ is the universe of the structure, and $M(x, S)$ is a binary predicate expressing membership of an element x in the set S in C.

  Solution: The cardinality of the minimum cover C', such that $C' \subseteq C$, and $\bigcup_{S \in C'} S = X$.

  $\mathrm{opt}_{\mathrm{MIN\ SET\ COVER}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall x)(\exists y)(\neg C(x) \rightarrow (S(y) \wedge M(x, y)))\}$.

- MIN VERTEX COVER.

  Instance: A graph $G = (V, E)$.

  Solution: The cardinality of the minimum vertex cover in G.

  $\mathrm{opt}_{\mathrm{MIN\ VC}}(G) = \min_S\{|S| : G \models (\forall x)(\forall y)(E(x, y) \rightarrow (S(x) \vee S(y)))\}$.

- MIN EDGE DOMINATING SET.

  Instance: A graph $G = (V, E)$.

  Solution: The cardinality of the minimum edge dominating set in G. An Edge Dominating Set E' is a subset of edges such that every edge in E shares at least one endpoint with some edge in E'.

  $$\mathrm{opt}_{\text{MIN EDGE DS}}(G) = \min_S\{|S| : \mathbf{A} \models (\forall x)(\forall y)(\exists z)(E(x,y) \rightarrow [(S(x,z) \wedge E(x,z)) \vee (S(y,z) \wedge E(y,z))]\}.$$

- MIN DOMINATING SET.

  Instance: A graph $G = (V, E)$.

  Solution: The cardinality of the minimum dominating set in G. A Dominating Set is a set of vertices such that every vertex is either in the set or has a neighbor in the set.

  $$\mathrm{opt}_{\text{MIN DOM. SET}}(G) = \min_S\{|S| : \mathbf{A} \models (\forall x)(\exists y)(S(x) \vee S(y) \wedge E(x,y))\}.$$

- MIN GRAPH COLORING.

  Instance: A graph $G = (V, E)$.

  Solution: The minimum number of colors used to color the vertices of G such that adjacent vertices have a different color.

  $$\mathrm{opt}_{\text{MIN COLORING}}(G) = \min_{C,T}\{|C| : G \models (\forall x)(\exists c)\ [T(x,c) \wedge C(c)] \wedge (\forall v_1)(\forall c_1)(\forall v_2)(\forall c_2)[(E(v_1,v_2) \wedge T(v_1,c_1) \wedge T(v_2,c_2)) \rightarrow c_1 \neq c_2]\}.$$

- MIN HITTING SET

  Instance: A collection C of subsets of a finite set X. It is encoded by a finite structure $\mathbf{A} = (A, C, M)$, where $A = X \cup C$ is the universe of the structure, and $M(x, S)$ is a binary predicate expressing membership of an element x in the set S in C.

  Solution: The cardinality of the smallest subset $X' \subseteq X$, such that X' contains at least one element from subset in C.

  $$\mathrm{opt}_{\text{MIN HITTING SET}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall x)(\exists y)(C(x) \rightarrow (S(y) \wedge M(y,x)))\}.$$

- LONGEST-PATH-with-FORBIDDEN PAIRS.

  Instance: A directed graph $G = (V, E)$ and a collection $C = \{(a_1, b_1), \cdots, (a_n, b_n)\}$ of pairs of vertices from V.

  Solution: The longest simple path that contains at most one vertex from each pair in C.

which is a contradiction.
(2) Towards a contradiction, assume that MAX CLIQUE can be defined as a minimization problem using a universal first-order formula $(\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S})$ with a single free-variable w.

Let $H_1 = (V_1, E_1)$ be a graph with $V_1 = \{a_1, b\}$ and the edge relation being empty. Assume that $\mathbf{S}_1^*$ witnesses the optimum value for the graph H_1 and let $A_1^* = \{w \in V_1 : H_1 \models (\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S}_1^*)\}$. Therefore, we have that $|A_1^*| = 1$. Let $\tilde{A}_1^* = V_1 - A_1^*$ be the complement of A_1^*. We may assume, without loss of generality, that $a_1 \in \tilde{A}_1^*$.

Now let $H_2 = (V_2, E_2)$ be a graph isomorphic to H_1, with $V_2 = \{a_2, b\}$, and with the isomorphism mapping a_1 to a_2. Let $\mathbf{S}_2^*$ be the image of $\mathbf{S}_1^*$ under the same isomorphism. Analogous to A_1^*, we define $A_2^* = \{w \in V_2 : H_2 \models (\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S}_2^*)\}$ and conclude that $a_2 \in \tilde{A}_2^*$, where $\tilde{A}_2^* = V_2 - A_2^*$ is the complement of A_2^*.

Notice that $H_i \models \neg(\forall \mathbf{x})\psi(a_i, \mathbf{x}, \mathbf{S}_i^*)$, for $i = 1, 2$. Let $G = (V, E)$ be the graph with $V = \{a_1, a_2, b\}$ and $E = \{\{a_1, a_2\}\}$. We let $\mathbf{S}^* = \mathbf{S}_1^* \cup \mathbf{S}_2^*$ and consider the set $A^* = \{w \in V : G \models (\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S}^*)\}$. Since H_1, a_1 and H_2, a_2 do not satisfy the universal sentences above, and since universal sentences are preserved under substructures, we have that $G \models \neg(\forall \mathbf{x})\psi(a_i, \mathbf{x}, \mathbf{S}^*)$, for $i = 1, 2$. Therefore, a_1, a_2 are elements of $\tilde{A}^*$, where $\tilde{A}^* = V - A^*$ is the complement of A^*. Consequently, $|A^*| \leq 1$, which is a contradiction, as the maximum clique of G is of size 2. □

We should point out that in the above proof we used in a crucial way the assumption that the universal first-order formula had a single free variable. Indeed, we used the fact that $\tilde{A}^* = \tilde{A}_1^* \cup \tilde{A}_2^*$, which would not be true if arity of w was greater than 1.

It remains an open problem to extend the previous result to universal first-order formulae with more than one free variable. Preliminary investigations suggest that such a result poses challenging combinatorial difficulties, even for the case of a universal first order formula with exactly two free variables.

Appendix

We state here the definitions of some optimization problems used in the paper.

- MAX CLIQUE.

  Instance: A graph $G = (V, E)$.

  Solution: The cardinality of the largest clique of G.

  $\text{opt}_{\text{MAX CLIQUE}}(G) = \max_S\{|S| : G \models (\forall x)(\forall y)(S(x) \wedge S(y) \wedge x \neq y) \rightarrow E(x, y)\}$.

that NP $\neq$ co-NP if and only if CLIQUE (or any NP-complete decision problem) can not be defined on finite structures by any *universal second-order* sentence. As far as we know, Corollary 6.1 is the only other characterization of the NP $\stackrel{?}{=}$ co-NP question in terms of logical definability alone.

The above comments suggest that it may be possible to shed some light on the NP vs. co-NP problem by examining subclasses of MIN Π_1 and showing that MAX CLIQUE is not in some of them. What makes this approach plausible is the fact that universal first-order formulae have well-understood model theoretic properties, such as *preservation under substructures*, which may be used in obtaining lower-bound expressibility results. In what follows we begin an investigation along these lines by considering two proper subclasses of MIN Π_1, namely the class MIN Σ_1 (cf. Theorem 2) and the subclass of MIN Π_1 consisting of all NP minimization problems that are definable using a universal first-order formula with a single free variable.

