Completeness in the Polynomial Time Hierarchy

Marcus Schaefer
School of CTI
DePaul University
243 South Wabash Avenue
Chicago, Illinois 60604, USA
schaefer@cs.depaul.edu

July 2, 2001

Abstract

We are trying to compile a list of problems that reside in the polynomial hieararchy above the second level. The current list is incomplete: it does not contain some recent results (such as Stockmeyer and Modha's results on block-coding), nor does it list any of the results on petri-nets, non-monotonic logics, and databases (to mention some areas not included so far). It should also be pointed out that the results themselves have not been verified for the purposes of this report.

1 The Second level

1.1 Logic

[L1] $\forall\exists$**3SAT**

Given: $\Psi = (\forall x)(\exists y)\varphi(x, y)$, where φ is in 3CNF.

Question: Is Ψ true?

Reference: Stockmeyer [Sto76], Wrathall [Wra76].

Comments: $\mathbf{\Pi_2^p}$-complete. Stockmeyer and Wrathall show that deciding QSAT_k, the set of true formulas with $k-1$ quantifier alternations beginning with a $\exists$ quantifier is $\mathbf{\Sigma_k^p}$-complete. Earlier, Meyer and Stockmeyer [MS72] had shown that QUANTIFIED BOOLEAN FORMULAE, the problem of deciding the truth of quantified Boolean formulas (without restriction on the number of alternations), is **PSPACE**-complete. See MINMAX SAT for the optimization variant.

[L2] NAE$\forall\exists$**3SAT**

Given: $\Psi = (\forall x)(\exists y)\varphi(x, y)$, where φ is in 3CNF.

Question: Is Ψ *nae-valid*, i.e. for every truth-assignment to x is there a truth-assignment to y such that each clause in $\varphi(x, y)$ contains both a true and a false literal.

Reference: Eiter, Gottlob [EG95]?

Comments: $\mathbf{\Pi_2^p}$-complete. Relativized version of NAE3SAT.

[L3] GENERALIZED 3-CNF CONSISTENCY

Given: Two sets A and B of Boolean formulas.

Question: Is there a Boolean formula φ such that $\varphi \wedge \psi$ is satisfiable for all $\psi \in A$, and unsatisfiable for all $\psi \in B$?

Reference: Ko, Tzeng [KT91].

Comments: $\mathbf{\Sigma_2^p}$-complete. Similar in structure to PATTERN CONSISTENCY, and GRAPH CONSISTENCY.

[L4] MIN DNF

Given: A DNF formula φ. The *size* $|\varphi|$ of a formula is the number of occurrences of literals in the formula.

Question: There is no DNF formula ψ such that $\psi \equiv \varphi$ and $|\psi| < |\varphi|$.

Reference: Umans [Uma98].

Comments: $\mathbf{\Sigma_2^p}$-complete. The problem is also known as $\mathsf{MEE_{DNF}}$, and MIN. Reduction from SHORTEST IMPLICANT CORE. If we drop the restriction to DNF formulas, we obtain MEE, whose complexity is not known. The main open variant is MINIMAL.

[L5] SHORTEST IMPLICANT CORE

Given: A DNF formula φ, an implicant C of ϕ, and a number k. An *implicant* of φ is a set of literals whose conjunction implies ϕ. The size of an implicant is its size as a set.

Question: There is no implicant $C' \subseteq C$ of φ of size at most k.

Reference: Umans [Uma98].

Comments: $\mathbf{\Sigma_2^p}$-complete.

[L6] MINMAX SAT

Given: $\Psi = (\forall x)(\exists y)\varphi(x, y)$ and integer k, where φ is in 3CNF.

Question: For every truth-assignment to x, is there a truth-assignment to y making at least k clauses in $\varphi(x, y)$ true?

Reference: Meyer, Stockmeyer [MS72].

Comments: This is the optimization version of $\forall\exists$3SAT. Let us call $f(\varphi)$ the largest k such that for every truth-assignment to x there is a truth-assignment to y making at least k clauses in $\varphi(x, y)$ true. Then there is a $c > 0$ such that approximating $f(\varphi)$ to within a factor of c is still $\mathbf{\Pi_2^p}$-complete. This follows from work of Condon, Feigenbaum, Lund and Shor on debate systems (generalizing the PCP characterization of **NP**) [KL95a]. Ko and Lin [KL95b] showed that the c-approximation problem remains $\mathbf{\Pi_2^p}$-complete if the number of occurrences of each variable is bounded by the same constant (MINMAX SAT B). They use this result to show that LONGEST DIRECTED CIRCUIT is $\mathbf{\Pi_2^p}$-complete.

[L7] $\exists^*\forall^*$ SATISFIABILITY IN FOL WITH ONE UNARY FUNCTION

Given: A first-order formula φ whose prenex belongs to $\exists^*\forall^*$, where φ may contain equality and one unary function, but no other constant, function, or relation symbols

Question: Is there a model for φ?

Reference: Börger, Grädel, Gurevich [BGG97, Theorem 6.4.19].

Comments: $\mathbf{\Sigma_2^p}$-complete; the harder part being membership in $\mathbf{\Sigma_2^p}$. This is a special case of Ramsey's decidability result of the satisfiability problem for $\exists^*\forall^*$ formulas with equality, but no other relation symbols (which is **NEXP**-complete). The following variants of the satisfiability problem are also $\mathbf{\Sigma_2^p}$-complete: prenex in $\exists^*\forall^*$, and no relation or function symbols except for equality; prenex in $\exists^*\forall^*$, and at most one unary relation (no function symbols, no equality); prenex in $\exists^2\forall^*$, and relations of arbitrary arity (no functions, no equality). See [BGG97, Theorem 6.4.7].

1.2 Graph Theory

[GT1] GRAPH CONSISTENCY

Given: Two sets A and B of (finite) graphs.

Question: Is there a graph G such that every graph in A is isomorphic to a subgraph of G, but no graph in B is isomorphic to a subgraph of G?

Reference: Ko, Tzeng [KT91] (GRAPH RECONSTRUCTION).

Comments: $\mathbf{\Sigma_2^p}$-complete. Similar in structure to PATTERN CONSISTENCY, and GENERALIZED 3-CNF CONSISTENCY.

[GT2] MINMAX CLIQUE

Given: Graph $G = (V, E)$, a partition $(V_{i,j})_{i\in I, j\in J}$ of V, integer k. For a function $t : I \to J$ let f_t be the size of the largest clique in G restricted to $\bigcup_{i\in I} V_{i,t(i)}$.

Question: Is $\min_{t\in J^I} f_t(G) \geq k$?

Reference: Ko, Lin [KL95a].

Comments: $\mathbf{\Pi_2^p}$-complete. There is a $c > 0$ such that approximating $\min_{t\in J^I} f_t(G)$ is still $\mathbf{\Pi_2^p}$-complete. Also see MAXMIN VERTEX COVER.

[GT3] MINMAX CIRCUIT

Given: Graph $G = (V, E)$, a partition $(V_{i,j})_{i\in I, j\in J}$ of V, integer k. For a function $t : I \to J$ let f_t be the length of the longest cycle in G restricted to $\bigcup_{i\in I} V_{i,t(i)}$.

Question: Is $\min_{t\in J^I} f_t(G) \geq k$?

Reference: Ko, Lin [KL95a].

Comments: $\mathbf{\Pi_2^p}$-complete.

[GT4] DYNAMIC HAMILTONIAN CIRCUIT

Given: Graph $G = (V, E)$, subset B of E. Let $G_D = (V, E - D)$.

Question: Is G_D Hamiltonian (is there a cycle containing all vertices) for all D with $|D| \leq |B|/2$?

Reference: Ko, Lin [KL95a].

Comments: $\mathbf{\Pi_2^p}$-complete.

[GT5] LONGEST DIRECTED CIRCUIT

Given: Directed graph $G = (V, E)$, and a subset E' of E of *alterable* edges, integer k. For $D \subseteq E$ let G_D be the graph obtained from G by substituting each edge (u, v) in D by its reverse edge (v, u). Define f_D to be the length of the longest cycle in G_D.

Question: Is $l(G) = \min_{D\subseteq E'} f_D \geq k$?

Reference: Ko, Lin [KL95b].