Proposition 5: Let σ be a vocabulary consisting of a single binary predicate symbol E. Then the following are true:

1. MAX CLIQUE is not in the class MIN Σ_1 over the vocabulary σ.
2. MAX CLIQUE can not be defined as a minimization problem using a universal first-order formula with a single free variable, i.e., if $\mathbf{S}$ is a sequence of predicate symbols and $(\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S})$ is a universal first-order formula over $\sigma \cup \mathbf{S}$ having w as its only free variable, then there is a graph G such that

$$\mathrm{opt}_{\mathrm{MAX\ CLIQUE}}(G) \neq \min_{\mathbf{S}} |\{w : G \models (\forall \mathbf{x})\psi(w, \mathbf{x}, \mathbf{S})\}|.$$

Proof: (1) Towards a contradiction, assume that MAX CLIQUE is in the class MIN Σ_1. Let $(\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})$ be an existential first-order formula such that for every graph G we have

$$\mathrm{opt}_{\mathrm{MAX\ CLIQUE}}(G) = \min_{\mathbf{S}} |\{\mathbf{w} : G \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})\}|.$$

For simplicity, in what follows we write opt(G) instead of $\mathrm{opt}_{\mathrm{MAX\ CLIQUE}}(G)$. Let G be a graph consisting of two vertices $\{v_1, v_2\}$ and no edges and assume that $\mathbf{S}^*$ witnesses opt (G), i.e., $\mathrm{opt}(G) = |\{\mathbf{w} : G \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S}^*)\}| = 1$. Let H_1 be the subgraph of G whose only vertex is v_1, and let H_2 be the subgraph of G whose only vertex is v_2. Let $\mathbf{S}_1^*$ and $\mathbf{S}_2^*$ be the restrictions of $\mathbf{S}^*$ to the sets $\{v_1\}$ and $\{v_2\}$ respectively. If $\mathbf{b}$ is a tuple from H_i, $i = 1, 2$, such that $H_i \models (\exists \mathbf{x})\psi(\mathbf{b}, \mathbf{x}, \mathbf{S}_i^*)$, then it is also the case that $G \models (\exists \mathbf{x})\psi(\mathbf{b}, \mathbf{x}, \mathbf{S}^*)$, because existential formulae are preserved under extensions. But,

$$|\{\mathbf{w} : H_i \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S}_i^*)\}| \geq 1, \text{ for } i = 1, 2.$$

Moreover, the sets $\{\mathbf{w} : H_1 \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S}_1^*)\}$ and $\{\mathbf{w} : H_2 \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S}_2^*)\}$ are disjoint. Therefore,

$$|\{\mathbf{w} : G \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S}^*)\}| \geq 2,$$

(cf. definition 2.1). This problem asks: Given $I \in \mathcal{I}_\mathcal{Q}$ and an integer k, does there exist a feasible solution $T \in \mathcal{F}_\mathcal{Q}(I)$ such that $f_\mathcal{Q}(I,T) \geq k$? We say $T \in \mathcal{F}_\mathcal{Q}(I)$ is an *optimum solution* for instance I of $\mathcal{Q}$ if $f_\mathcal{Q}(I,T) = \max_\mathcal{Q}(I)$. Since NP = co-NP, the complement of the above decision problem is also in NP, and, consequently, the following problem is in NP: Given $I \in \mathcal{I}_\mathcal{Q}$ and an integer k, is there an optimum solution T of I for the problem $\mathcal{Q}$ such that $f_\mathcal{Q}(I,T) < k$?

Let $\mathcal{Q}^* = (\mathcal{I}_{\mathcal{Q}^*}, \mathcal{F}_{\mathcal{Q}^*}, f_{\mathcal{Q}^*}, \min)$ be the minimization problem with

$$\begin{aligned}\mathcal{I}_{\mathcal{Q}^*} &= \mathcal{I}_\mathcal{Q},\\ \mathcal{F}_{\mathcal{Q}^*}(I) &= \{T : T \text{ is an optimum solution of } I \text{ for } \mathcal{Q}\}, \text{ for all } I \in \mathcal{I}_{\mathcal{Q}^*},\\ f_{\mathcal{Q}^*}(I,T) &= f_\mathcal{Q}(I,T) \text{ for all } T \in \mathcal{F}_{\mathcal{Q}^*}(I).\end{aligned}$$

The associated decision problem for $\mathcal{Q}^*$ is: Given $I \in \mathcal{I}_{\mathcal{Q}^*}$ and an integer k, is there a feasible solution $T \in \mathcal{F}_{\mathcal{Q}^*}(I)$ such that $f_{\mathcal{Q}^*}(I,T) \leq k$? Note that this is in NP, because the above decision problem is in NP. Hence, $\mathcal{Q}^*$ is an NP minimization problem.

Notice that for all feasible solutions T of I for $\mathcal{Q}$ it is the case that $f_{\mathcal{Q}^*}(I,T) = \mathrm{opt}_\mathcal{Q}(I)$. As a result, $\min_{T \in \mathcal{F}_{\mathcal{Q}^*}(I)} f_{\mathcal{Q}^*}(I,T) = \mathrm{opt}_\mathcal{Q}(I)$. Therefore, $\min_{\mathcal{Q}^*}(I) = \max_\mathcal{Q}(I)$ for all instances I. □

As mentioned earlier, Kolaitis and Thakur [KT90] showed that the class MIN $\mathcal{PB}$ of all polynomially bounded minimization problems coincides with the class MIN Π_1 of minimization problems definable using universal first-order formulae (cf. Theorem 2). By combining this result with the preceding Proposition 4, we obtain the following reformulation of the NP $\stackrel{?}{=}$ co-NP question.

Corollary 6.1: The following two statements are equivalent:

1. NP $\neq$ co-NP.
2. MAX CLIQUE is *not* in the class MIN Π_1, i.e., there is no universal first-order formula $(\forall \mathbf{x})\psi(\mathbf{w},\mathbf{x},\mathbf{S})$ (where $\psi(\mathbf{w},\mathbf{x},\mathbf{S})$ is a quantifier-free formula) over a vocabulary $\{E\} \cup \mathbf{S}$ such that for every graph $G = (V,E)$ we have that

$$\max\nolimits_{\text{MAX CLIQUE}}(G) = \min_{\mathbf{S}} |\{\mathbf{w} : G \models (\forall \mathbf{x})\psi(\mathbf{w},\mathbf{x},\mathbf{S})\}|.$$

The preceding Corollary 6.1 holds also with any NP-hard maximization problem in place of MAX CLIQUE. We chose to use MAX CLIQUE here, because it is in the class MAX Π_1 and, thus, the result makes the difference between the classes MAX Π_1 and and MIN Π_1 more striking.

Corollary 6.1 yields a machine-independent characterization of the NP $\stackrel{?}{=}$ co-NP question. Fagin [Fag74] characterized NP computability in terms of definability in second-order logic on finite structures. From Fagin's [Fag74] main result, it follows this conjecture can be established directly using combinatorial and model theoretic arguments, and without any complexity theoretic assumptions.

5 Maximization Problems vs. Minimization Problems

In some cases, given a maximization problem $\mathcal{Q}$, one can find a minimization problem $\mathcal{Q}^*$ with the property that the optimum value of $\mathcal{Q}$ is equal to the optimum value of $\mathcal{Q}^*$. LINEAR PROGRAMMING provides the canonical manifestation of this phenomenon. Indeed, *duality theory* makes it possible to rewrite a given maximization linear programming problem as a minimization linear programming problem, and vice versa (cf. [PS82]).

When it comes to NP optimization problems, a folklore result in complexity theory asserts that, unless P = NP, it not possible to rewrite every NP maximization problem as an NP minimization problem. We now state this result more formally and prove it, since we were not able to pinpoint an exact reference in the literature for it.

Proposition 4: The following statements are equivalent.

1. For every NP maximization problem $\mathcal{Q}$, there is an NP minimization problem $\mathcal{Q}^*$ such that $\mathcal{Q}$ and $\mathcal{Q}^*$ have the same instances and for every instance I, $\max_{\mathcal{Q}}(I) = \min_{\mathcal{Q}^*}(I)$.
2. NP = co-NP.
3. For every NP minimization problem $\mathcal{Q}^*$, there is an NP maximization problem $\mathcal{Q}$ such that $\mathcal{Q}^*$ and $\mathcal{Q}$ have the same instances and for every instance I, $\min_{\mathcal{Q}^*}(I) = \max_{\mathcal{Q}}(I)$.

Proof: We prove here that statements (1) and (2) above are equivalent. The remaining equivalences can be proved with a similar argument.

Assume first that every NP maximization problem can be rewritten as a minimization problem with the same instances. Since MAX CLIQUE is an NP maximization problem with graphs as instances, there is an NP minimization problem $\mathcal{Q}^* = (\mathcal{I}_{\mathcal{Q}^*}, \mathcal{F}_{\mathcal{Q}^*}, f_{\mathcal{Q}^*}, \min)$ on graphs such that $\max_{\text{MAX CLIQUE}}(G) = \min_{\mathcal{Q}^*}(G)$.

Consider now the NP-complete decision problem CLIQUE: Given a graph G and an integer k, does G have a clique of size greater than or equal to k? It follows from the above that CLIQUE has a YES answer on a graph G if and only if $\min_{\mathcal{Q}^*}(G) = \min_{T \in \mathcal{F}_{\mathcal{Q}^*}} f_{\mathcal{Q}^*}(G,T) \geq k$. Thus, CLIQUE has a YES answer on a graph G if and only if for every feasible solution T of $\mathcal{Q}^*$ on G we have that $f_{\mathcal{Q}^*}(G,T) \geq k$. Since $\mathcal{Q}^*$ is an NP minimization problem, the latter decision problem is in co-NP (cf. definition 2.1). As a result, CLIQUE is in co-NP and, consequently, NP = co-NP.

Assume now that NP = co-NP and let $\mathcal{Q} = (\mathcal{I}_{\mathcal{Q}}, \mathcal{F}_{\mathcal{Q}}, f_{\mathcal{Q}}, \max)$ be an NP maximization problem. Therefore, the decision problem, associated with $\mathcal{Q}$ is in NP

Let $\mathbf{S} = (S_1, S_1^*, \cdots, S_{t'}^*)$ and let $\phi(\mathbf{S})$ be the formula $(\forall \mathbf{w})(S_1^*(\mathbf{w}) \rightarrow S_1(\mathbf{w})) \wedge \psi(\mathbf{S}^*)$. It follows that

$$\text{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\}.$$

Note that S_1 has only positive occurrences in $\phi(\mathbf{S})$ and note also that the quantifier-free part of $\phi(\mathbf{S})$, when expressed in DNF, has at most one occurrence of S_1 per disjunct. □

The preceding Proposition 2 illuminates the differences between the class MIN $\text{F}^+\Pi_2(1)$ and the class MIN $\mathcal{PB}$ of all polynomially bounded minimization problems. Indeed, it follows that it is the presence of additional predicates in the sequence $\mathbf{S}$ that makes the difference between MIN $\mathcal{PB}$ and the MIN $\text{F}^+\Pi_2(1)$. In other words, the optimization problems in the class MIN $\text{F}^+\Pi_2(1)$ have the property that the feasible solution is represented by only one predicate, the cardinality of which is the objective function, while in the larger class MIN $\mathcal{PB}$ the feasible solution is a sequence of predicates and the objective function is the cardinality of one of these predicates. For both classes the quantifier complexity of the formulae and the syntactic restrictions on the occurrences of S_1 in the formulae are the same.

As mentioned earlier, there are polynomially bounded minimization problems, such as MIN 3NON-TAUTOLOGY, that are not log-approximable, unless P = NP. Since these problems are not in the class MIN $\text{F}^+\Pi_2(1)$, it follows that the additional predicates are indispensable in order to express all polynomially bounded NP minimization problems in the logical definability framework.

So far, we focused on the approximation properties of optimization problems in the class MIN $\text{F}^+\Pi_2(1)$. In the beginning of this section, we introduced also the classes MIN $\text{F}^+\Pi_2(k)$, $k > 1$, containing minimization problems that are defined using formulae in which S_1 is allowed to have up to k occurrences (all of them positive) in every disjunct of the quantifier-free part (in DNF). It turns out that this relaxed condition gives rise to weaker approximation properties.

Proposition 3: Let $\mathcal{Q}$ be an optimization problem in the class MIN $\text{F}^+\Pi_2(k)$, for some $k > 1$. Then there is a polynomial-time approximation algorithm and a constant c such that for every instance $\mathbf{A}$ of $\mathcal{Q}$ the algorithm produces a feasible solution on which the objective function takes value less than or equal to $c(\text{opt}(\mathbf{A})^k)\log(|A|)$.

Proof: (*Hint:*) The approximation algorithm is an extension of Johnson's greedy algorithm [Joh74] for the MIN SET COVER problem. Details will appear in the full paper. □

We conclude this section by conjecturing that the MIN GRAPH COLORING problem is not in the class MIN $\text{F}^+\Pi_2 = \bigcup_k \text{MIN F}^+\Pi_2(k)$. We believe that

It can be seen that C^* is a set cover for $I_{\mathbf{A}}$ if and only if $\{\mathbf{a} : s_{\mathbf{a}} \in C^*\}$ covers every element of A^k. Therefore, $\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \mathrm{opt}_{\mathrm{SET\ COVER}}(I_{\mathbf{A}})$. It is clear that this is an A-reduction. Hence, MIN SET COVER is complete for MIN $\mathrm{F}^+\Pi_2(1)$ under A-reductions.

We can show that MIN SET COVER $\leq_A$ MIN DOMINATING SET. As a result MIN DOMINATING SET is also complete for MIN $\mathrm{F}^+\Pi_2(1)$ under $A-$reductions. □

We now comment on the differences between the class MIN $\mathrm{F}^+\Pi_2(1)$ and the class MIN $\mathcal{PB}$ of all polynomially bounded NP minimization problems.

Recall the earlier Theorem 1, which asserts that MIN $\mathcal{PB}$ coincides with the class MIN $\mathrm{F}\Pi_2$. In other words, if $\mathcal{Q}$ is a polynomially bounded minimization problem with instances finite structures $\mathbf{A}$ over a vocabulary σ, then there is a Π_2 sentence $\psi(\mathbf{S})$ over the vocabulary $\sigma \cup S$, where $\mathbf{S} = (S_1, \cdots, S_t)$ is a sequence of predicate symbols not in σ, such that

$$\mathrm{opt}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \psi(\mathbf{S})\},$$

for every finite structure $\mathbf{A}$ over σ. The following simple result shows that a syntactically proper subclass of MIN $\mathrm{F}\Pi_2$ captures every polynomially bounded minimization problem.

Proposition 2: Let $\mathcal{Q}$ be a polynomially bounded NP minimization problem with instances finite structures $\mathbf{A}$ over a vocabulary σ. Then there is a Π_2 sentence $\phi(\mathbf{S})$ over the vocabulary $\sigma \cup S$, where $\mathbf{S}$ is the sequence $(S_1, \cdots, S_t)$ of predicate symbols not in σ, such that

1. $\mathrm{opt}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\}$, for every finite structure $\mathbf{A}$ over σ;
2. the predicate symbol S_1 has only *positive* occurrences in $\phi(S)$;
3. the quantifier-free part of the sentence $\psi(\mathbf{S})$ is equivalent to a formula in DNF with at most one occurrence of S_1 per disjunct.

Proof: Since $\mathcal{Q}$ is a polynomially bounded NP minimization theorem, Theorem 1 implies that optima on instances $\mathbf{A}$ over σ can be expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{\mathbf{S}^*}\{|S_1^*| : \mathbf{A} \models \psi(\mathbf{S}^*)\},$$

where $\psi(\mathbf{S}^*)$ is a Π_2 sentence with predicate symbols amongst those in σ and the sequence of predicates $\mathbf{S}^* = (S_1^*, \cdots, S_{t'}^*)$. Notice that at this point the predicate S_1^* may have negative occurrences in the formula $\psi(\mathbf{S}^*)$. We now introduce a new predicate S_1 and insist that $S_1^* \subseteq S_1$. In other words, we express the optimum value of $\mathcal{Q}$ as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{S_1, \mathbf{S}^*}\{|S_1| : \mathbf{A} \models (\forall \mathbf{w})(S_1^*(\mathbf{w}) \rightarrow S_1(\mathbf{w})) \wedge \psi(\mathbf{S}^*)\}.$$

It is clear that MIN $F^+\Pi_1$ is a subclass of MIN $F\Pi_2$. Moreover, MIN $F^+\Pi_1$ is contained in the class MIN $F^+\Pi_2$ and, in fact, this containment is a proper one, because MIN SET COVER witnesses the separation of the two classes.

We will be focusing on the class MIN $F^+\Pi_2(1)$, which contains MIN SET COVER and other natural optimization problems, such as MIN DOMINATING SET, MIN EDGE DOMINATING SET, and MIN HITTING SET [GJ79]. Actually, it will turn out that some of these problems are complete for the class MIN $F^+\Pi_2(1)$, via a certain *approximability preserving reduction.*

The notion of an approximabilty preserving reduction was introduced by Papadimitriou and Yannakakis [PY88], who considered *L-reductions* between optimization problems. Panconesi and Ranjan [PR90] generalized this to the notion of A-*reduction* ($\leq_A$). These reductions have the property that if $\mathcal{P}$ and $\mathcal{Q}$ are optimization problems such that $\mathcal{P} \leq_A \mathcal{Q}$ and $\mathcal{Q}$ is g-approximable, for some function g, then $\mathcal{P}$ is also g-approximable. We now establish the following result.