Comments: $\mathbf{\Pi_2^p}$-complete. There is a constant $c > 0$ such that approximating $l(G)$ to within a factor of c is still $\mathbf{\Pi_2^p}$-complete. Reduction from MINMAX SAT B (see MINMAX SAT).

[GT6] GENERALIZED GRAPH COLORING

Given: Graphs F, G.

Question: Is there a two-coloring of the vertices of F which does not contain a monochromatic G as a subgraph?

Reference: Rutenburg [Rut86].

Comments: $\mathbf{\Sigma_2^p}$-complete even if G is restricted to be complete. The completeness proof also works for other **coNP**-complete families of graphs, see, for example, the GENERALIZED NODE DELETION problem. For edge colorings compare ARROWING, and STRONG ARROWING.

[GT7] GENERALIZED NODE DELETION

Given: Graphs F, G, integer k.

Question: Can we remove at most k vertices from F such that the resulting graph does not contain G as a subgraph?

Reference: $\mathbf{\Sigma_2^p}$-completeness is claimed in Rutenburg [Rut86] without proof.

Comments: $\mathbf{\Sigma_2^p}$-complete even if G is restricted to be complete.

[GT8] GENERALIZED RAMSEY NUMBER

Given: Graph F, a partial two-coloring of the edges of F, integer k.

Question: Does every two-coloring of F which extends the given two-coloring contain a clique on k vertices.

Reference: Ko and Lin [KL95a]. A proof can also be found in [DK00].

Comments: $\mathbf{\Pi_2^p}$-complete. See ARROWING.

[GT9] ARROWING

Given: Graphs F, G, and H.

Question: Does $F \to (G, H)$, i.e. does every edge-coloring of F with colors red and green contain either a red G, or a green H as a subgraph?

Reference: Schaefer [Sch99a]

Comments: $\mathbf{\Pi_2^p}$-complete even if G is a fixed tree on at least three vertices, and H a complete graph. The problem is **coNP**-complete for fixed three-connected graphs G and H [Bur90]. If F is a complete graph, then the problem is **NP**-hard [Bur84], but not known to be $\mathbf{\Pi_2^p}$-complete. $K_n \to (K_m, K_\ell)$ is unlikely to be $\mathbf{\Pi_2^p}$-complete, since it lies in $\mathbf{coNP}^{\mathsf{LOGCLIQUE}}$, where LOGCLIQUE is the problem of deciding whether a graph F has a clique of size at least $\log |F|$. This version is particularly interesting since it corresponds to computing Ramsey numbers. Also see STRONG ARROWING, and GENERALIZED RAMSEY NUMBER. The vertex-coloring version of this problem is called GENERALIZED GRAPH COLORING.

[GT10] STRONG ARROWING

Given: Graphs F, G, and H.

Question: Does $F \rightarrowtail (G, H)$, i.e. does every edge-coloring of F with colors red and green contain either a red G, or a green H as an induced subgraph of F?

Reference: Schaefer [Sch99a]

Comments: $\mathbf{\Pi_2^p}$-complete if G is a fixed star $K_{1,p}$, and H a complete graph, or $G = H = K_{1,n}$ (the diagonal case). The noninduced version $F \to (K_{1,n}, K_{1,m})$ is in **P** [Bur90]. Also see ARROWING.

1.3 Sets and Partitions

[SP1] MAXMIN VERTEX COVER

Given: Graph $G = (V, E)$, a partition $(V_{i,j})_{i \in I, j \in J}$ of V, integer k. For a function $t : I \to J$ let f_t be the size of a smallest vertex cover of G restricted to $\bigcup_{i \in I} V_{i,t(i)}$.

Question: Is $\max_{t \in J^I} f_t(G) \le k$?

Reference: Ko, Lin [KL95a].

Comments: $\mathbf{\Pi_2^p}$-completeness follows from $\mathbf{\Pi_2^p}$-completeness of MINMAX CLIQUE using the standard transformation between vertex covers and cliques. According to Ko and Lin the nonapproximability result for MINMAX CLIQUE does not carry over.