Theorem 6: The MIN SET COVER problem and the MIN DOMINATING SET problem are complete for the class MIN $F^+\Pi_2(1)$ under $A-$reductions. As a result, every problem in the class MIN $F^+\Pi_2(1)$ is log-approximable.

Proof: (Sketch) Let $\mathcal{Q}$ be a problem in the class MIN $F^+\Pi_2(1)$. Hence, its optimum is expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall \mathbf{x})(\exists \mathbf{y})\psi(\mathbf{x}, \mathbf{y}, S)\},$$

where S is an m-ary predicate symbol, $\psi(\mathbf{x}, \mathbf{y}, S)$ is a quantifier-free DNF formula in which all occurrences of S are positive, and S occurs at most once in each disjunct. Assume that the arity of $\mathbf{x}$ is k, the arity of $\mathbf{y}$ is l, and let N denote $|A|^l$.

Given an instance $\mathbf{A}$ of $\mathcal{Q}$, we express its optimum as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall \mathbf{x})(\bigvee_{j=1}^{N} \psi(\mathbf{x}, \mathbf{c}_j, S))\},$$

where $A^l = \{\mathbf{c}_1, \cdots, \mathbf{c}_N\}$. We say a set $S \subseteq A^m$ *covers* a tuple $\mathbf{b}$, if $\mathbf{A} \models \bigvee_{j=1}^{N} \psi(\mathbf{b}, \mathbf{c}_j, S)$. Observe that $\mathrm{opt}_{\mathcal{Q}}(\mathbf{A})$ is the cardinality of the smallest set $S \subseteq A^m$ such that S covers every element of A^k.

We now use the structure $\mathbf{A}$ to construct the following instance $I_{\mathbf{A}} = (X, C)$ of the MIN SET COVER problem:

$$X = A^k, \quad C = \{s_{\mathbf{a}} : \mathbf{a} \in A^m\},$$

where $\mathbf{b}$ is an element of the set $s_{\mathbf{a}}$ if and only if $\mathbf{A} \models \bigvee_{j=1}^{N} \psi(\mathbf{b}, \mathbf{c}_j, S/\{\mathbf{a}\})$. In other words a set $s_{\mathbf{a}}$ contains $\mathbf{b}$ if and only if $\{\mathbf{a}\}$ covers $\mathbf{b}$.

Proposition 1: Let σ be a vocabulary consisting of a unary predicate symbol C and a binary predicate symbol M. The MIN SET COVER problem is not in the class MIN FΠ_1 over σ, and a fortiori, not in the class MIN $\mathrm{F}^+\Pi_1$.

Proof: Towards a contradiction, assume that MIN SET COVER is defined as

$$\mathrm{opt}_{\mathrm{SET\ COVER}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models \phi(S)\},$$

where $\phi(S)$ is a universal first-order sentence over $\sigma \cup \{S\}$. Let $I = (X, C)$ be the following instance of MIN SET COVER:

$$X = \{x_1, x_2\}, \quad C = \{c_1, c_2, c_3\},$$

$$\text{where} \quad c_1 = \{x_1\}, \quad c_2 = \{x_2\}, \quad c_3 = \{x_1, x_2\}.$$

Assume that S^* witnesses the optimum of MIN SET COVER on the encoding $\mathbf{A}(I)$ of the instance I, i.e., $\mathbf{A}(I) \models \phi(S^*)$ and $|S^*| = \mathrm{opt}(\mathbf{A}(I)) = 1$. Let $\mathbf{A}_0$ be the substructure of $\mathbf{A}(I)$ with universe $\{x_1, x_2, c_1, c_2\}$, and let S_0^* be the restriction of S^* to the universe of $\mathbf{A}_0$. It is clear that $\mathrm{opt}(\mathbf{A}_0) = 2$ and, consequently, $|S_0^*| \geq 2$. Since $\mathbf{A}(I) \models \phi(S^*)$ and $\phi(S^*)$ is a universal sentence, it is also the case that $\mathbf{A}_0 \models \phi(S_0^*)$. As a result, $|S^*| \geq 2$, which is a contradiction. □

4.2 The class MIN $\mathrm{F}^+\Pi_2$

We now introduce the class MIN $\mathrm{F}^+\Pi_2$, which is a subclass of MIN FΠ_2 containing the MIN SET COVER problem. Observe that the optimum of the MIN SET COVER problem on a structure $\mathbf{A}$ over the vocabulary $\{C, M\}$ is defined as follows:

$$\mathrm{opt}_{\mathrm{SET\ COVER}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall x)(\exists y)(\neg C(x) \rightarrow (S(y) \wedge M(x, y)))\}.$$

This observation motivates the following definition.

Definition 4.1: Let MIN $\mathrm{F}^+\Pi_2(k)$, $k \geq 1$, be the class of minimization problems $\mathcal{Q}$ whose optimum on a structure $\mathbf{A}$ over a vocabulary σ can be expressed as:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall \mathbf{x})(\exists \mathbf{y})\psi(\mathbf{x}, \mathbf{y}, S)\},$$

where S is a single predicate of some fixed arity, $\psi(\mathbf{x}, \mathbf{y}, S)$ is a quantifier-free *DNF* formula in which all occurrences of S are positive, and S occurs at most k times in each disjunct. We also let

$$\text{MIN } \mathrm{F}^+\Pi_2 = \bigcup_k \text{MIN } \mathrm{F}^+\Pi_2(k)$$

denote the union of these classes.

As mentioned earlier, MIN VERTEX COVER is a constant-approximable problem (cf. [PS82]). The following result by Kolaitis and Thakur [KT90] shows that this property is shared by every member of the class MIN $F^+\Pi_1$.

Theorem 5: Every problem in the class MIN $F^+\Pi_1$ is constant-approximable.

We now discuss the expressive power of MIN $F^+\Pi_1$. On the positive side, in addition to MIN VERTEX COVER, the class MIN $F^+\Pi_1$ contains a large number of *node-deletion* and *edge-deletion* graph problems (cf. [Yan81a,Yan81b]).

If Ψ is a property of graphs, then the *node (edge) deletion problem* NODE-DEL$_\Psi$ (EDGE-DEL$_\Psi$) *associated with* Ψ is defined as follows: Given a graph G, find a set of nodes (edges) of minimum cardinality whose deletion from G results into a graph satisfying Ψ [Yan81a,Yan81b].

Several well known NP-hard optimization problems, such as MIN VERTEX COVER and MIN FEEDBACK ARC SET [GJ79], can be stated as node or edge deletion problems by specifying the property Ψ appropriately. Assume now that Ψ is a property of finite graphs that is definable using a universal first-order sentence. Then the node (edge) deletion problem NODE-DEL$_\Psi$ (EDGE-DEL$_\Psi$) associated with Ψ is contained in the class MIN $F^+\Pi_1$. Indeed, it is easy to verify that if Ψ is definable by the universal sentence $(\forall x_1)\cdots(\forall x_t)\psi(x_1,\cdots,x_t)$, then the optimum of NODE-DEL$_\Psi$ on a graph G can be expressed as

$$\mathrm{opt}(G) = \min_S\{|S| : G \models (\forall x_1)\cdots(\forall x_t)(\neg\psi(x_1,\cdots,x_t) \rightarrow (S(x_1) \vee \cdots \vee S(x_t)))\}.$$

Edge Deletion problems can be formulated similarly. Yannakakis [Yan81a,Yan81b] showed that if Ψ is one of the following properties, then the node or edge deletion decision problem associated with Ψ is NP-complete:

1. transitive digraph (edge deletion problem).
2. without cycles of specified length l, for any fixed $l \geq 3$ (edge deletion problem).
3. maximum degree 1 (node deletion problem).

Each of the above properties is definable by a universal first-order sentence and, thus, the associated minimization problem is in the class MIN $F^+\Pi_1$.

On the negative side, we will show below that the MIN SET COVER problem is not in the class MIN $F^+\Pi_1$ (no complexity theoretic assumptions will be used in the proof of this result). MIN SET COVER is a log-approximable problem [Joh74] and no better approximation properties for it are known.