[SP2] MINMAX THREE DIMENSIONAL MATCHING

Given: Set W, partition $(W_{i,j})_{i \in I, j \in J}$ of W, set S of three-element subsets of W, and a number k. Call a set $S' \subseteq S$ a *matching in* $W' \subseteq W$, if the sets in S' are mutually disjoint subsets of W'. For a function $t : I \to J$ let $f_t(W)$ be the size $|S'|$ of a largest matching S' in $\bigcup_{i \in I} W_{i,t(i)}$.

Question: Is $\min_{t \in J^I} f_t(W) \ge k$?

Reference: Ko, Lin [KL95a].

Comments: $\mathbf{\Pi_2^p}$-complete; reduction from MINMAX SAT YB. There is a $c > 0$ such that approximating $\min_{t \in J^I} f_t(W)$ is still $\mathbf{\Pi_2^p}$-complete.

1.4 Algebra and Number Theory

[AN1] INTEGER EXPRESSION INEQUIVALENCE

Given: Two integer expressions e_1, and e_2 built from binary numbers with operators $+$, and $\cup$. For an integer expression e define $L(e) = \{n\}$ if e is the binary representation of n, $L(e+f) = \{n+m : n \in L(e), m \in L(f)\}$, and $L(e \cup f) = L(e) \cup L(f)$.

Question: Is $L(e_1) \neq L(e_2)$?

Reference: Stockmeyer, Meyer [GJ79]. The result appears in a 1973 conference paper by Stockmeyer and Meyer, and a 1976 paper by Stockmeyer.

Comments: Probably the first natural problem to be shown $\mathbf{\Sigma_2^p}$-complete. The subset problem $L(e_1) \subseteq L(e_2)$ is also $\mathbf{\Pi_2^p}$-complete. The same is true for expressions represented in the general hierarchy input language (GHIL) which according to Wagner [Wag86] was shown by Bentley, Ottmann, and Widmayer (1983). See INTEGER EXPRESSION CONNECTEDNESS, and INTEGER EXPRESSION COMPONENT LENGTH.

[AN2] INTEGER EXPRESSION CONNECTEDNESS

Given: An integer expression e built from binary numbers with operators $+$, and $\cup$. For an integer expression e define $L(e) = \{n\}$ if e is the binary representation of n, $L(e+f) = \{n+m : n \in L(e), m \in L(f)\}$, and $L(e \cup f) = L(e) \cup L(f)$. A set of numbers L is called *connected* if for every $x, z \in L$ and any y, if $x < y < z$ then $y \in L$.

Question: Is $L(e)$ connected?

Reference: Wagner [Wag86].

Comments: $\mathbf{\Pi_2^p}$-complete. The result also holds if the input is specified using the general hierarchic input language (GHIL).

[AN3] BOOLEAN EXPRESSION COMPONENT LENGTH

Given: A Boolean formula φ, integer k. If φ has n Boolean input variables $x_1, \ldots, x_n$ we let $L(\varphi) = \{x_1 \cdots x_n : \varphi(x_1, \ldots, x_n)\}$ interpreting the binary vector as a natural number. A set of numbers L is called *connected* if for every $x, z \in L$ and any y, if $x < y < z$ then $y \in L$. A maximal connected subset of a set is called a *component*.

Question: Does $L(\varphi)$ have a component of size at least k?

Reference: Wagner [Wag86].

Comments: $\mathbf{\Sigma_2^p}$-complete. For integer expressions the problem is $\mathbf{\Sigma_3^p}$-complete (INTEGER EXPRESSION COMPONENT LENGTH).

[AN4] BOUNDED EIGENVECTOR

Given: $n \times n$ integer matrix M, eigenvalue λ of M, $I \subseteq \{1, \ldots, n\}$, number y.

Question: Is there an eigenvector $\mathbf{x} = (x_1, \ldots, x_n)$ (for λ) such that $x_1 = y$, $|x_i| \leq c$ (for some fixed c), and $\mathbf{x}$ has maximal ℓ_2-norm among vectors identical to $\mathbf{x}$ on I?

Reference: Eiter, Gottlob [EG95].

Comments: $\mathbf{\Sigma_2^p}$-complete for any fixed $c \geq 1$, and $z = 0$. Reduction from NAE∀∃3SAT, a relativized version of NAE3SAT.

[AN5] SEMILINEAR SET EQUIVALENCE

Given: Finite sets $C_i, P_i, C'_i, P'_i \subseteq \mathbb{N}^k$ $(1 \leq i \leq n)$. Let $L(C, P) = \{c + \sum_{p \in P} \lambda_p p : c \in C, p \in P, \lambda \in \mathbb{N}\}$, and $SL(C_1, \ldots, C_n; P_1, \ldots P_n) = \bigcup_{i=1}^n L(C_i, P_i)$. Sets of the form L are called *linear*, sets of the form SL *semilinear*.

Question: Is $SL(C_1, \ldots, C_n; P_1, \ldots P_n) = SL(C'_1, \ldots, C'_n; P'_1, \ldots P'_n)$?

Reference: Huynh [Huy82].

Comments: $\mathbf{\Pi_2^p}$-complete, even for $k = 1$.

1.5 Automata and Languages

[AL1] PATTERN CONSISTENCY

Given: Two sets A and B of strings over $\{0, 1\}$. A *pattern* is a string over $\{0, 1\}$ and a set of variables. The language $L(p)$ associated with a pattern p is the set of strings that can be obtained from p by substituting all variables in p by strings over $\{0, 1\}$.

Question: Is there a pattern p such that $A \subseteq L(p) \subseteq \overline{B}$.

Reference: Ko, Tzeng [KT91].

Comments: $\mathbf{\Sigma_2^p}$-complete. Similar in structure to GRAPH CONSISTENCY, and GENERALIZED 3-CNF CONSISTENCY.

[AL2] POMSET LANGUAGE CONTAINMENT

Given: Two POMSETS P, Q. A POMSET (partially ordered multiset) is a directed acyclic graph (V, E) whose vertices have labels in Σ. The language $L(P)$ associated with a POMSET P is the set of words of length $n = |V|$ over Σ that corresponds to a permutation of the vertices in V which is consistent with the partial order generated by (V, E).

Question: Is $L(P) \subseteq L(Q)$?

Reference: Feigenbaum, Kahn, Lund [FKL93].

Comments: $\mathbf{\Pi_2^p}$-complete. The language membership problem is **NP**-complete, and determining the size of $L(P)$ is span-P complete. Determining whether $L(P) = L(Q)$ obviously lies in $\mathbf{\Pi_2^p}$, and Feigenbaum, Kahn, and Lund showed that it is at least as hard as GRAPH ISOMORPHISM.

[AL3] STAR-FREE REGULAR EXPRESSION WITH INTERLEAVING EQUIVALENCE

Given: Two regular expressions e_1, e_2 using union, concatenation, and interleaving. For two words $x, y \in \{0,1\}^*$ the operation $|$ of interleaving x and y results in the set $x|y$ containing all words $x_1 y_1 \dots x_k y_k$ such that $x = x_1 \dots x_k$, and $y = y_1 \dots y_k$, where the y_i can have any length (including zero).

Question: Are e_1 and e_2 equivalent, i.e. do they describe the same set of words?

Reference: Mayer, Stockmeyer [MS94].

Comments: $\mathbf{\Pi_2^p}$-complete. The proof is based on Stockmeyer's INTEGER EXPRESSION INEQUIVALENCE result without being a reduction. There are many versions of the regular expression problem. The standard version has union, concatenation, and Kleene star, and it is **PSPACE**-complete [GJ79, AL9]. Adding interleaving, or intersection (Hunt, 1973; according to [MS94]) makes is exponential space-complete. Removing both the Kleene star and interleaving gives an **NP**-complete problem (Hunt, Stockmeyer and Mayer, 1973; according to [MS94]).

[AL4] TRIE$_2$

Given: A sequence Π of patterns of length n and an integer k. A *pattern* is a string in $\{0, 1, *\}^*$. A *call* is a string over $\{0, 1\}$ ($*$ matches both 0 and 1). A *TRIE T* is an ordered rooted tree (i.e. the order of a depth-first search traversal is specified) whose edges have labels in $\{1, 2, \dots\} \times \{0, 1, *\}$. A *TRIE T for* Π is a TRIE which has as many leaves as Π has patterns. Furthermore if τ is the set of labels along the path to the jth leaf reached in the fixed depth-first search traversal of T, then τ needs to be equal to the jth pattern in Π (for all j). For a call c let $m(c, T)$ be the number of edges that a depth-first traversal of T will visit (we do not continue along an edge whose label is not consistent with c). Intuitively $m(c, T)$ is the number of matches performed by the TRIE to find all patterns in Π matching c.