Recall that an instance $I = (X, C)$ of the MIN SET COVER problem is viewed as a finite structure $\mathbf{A}(I) = (X \cup C, C, M)$, where M is a binary relation expressing membership of an element $x \in X$ in a set $S \in C$.

where $\mathbf{A}$ is the structure $(A^* \cup V_0, V_0, E_0, E^*)$ over σ. Since $\mathcal{Q}_\chi$ is not constant-approximable, unless P = NP, it follows that if P $\neq$ NP, then $\mathcal{Q}_\phi$ is not constant-approximable. □

This proof can be extended to the undecidability of g-approximability for any function g for which there is an optimization problem $\mathcal{Q}$ that is not g-approximable, modulo P $\neq$ NP or some other complexity theory assumption. In particular, we can show that it is undecidable if a first-order formula defines a log-approximable problem.

4 Approximation Properties of Minimization Problems

The preceding Theorem 4 implies that, unless P = NP, it is not possible to find a "nice" necessary and sufficient condition that characterizes which first-order formulae give rise to approximable optimization problems. In view of this, we can only hope to isolate sufficient conditions for approximability. In this section we investigate certain syntactically defined classes of minimization problems and study their approximation properties.

Papadimitriou and Yannakakis [PY88] showed that every problem in the class MAX Σ_1 is constant-approximable. On the other hand, Kolaitis and Thakur [KT90] showed that the class MIN Σ_1 contains problems, such as MIN 3NON-TAUTOLOGY, that are not log-approximable, unless P = NP. This suggests that if one wants to isolate sufficient conditions for approximability of minimization problems in the logical definability framework, then one has to look for syntactic features beyond the quantifier patterns. It turns out that certain natural subclasses of MIN FΠ_1 and MIN FΠ_2 possess good approximation properties. The first of these is the class MIN F$^+\Pi_1$ below, which was introduced and studied in [KT90].

4.1 The class MIN F$^+\Pi_1$

We define MIN F$^+\Pi_1$ to be the collection of all minimization problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ can be expressed as:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_S\{|S| : \mathbf{A} \models (\forall \mathbf{x})\psi(\mathbf{x}, S)\},$$

where S is a single predicate of some fixed arity and $\psi(\mathbf{x}, S)$ is a quantifier-free formula over the vocabulary $\sigma \cup \{S\}$ in which all occurrences of S are *positive*.

It is clear that MIN F$^+\Pi_1$ is a subclass of MIN FΠ_1. Moreover, MIN F$^+\Pi_1$ contains MIN VERTEX COVER as a member, since the optimum of MIN VERTEX COVER on a graph $G = (V, E)$ is given by

$$\mathrm{opt}_{\mathrm{VC}}(G) = \min_S\{|S| : G \models (\forall x)(\forall y)(E(x, y) \rightarrow S(x) \vee S(y))\}.$$

Proof: We use Trakhtenbrot's classical theorem [Tra50], which asserts that the set of first-order sentences true on all finite structures over a vocabulary τ is not recursive, provided τ contains at least one non-unary predicate symbol. We reduce the question of "truth on all finite structures" to that of "constant-approximabilty".

Let $\mathbf{S} = (S_1, \cdots, S_t)$ be a sequence of predicate symbols and let $\chi(\mathbf{S})$ be a first-order formula over the vocabulary $\{E^*\} \cup \mathbf{S}$ such that $\mathcal{Q}_\chi$ is a minimization problem that is not constant-approximable, unless P = NP. The instances of Q_χ are identified with finite structures $\mathbf{A}^* = (A^*, E^*)$.

Given a first-order formula $\psi(E)$ over the vocabulary $\{E\}$, we consider the following formula $\phi(\mathbf{S})$ with predicate symbols from $\{V, E, E^*\} \cup \mathbf{S}$:

$$\phi(\mathbf{S}) \stackrel{def}{\equiv} \neg\psi^V(E) \wedge \chi^{\tilde{V}}(\mathbf{S}) \wedge (\bigwedge_{i=1}^{t} S_i \subseteq \tilde{V}^{\alpha_i}),$$

where $\tilde{V}$ is the complement $\neg V$ of V, the expression $\psi^V(E)$ denotes the formula $\psi(E)$ *relativized* to V, the expression $\chi^{\tilde{V}}$ denotes the formula χ relativized to $\tilde{V}$, and α_i is the arity of S_i, $1 \leq i \leq t$. The concept of *relativization* used here is from mathematical logic, namely, if φ is a formula and R is a unary predicate, then the *relativized* formula φ^R is obtained from φ by replacing every subformula $(\forall x)\varphi'(x)$ of φ by $(\forall x)(R(x) \rightarrow \varphi'(x))$ and by replacing every subformula $(\exists x)\varphi'(x)$ of φ by $(\exists x)(R(x) \wedge \varphi'(x))$.

We now consider the minimization problem $\mathcal{Q}_\phi$ defined by $\phi(\mathbf{S})$. The instances of $\mathcal{Q}_\phi$ are finite structures of the form $\mathbf{A} = (A, V, E, E^*)$ over the vocabulary σ and the optimum of $\mathcal{Q}_\phi$ is given as

$$\mathrm{opt}_{\mathcal{Q}_\phi}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\}.$$

We will show that the truth of $\psi(E)$ on all finite structures is equivalent to the constant-approximabilty of $\mathcal{Q}_\phi$, modulo P $\neq$ NP.

If the sentence $\psi(E)$ is true on all finite structures, then $\mathbf{A} \not\models \phi(\mathbf{S})$, for every finite structure $\mathbf{A}$ over σ and for every sequence of relations $\mathbf{S}$ on A. Hence, for all finite structures $\mathbf{A}$, we have that

$$\mathrm{opt}_{\mathcal{Q}_\phi}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\} = \mathit{triv}_{\mathcal{Q}_\phi}.$$

As a result, in this case $\mathcal{Q}_\phi$ is trivially constant-approximable.

For the other direction, assume that (V_0, E_0) is a finite structure over the vocabulary $\{E\}$ on which the sentence $\psi(E)$ is false. Then, given any finite structure $\mathbf{A}^* = (A^*, E^*)$ over the vocabulary $\{E^*\}$, we have that

$$\mathrm{opt}_{\mathcal{Q}_\chi}(\mathbf{A}^*) = \mathrm{opt}_{\mathcal{Q}_\phi}(\mathbf{A}),$$

and

$$\text{MIN}\,\Pi_n \quad \subseteq \quad \text{MIN F}\Pi_{n+1}, \qquad\qquad \text{MIN}\,\Sigma_n \quad \subseteq \quad \text{MIN F}\Pi_n, \quad n \geq 1.$$

It follows that

$$\text{MAX}\,\Pi_n = \text{MAX F}\Pi_n \quad \text{and} \quad \text{MIN}\,\Sigma_n = \text{MIN F}\Pi_n,$$

for $n \geq 1$.

The preceeding Theorem 1 can now be restated as follows.

Theorem 3: The class MAX $\mathcal{PB}$ of all polynomially bounded NP maximization problems coincides with the class MAX $\text{F}\Pi_2$. Thus,

$$\text{MAX } \mathcal{PB} = \text{MAX F}\Pi_2 = \text{MAX F}\Pi_n, \quad n \geq 2.$$

The class MIN $\mathcal{PB}$ of all polynomially bounded NP minimization problems coincides with the class MIN $\text{F}\Pi_2$. Thus,

$$\text{MIN } \mathcal{PB} = \text{MIN F}\Pi_2 = \text{MIN F}\Pi_n, \quad n \geq 2.$$

3 On the Undecidability of Approximation Properties

We showed before that we can express all polynomially bounded NP optimization problems using logic. It is natural to ask whether or not logic can be also used to capture all g-approximable problems, for a given function g. In this section we provide a negative answer to this question by establishing that, assuming P≠NP, it is an *undecidable* problem to tell if a first-order formula gives rise to a constant-approximable problem.

Let σ be a vocabulary and let $\phi(\mathbf{S})$ be a first-order sentence with predicate symbols from $\sigma \cup \mathbf{S}$. We write $\mathcal{Q}_\phi$ to denote the minimization problem whose optimum on a structure $\mathbf{A}$ is expressed as

$$\text{opt}_{\mathcal{Q}_\phi}(\mathbf{A}) = \min_{\mathbf{S}}\{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\}.$$

We now state and prove the main theorem of this section.

Theorem 4: Let σ be a vocabulary with one unary predicate symbol $\{V\}$ and two binary predicate symbols $\{E, E^*\}$. Assuming P≠NP, the following is an undecidable problem: Given a first-order sentence $\phi(\mathbf{S})$ over $\sigma \cup \mathbf{S}$, is the minimization problem $\mathcal{Q}_\phi$ constant-approximable?

where $opt \in \{\max, \min\}$, and $\phi(\mathbf{S})$ is a first-order sentence. Let $\mathbf{w}$ range over tuples with arity the same as the arity of S_1. If $\mathcal{Q}$ is a maximization problem, then its optimum is expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models S_1(\mathbf{w}) \wedge \phi(\mathbf{S})\}|.$$

On the other hand, if $\mathcal{Q}$ is a minimization problem, then its optimum is expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models \phi(\mathbf{S}) \rightarrow S_1(\mathbf{w})\}|.$$

Notice that, unlike the case of maximization problems, we cannot express the optima of minimization problem as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models S_1(\mathbf{w}) \wedge \phi(\mathbf{S})\}|,$$

because the minimum cardinality of the above set is zero, which occurs when S_1 is empty.

From the preceding remarks, it follows that, for $n \geq 1$,

$$\text{MAX F}\Pi_n \subseteq \text{MAX}\,\Pi_n, \qquad \text{MAX F}\Sigma_n \subseteq \text{MAX}\,\Sigma_n, \quad n \geq 1,$$

and

$$\text{MIN F}\Pi_n \subseteq \text{MIN}\,\Sigma_n, \qquad \text{MIN F}\Sigma_n \subseteq \text{MIN}\,\Pi_n, \quad n \geq 1.$$

In the opposite direction, assume that $\mathcal{Q}$ is an optimization problem with optimum on a structure $\mathbf{A}$ expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \mathop{opt}_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models \phi(\mathbf{w}, \mathbf{S})\}|,$$

where $opt \in \{\max, \min\}$, and $\phi(\mathbf{w}, \mathbf{S})$ is a first-order formula. By introducing a new predicate symbol T with arity the same as that of $\mathbf{w}$, we can express the optimum as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \mathop{opt}_{T,\mathbf{S}} \{|T| : \mathbf{A} \models (\forall \mathbf{w})(T(\mathbf{w}) \leftrightarrow \phi(\mathbf{w}, \mathbf{S}))\}.$$

It follows that

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \max_{T,\mathbf{S}} \{|T| : \mathbf{A} \models (\forall \mathbf{w})(T(\mathbf{w}) \rightarrow \phi(\mathbf{w}, \mathbf{S}))\}$$

for a maximization problem $\mathcal{Q}$, while

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \min_{T,\mathbf{S}} \{|T| : \mathbf{A} \models (\forall \mathbf{w})(\phi(\mathbf{w}, \mathbf{S}) \rightarrow T(\mathbf{w}))\}$$

for a minimization problem $\mathcal{Q}$. As a result, for $n \geq 1$, we have the containments

$$\text{MAX}\,\Pi_n \subseteq \text{MAX F}\Pi_n, \qquad \text{MAX}\,\Sigma_n \subseteq \text{MAX F}\Pi_{n+1}, \quad n \geq 1,$$

2.2 Logic and Feasible Solutions

In this section we introduce a different approach to defining optimization problems using logic. For many natural optimization problems, a feasible solution is a collection of relations and the objective function is the cardinality of one of these relations. For example, a feasible solution of the MIN VERTEX COVER problem is a set of vertices forming a vertex cover and the objective function is its cardinality. A feasible solution of the MAX CLIQUE problem is a set of vertices forming a clique and the objective function is the cardinality of the clique. In both the above examples, a feasible solution consists of a single relation. On the other hand, a feasible solution in the MIN GRAPH COLORING problem is a pair of two relations C, T, where C is a set that contains colors and T is a binary relation that denotes a legal assignment of colors to the vertices of the graph. In this case, the objective function is the cardinality of C. In what follows, we use this observation to introduce classes of optimization problems.

Definition 2.5: For each $n \geq 1$, let MAX FΠ_n (F stands for *feasible*) be the class of maximization problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ are defined as follows:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \begin{cases} \max_{\mathbf{S}} \{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\} & \text{if there is an } \mathbf{S} \text{ such that } \mathbf{A} \models \phi(\mathbf{S}), \\ triv_{\mathcal{Q}} & \text{otherwise,} \end{cases}$$

where $\mathbf{S} = (S_1, \cdots, S_t)$ is a sequence of predicate symbols, $\phi(\mathbf{S})$ is a Π_n *sentence* (i.e., a formula with no free variables) with predicate symbols from $\sigma \cup \mathbf{S}$, and $triv_{\mathcal{Q}}$ is a trivial value for the optimum. Similarly, for each $n \geq 1$, we define the class MAX FΣ_n using Σ_n sentences. We also define the classes MIN FΠ_n and MIN FΣ_n, $n \geq 1$ analogously. For the trivial value $triv_{\mathcal{Q}}$ we have that

$$triv_{\mathcal{Q}} = \begin{cases} |A|^m & \text{if } \mathcal{Q} \text{ is a minimization problem,} \\ 1 & \text{if } \mathcal{Q} \text{ is a maximization problem,} \end{cases}$$

where m is the arity of the predicate S_1.

For the sake of brevity, we shall denote the optimum as $\underset{\mathbf{S}}{opt} \{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\}$, where $opt \in \{\max, \min\}$, but implicitly refer to the precise definition above.

Intuitively, if $\mathcal{Q}$ is an optimization problem in one of the classes defined above, then a feasible solution for an instance $\mathbf{A}$ of $\mathcal{Q}$ is a sequence $\mathbf{S} = (S_1, \cdots, S_t)$ of relations satisfying $\phi(\mathbf{S})$ and the objective function is the cardinality of S_1.

We exhibit next the relationships between the classes of optimization problems defined above and those defined in the preceding Section 2.1.

Let $\mathcal{Q}$ be an optimization problem with optimum on a structure $\mathbf{A}$ expressed as

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \underset{\mathbf{S}}{opt} \{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\},$$

Recall that $\Sigma_n, n \geq 1$, is the class of first-order formulae in prenex normal form that have n alternations of quantifiers, starting with a block of existential quantifiers. For example, Σ_1 is the collection of existential formulae, while Σ_2 is the class of existential-universal formulae. Similarly, Π_n, $n \geq 1$, is the class of first-order formulae in prenex normal form with n alternations of quantifiers, starting with a block of universal quantifiers. Thus, a Π_1 formula has universal quantifiers only, while Π_2 is the collection of universal-existential formulae. The class of quantifier-free formulae is denoted by Σ_0 or Π_0.

Definition 2.4: For each $n \geq 0$, let MAX Π_n be the set of problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ are defined as follows:

$$\text{opt}_{\mathcal{Q}}(\mathbf{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models \phi(\mathbf{w}, \mathbf{S})\}|,$$

where $\phi(\mathbf{w}, \mathbf{S})$ is a Π_n formula with predicate symbols from $\sigma \cup \mathbf{S}$. Similarly, for each $n \geq 0$, we define the class MAX Σ_n by using Σ_n formulae. The classes MIN Π_n and MIN Σ_n, $n \geq 0$ are defined analogously with min in place of max.

The relationships between the various classes of maximization and minimization problems defined above are described in the following results, which were obtained in [KT90].

Theorem 1: The class MAX $\mathcal{PB}$ of all polynomially bounded NP maximization problems coincides with the class MAX Π_2. Thus,

$$\text{MAX}\,\mathcal{PB} = \text{MAX}\,\Pi_2 = \text{MAX}\,\Pi_n, \;\; n \geq 2.$$

The class MIN $\mathcal{PB}$ of all polynomially bounded NP minimization problems coincides with the class MIN Σ_2. Thus,

$$\text{MIN}\,\mathcal{PB} = \text{MIN}\,\Sigma_2 = \text{MIN}\,\Sigma_n, \;\; n \geq 2.$$

Theorem 2: MAX $\mathcal{PB}$ is a hierarchy with four distinct levels, namely,

$$\text{MAX}\,\Sigma_0 \subset \text{MAX}\,\Sigma_1 \subset \text{MAX}\,\Pi_1 \subset \text{MAX}\,\Pi_2.$$

MIN $\mathcal{PB}$ is a hierarchy with three distinct levels, namely,

$$\text{MIN}\,\Sigma_0 \;\subset\; \text{MIN}\,\Sigma_1 \;\subset\; \text{MIN}\,\Pi_1 \;=\; \text{MIN}\,\Sigma_2.$$

In the above, $\subset$ denotes proper containment.