Question: Is there a call c such that $m(c, T) \geq k$ for all TRIEs T for Π.

Reference: Lin [Lin95].

Comments: $\mathbf{\Pi_2^p}$-complete. There is a constant $0 < \delta < 1$ for which approximating $f(\Pi) = \min_T \max_c m(c, T)$ to within a factor of δ is still $\mathbf{\Pi_2^p}$-complete. For a fixed TRIE the problem is **NP**-complete (Dawson, Ramakrishnan, Ramakrishnan, Swift, 1994; according to [Lin95]).

2 The Third Level

2.1 Graph Theory

[GT1] PATH VC DIMENSION

Given: A graph $G = (V, E)$, and a number k. Let $VC_{\text{path}}(G)$ be the size of the largest set $X \subseteq V$ which is shattered by subpaths of G, i.e. such that for each $S \subseteq X$ there is a subpath of G containing all vertices in S, but no vertex of $X \setminus S$.

Question: Is $VC_{\text{path}}(G) \geq k$?

Reference: Schaefer [Sch00].

Comments: Special case of the GRAPH VC DIMENSION problem defined for types of subgraphs of a given graph. Introduced by Kranakis, *et al.* [KKR$^+$97] building on an idea of Haussler and Welzl [HW87]. The problem is also $\mathbf{\Sigma_3^p}$-complete for cyles instead of paths [Sch00]. All other cases investigated so far turn out to be in **P** (stars, neighborhoods), or **NP**-complete (trees, connected sets) [KKR$^+$97]. Also see VC DIMENSION.

2.2 Sets and Partitions

[SP1] VC DIMENSION

Given: A Boolean circuit C with $n+\ell$ input gates, and a number k. Let $C_i := \{x \in \{0,1\}^\ell : C(i,x) = 1\}$ ($i \in \{0,1\}^n$), and $\mathcal{C} := \{C_i : i \in \{0,1\}^n\}$.

Question: Is $VC(\mathcal{C}) \geq k$, i.e. is there a set $X \subseteq \{0,1\}^\ell$ containing at least k elements, such that for every $S \subseteq X$ there is an $i \in \{0,1\}^n$ such that $S = C_i \cap X$?

Reference: Schaefer [Sch99b].

Comments: $\mathbf{\Sigma_3^p}$-complete. Remains $\mathbf{\Sigma_3^p}$-complete to approximate to within any factor of $2-\varepsilon$, but can be approximated to within a factor of 2 in **AM** (both results are due to Mossel and Umans [MU01]). If the concept class $\mathcal{C}$ is represented by a matrix instead of a Boolean circuit the problem is **LOGNP**-complete as shown by Papadimitriou and Yannakakis [PY93]. Also see GRAPH VC_{path} DIMENSION.

2.3 Algebra and Number Theory

[AN1] INTEGER EXPRESSION COMPONENT LENGTH

Given: An integer expression e built from binary numbers with operators $+$, and $\cup$, and a number k. For an integer expression e define $L(e) = \{n\}$, if e is the binary representation of n, $L(e+f) = \{n+m : n \in L(e), m \in L(f)\}$, and $L(e \cup f) = L(e) \cup L(f)$. A set of numbers L is called *connected* if for every $x, z \in L$ and any y, if $x < y < z$ then $y \in L$. A maximal connected subset of a set is called a *component*.

Question: Does $L(e)$ have a component of size at least k?

Reference: Wagner [Wag86].

Comments: $\mathbf{\Sigma_3^p}$-complete. The result also holds if using the general hierarchic input language (GHIL) for specifying the input. If the set of integers is specified by a Boolean formula, the problem is $\mathbf{\Sigma_2^p}$-complete (BOOLEAN EXPRESSION COMPONENT LENGTH). See INTEGER EXPRESSION INEQUIVALENCE, INTEGER EXPRESSION CONNECTEDNESS.