COVER and GRAPH COLORING are problems about finite graphs, while instances of SET COVER can be identified with finite structures as follows. An instance $I = (X, C)$ of the SET COVER problem consists of a set X and a collection C of subsets of X such that $X = \bigcup_{S \in C} S$. It is encoded by a finite structure $\mathbf{A}(I) = (A, C, M)$, where $A = X \cup C$ is the universe of the structure and $M(x, S)$ is a binary predicate expressing membership of an element $x \in X$ in a set $S \in C$.

From now on we assume that the instances of the optimization problems we consider are given as finite structures over some vocabulary σ.

Papadimitriou and Yannakakis [PY88] were the first to use logic in order to define classes of maximization problems and study their approximation properties. They introduced the class MAX NP of maximization problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ can be defined using an existential first-order formula as follows:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})\}|,$$

where $\mathbf{S}$ is a sequence of predicate symbols that are not in σ, and ψ is a quantifier-free formula with predicate symbols amongst those in $\sigma \cup \mathbf{S}$. MAX SAT is a typical example of an optimization problem in MAX NP. Papadimitriou and Yannakakis [PY88] showed that every problem in MAX NP is constant-approximable.

Panconesi and Ranjan [PR90] showed that MAX CLIQUE is not in the class MAX NP. Moreover, they proved that certain polynomial-time optimization problems are not in MAX NP. In an attempt to find a syntactic class of maximization problems containing MAX CLIQUE, they introduced the class MAX Π_1 of maximization problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ can be defined using a universal first-order formula as follows:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models (\forall \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})\}|,$$

where ψ is a quantifier-free formula with predicate symbols from $\sigma \cup \mathbf{S}$. Panconesi and Ranjan [PR90] showed that this class contains problems that are not constant-approximable, unless P = NP.

Kolaitis and Thakur [KT90] introduced a general framework for studying the logical definability of both maximization and minimization problems. More specifically, [KT90] considered optimization problems $\mathcal{Q}$ whose optima on finite structures $\mathbf{A}$ over a vocabulary σ are defined using an *arbitrary* first-order a formula $\phi(\mathbf{w}, \mathbf{S})$ as follows:

$$\mathrm{opt}_{\mathcal{Q}}(\mathbf{A}) = \mathop{opt}_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models \phi(\mathbf{w}, \mathbf{S})\}|.$$

Various classes of optimization problems, including MAX NP and MAX Π_1, can be obtained in this framework by restricting the quantifier complexity of the first-order formulae used.

Definition 2.2: Let $g(n)$ be a function from positive integers to positive reals. We say that an algorithm is a *g-approximation algorithm* for an optimization problem $\mathcal{Q}$ if, given an instance I of $\mathcal{Q}$, the algorithm produces a feasible solution T such that

$$g(|I|) \geq \begin{cases} \frac{c \cdot f(I,T)}{\mathrm{opt}_{\mathcal{Q}}(I)} & \text{if } \mathcal{Q} \text{ is a minimization problem} \\ \frac{c \cdot \mathrm{opt}_{\mathcal{Q}}(I)}{f(I,T)} & \text{if } \mathcal{Q} \text{ is a maximization problem,} \end{cases}$$

where $|I|$ denotes the size of the instance I and c is a constant that depends on the algorithm only. We say that an optimization problem is *g-approximable* if there is a polynomial time g-approximation algorithm for it. An optimization problem is *constant-approximable*, if it is g-approximable for a constant function $g(n)$.

In this paper we will be considering constant-approximable problems and log-approximable problems. The Appendix contains precise definitions of various optimization problems that will be encountered in the sequel. MAX SAT [Joh74] and MIN VERTEX COVER (cf. [PS82]) are examples of constant-approximable problems. MIN SET COVER and MIN DOMINATING SET are examples of log-approximable problems [Joh74]. There are graph problems that are not constant-approximable, unless P = NP, such as the LONGEST-PATH-with-FORBIDDEN-PAIRS [Ber89]. Also the MIN 3NON-TAUTOLOGY problem is not log-approximable, unless P = NP [KT90].

We now restrict attention to *polynomially bounded* NP optimization problems. These are NP optimization problems in which the optimal solution is bounded by a polynomial in the length of the corresponding instance [BJY89,LM81].

Definition 2.3: An NP optimization problem $\mathcal{Q}$ is said to be *polynomially bounded* if there is a polynomial p such that

$$\mathrm{opt}_{\mathcal{Q}}(I) \leq p(|I|) \text{ for all } I \in \mathcal{I}_{\mathcal{Q}}.$$

We denote the class of all polynomially bounded NP maximization (minimization) problems by MAX $\mathcal{PB}$ (MIN $\mathcal{PB}$).

Examples of polynomially bounded NP optimization problems are MAX CLIQUE, MAX SAT, MIN VERTEX COVER, MIN SET COVER, and MIN TRAVELING SALESMAN problem with weights 1 or 2. On the other hand, the unrestricted version of the TRAVELING SALESMAN problem and INTEGER PROGRAMMING are examples of NP optimization problems that are not polynomially bounded.

NP decision problems and, consequently, NP optimization problems can be viewed as problems on finite structures over some vocabulary σ consisting of predicate symbols. In most cases, either an NP problem is described directly as a problem on finite structures, or it can be easily encoded by such a problem. For example, VERTEX

SET, and many other natural optimization problems. We show that some of these problems are actually complete for MIN $F^+\Pi_2(1)$ under appropriate reductions. As a result, every problem in the class MIN $F^+\Pi_2(1)$ is log-approximable. We also consider classes that extend MIN $F^+\Pi_2(1)$ and have weaker approximation properties.

Finally, we compare classes of maximization problems to classes of minimization problems by examining under what conditions a maximization problem can be represented as a minimization problem, and vice versa. In particular, we obtain a machine-independent characterization of the NP $\stackrel{?}{=}$ co-NP question, by showing that NP $\neq$ co-NP if and only if MAX CLIQUE is not in the class MIN Π_1 of minimization problems definable using universal first-order formulae. We also show that MAX CLIQUE is not in certain proper subclasses of MIN Π_1.

2 Logic and NP Optimization Problems

2.1 Background

This section contains the basic definitions and the necessary background material explaining the relationship between logical definability and optimization problems.

Definition 2.1: An NP optimization problem is a tuple $\mathcal{Q} = (\mathcal{I}_\mathcal{Q}, \mathcal{F}_\mathcal{Q}, f_\mathcal{Q}, opt)$ such that

- $\mathcal{I}_\mathcal{Q}$ is the set of input instances. It is assumed that $\mathcal{I}_\mathcal{Q}$ can be recognized in polynomial time.
- $\mathcal{F}_\mathcal{Q}(I)$ is the set of feasible solutions for the input I.
- $f_\mathcal{Q}$ is a polynomial time computable function, called the *objective function*. It takes integer values and is defined on pairs (I, T), where I is an input instance and T is a feasible solution of I.
- $opt \in \{\max, \min\}$.
- The following decision problem is in NP : Given $I \in \mathcal{I}_\mathcal{Q}$ and an integer k, does there exist a feasible solution $T \in \mathcal{F}_\mathcal{Q}(\mathcal{I})$ such that $f_\mathcal{Q}(I, T) \leq k$, when $opt =$ min? (or, $f_\mathcal{Q}(I, T) \geq k$, when $opt = \max$)

We write $\text{opt}_\mathcal{Q}(I)$ to denote the optimum value of $\mathcal{Q}$, on an instance I.

The above definition is essentially due to [PR90] and is broad enough to encompass every known optimization problem arising from the class NP. If the underlying decision problem of an optimization problem is NP-complete, then one cannot expect to have polynomial time algorithms to compute the optimum. In view of this, researchers have studied approximation algorithms and approximation properties of such problems.

In view of the above, it is natural to ask: are there other syntactic properties of formulae that may have implications on the approximation properties of the problems they define? In this paper we introduce an alternative framework for defining and studying optimization problems using logic. Notice that many natural NP optimization problems, including MAX CLIQUE, MIN VERTEX COVER, MIN SET COVER, and MIN GRAPH COLORING, have the following two properties:
(1) a feasible solution is a finite sequence of sets satisfying a first-order sentence (often, the sequence consists of a single set);
(2) the objective function is the cardinality of one of these sets.

Motivated from the above observation, we consider the class of all optimization problems on finite structures $\mathbf{A}$ whose optimum can be defined as

$$\mathop{opt}_{\mathbf{S}} \{|S_1| : \mathbf{A} \models \phi(\mathbf{S})\},$$

where *opt* is either max or min, $\phi(\mathbf{S})$ is an arbitrary first-order sentence and $\mathbf{S} = (S_1, \cdots, S_t)$ is a sequence of predicates. Intuitively, the sequence $\mathbf{S}$ of predicates represents the feasible solution and the objective function is the cardinality of S_1.

We study first the relationships between classes of optimization problems defined in the new framework and those defined in [PY88,PR90,KT90]. In particular, we show that using universal-existential (Π_2) first-order sentences in the new framework we can capture all *polynomially bounded* NP maximization and NP minimization problems.

Can logic be also used to capture all optimization problems having good approximation properties? We address this question here and provide a negative answer to it by establishing that, assuming P $\neq$ NP, it is an *undecidable* problem to tell if a first-order formula gives rise to an optimization problem that is constant-approximable. Similarly, it is an undecidable problem to tell if a first-order formula gives rise to an optimization problem that is log-approximable. As a consequence of the above result, we cannot expect to have "nice" necessary and sufficient conditions that characterize which first-order formulae give rise to optimization problems with good approximation properties. Thus, we can only hope to isolate and study sufficient conditions for approximability. With this insight in mind, we consider certain syntactic conditions on first-order sentences and investigate the approximation properties of the minimization problems defined by such sentences in the new framework.

We reexamine here the class MIN $F^+\Pi_1$, introduced in [KT90]. This class contains MIN VERTEX COVER and has the property that all its members are constant-approximable problems. We observe that MIN $F^+\Pi_1$ contains contains several *node-deletion* and *edge-deletion* problems on graphs (cf. [Yan81a,Yan81b]), but also show that MIN SET COVER is not in this class.

We then introduce the class MIN $F^+\Pi_2(1)$, which is a class of NP minimization problems that contains MIN SET COVER, MIN DOMINATING SET, MIN HITTING

CUT. Papadimitriou and Yannakakis [PY88] showed that every problem in MAX NP is constant-approximable. Thus, perhaps for the first time, we have a structural result that accounts for the common approximation properties of many natural optimization problems.

After this, Panconesi and Ranjan [PR90] showed that MAX CLIQUE does not belong to the class MAX NP. On the other hand, MAX CLIQUE belongs to the class MAX Π_1 of maximization problems whose optimum is definable using universal first-order formulae, i.e., it is of the form

$$\max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models (\forall \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})\}|,$$

where $\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})$ is a quantifier-free formula. Panconesi and Ranjan [PR90] showed also that MAX Π_1 contains optimization problems for which no constant approximation algorithm exists, unless P=NP.

Motivated by the above developments, Kolaitis and Thakur [KT90] undertook a systematic investigation of optimization problems from the logical definability perspective. In particular, they became interested in understanding the exact expressive power of this framework and in discovering other natural classes of optimization problems that can be obtained using this perspective. For this, they examined the class of all maximization problems on finite structures with optimum definable as

$$\max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models \phi(\mathbf{w}, \mathbf{S})\}|,$$

where $\phi(\mathbf{w}, \mathbf{S})$ is an arbitrary first-order formula. Kolaitis and Thakur [KT90] showed that this class coincides with the collection of all *polynomially bounded* NP-maximization problems on finite structures, namely the maximization problems whose decision problem is in NP and whose optimum value is less than or equal to a polynomial of the input size. In the same paper, these problems were classified according to the quantifier complexity of the formulae used in defining them, and it was shown that the polynomially bounded NP-maximization problems form a proper hierarchy with exactly four distinct levels. Kolaitis and Thakur [KT90] also examined minimization problems and showed that logical definability has different implications for NP-minimization problems than it has for NP-maximization problems in terms of both expressive power and approximation properties. In particular, the polynomially bounded NP-minimization problems form a proper hierarchy with exactly three distinct levels. In addition, it turns out that the quantifier pattern of the formulae used to define NP-minimization problems does not impact on the approximation properties of the problems, unlike the case of NP-maximization problems considered in [PY88]. More specifically, unlike MAX NP, the class of minimization problems definable using existential first-order formulae contains problems that are not constant-approximable, if P $\neq$ NP.

1 Introduction and Summary of Results

Optimization problems had a major influence on the development of the theory of NP-completeness. Indeed, many NP-complete problems are decision problems that are derived from natural optimization problems by imposing a bound on the objective function (cf. [GJ79]). Turing machines and polynomial-time reductions have provided a robust computational model for classifying and studying decision problems. In contrast, the absence of a robust model for optimization problems has been a serious impediment to the development of a structural optimization theory.

Johnson [Joh74] initiated a classification of optimization problems according to their *approximation* properties. For some concrete problems, such as MAX SAT, Johnson [Joh74] found algorithms that provide approximate solutions within a constant factor of the optimum (*constant-approximable problems*). For other concrete problems, such as MIN SET COVER, he exhibited algorithms with the property that the ratio of the worst case to the optimum is bounded by the logarithm of the input size (*log-approximable problems*). At the end of that paper, Johnson [Joh74] raises a number of interesting questions, all of which have to do with the development of a structural optimization theory that would explain the similarities and the differences between the approximation properties of various optimization problems. In Johnson's words: *"What is it that makes algorithms for different problems behave the same way? Is there some stronger kind of reducibility than the simple polynomial reducibility that will explain these results, or are they due to some structural similarity between the problems as we define them? And what other types of behavior and ways of analyzing and measuring it are possible?"*

Johnson's questions remained largely unanswered for a number of years. In 1988, Papadimitriou and Yannakakis [PY88] brought a new insight to this area of research by focusing on the logical definability of optimization problems. It is worth pointing out that they were motivated by Fagin's [Fag74] theorem, which asserts that a class of finite structures is NP-computable if and only if it is definable by an existential second-order sentence. Papadimitriou and Yannakakis [PY88] introduced the class MAX NP of optimization problems on finite structures **A** whose optimum can be defined as

$$\max_{\mathbf{S}} |\{\mathbf{w} : \mathbf{A} \models (\exists \mathbf{x})\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})\}|,$$

where $|...|$ denotes the cardinality of a set, $\psi(\mathbf{w}, \mathbf{x}, \mathbf{S})$ is a quantifier-free formula, $\exists \mathbf{x}$ is an existential first-order quantifier, and $\mathbf{S}$ is a sequence of second-order variables, i.e., it ranges over relations of fixed arities on the structure $\mathbf{A}$. Intuitively, the second-order variables $\mathbf{S}$ correspond to the existential second-order quantifiers in Fagin's [Fag74] characterization of NP.

The class MAX NP contains many natural optimization problems, including MAX 3SAT, MAX SAT (appropriately encoded as a problem on finite structures), and MAX

Approximation Properties of NP Minimization Classes *

Phokion G. Kolaitis
Madhukar N. Thakur†

UCSC-CRL-91-10
April 1991

Computer and Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064

Abstract: In this paper we introduce a new approach to the logical definability of NP optimization problems by focusing on the expressibility of feasible solutions. We show that in this framework first-order sentences capture exactly all polynomially bounded optimization problems. We also show that, assuming P $\neq$ NP, it is an undecidable problem to tell if a given first-order sentence defines an approximable optimization problem. We then isolate a syntactically defined class of NP minimization problems that contains the MIN SET COVER problem and has the property that every problem in it has a logarithmic approximation algorithm. We conclude by giving a machine-independent characterization of the NP $\stackrel{?}{=}$ co-NP problem in terms of logical expressibility of the MAX CLIQUE problem.

---

*Research partially supported by NSF Grant CCR-8905038
To appear in Proceedings of the IEEE Structure in Complexity Theory, Sixth Annual Conference, 1991
†email addresses: kolaitis@cs.ucsc.edu, thakur@cs.ucsc.edu