3 The fourth level

4 Open Problems

[O2] RAMSEY

Given: Finite Graphs G, and H.

Question: Does $K_n \to (G, H)$, i.e. does every edge-coloring of K_n with colors red and green contain either a red G, or a green H as a subgraph.

Comments: See ARROWING.

[O3] MINIMUM EQUIVALENT EXPRESSION

Given: A Boolean formula φ, integer k. The *size* $|\varphi|$ of a formula is the number of occurrences of literals in the formula.

Question: Is there a Boolean formula ψ for which $\psi \equiv \varphi$, and $|\psi| < k$?

Reference: Garey, Johnson [GJ79].

Comments: **coNP**-hard (trivially), and in $\mathbf{\Sigma_2^p}$. Shown to be hard for $\mathbf{P}_{||}^{\mathbf{NP}}$ (**P** with parallel access to **NP**) by Hemaspaandra and Wechsung [HW97]. Restricted to DNF formulas the problem is $\mathbf{\Sigma_2^p}$-complete (MIN DNF). Also see MINIMAL.

[O4] MINIMAL

Given: A Boolean formula φ. The *size* $|\varphi|$ of a formula is the number of occurrences of literals in the formula.

Question: There is no formula ψ such that $\psi \equiv \varphi$ and $|\psi| < |\varphi|$.

Reference: Meyer, Stockmeyer [MS72].

Comments: **coNP**-hard [HW97], and in $\mathbf{\Pi_2^p}$. Also see MEE, and MIN DNF.

[O5] POMSET LANGUAGE EQUALITY

Given: Two POMSETS P, Q. A POMSET (partially ordered multiset) is a directed acyclic graph (V, E) whose vertices have labels in Σ. The language $L(P)$ associated with a POMSET P is the set of words of length $n = |V|$ over Σ that corresponds to a permutation of the vertices in V which is consistent with the partial order generated by (V, E).

Question: Is $L(P) = L(Q)$?

Reference: Feigenbaum, Kahn, Lund [FKL93].

Comments: In $\mathbf{\Pi_2^p}$, and at least as hard as GRAPH ISOMORPHISM. See POMSET LANGUAGE CONTAINMENT.

[O6] MASTER TOUR

Given: Graph $G = (V, E)$ in the Euclidean plane.

Question: Does G have the *master tour* property, i.e. is there a path P in G such that for any set of vertices $U \subseteq V$ an optimal tour through all the cities in U can be obtained by restricting the path P to the vertices in U.

Reference: Papadimitriou [Pap94, Problem 17.3.11].

Comments: In $\mathbf{\Sigma_2^p}$.

[O7] SHORT CNF

Given: A Boolean expression φ in disjunctive normal form, integer k.

Question: Is φ equivalent to a Boolean expression in conjunctive normal form of at most k clauses.

Reference: Papadimitriou [Pap94, Problem 17.3.12].

Comments: In $\mathbf{\Sigma_2^p}$.

References

[BGG97] Egon Börger, Erich Grädel, and Yuri Gurevich. *The Classical Decision Problem*. Springer Verlag, Berlin, 1997.

[Bur84] Stefan A. Burr. Some undecidable problems involving the edge-coloring and vertex-coloring of graphs. *Discrete Mathematics*, 50:171–177, 1984.

[Bur90] Stefan A. Burr. On the Computational Complexity of Ramsey-Type Problems. In Nešetřil & Rödl, editor, *Mathematics of Ramsey Theory*. Springer-Verlag, 1990.

[DK00] Ding-Zhu Du and Ker-I Ko. *Theory of Computational Complexity*. Wiley, 2000.

[EG95] Thomas Eiter and Georg Gottlob. Note on the complexity of some eigenvector problems. Technical Report CD-TR 95/89, Christian Doppler Laboratory for Expert Systems, TU Vienna, December 1995.

[FKL93] Joan Feigenbaum, Jeremy A. Kahn, and Carsten Lund. Complexity Results for POMSET Languages. *SIAM Journal on Discrete Mathematics*, 6:432–442, 1993.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and Intractability.* Freeman, San Francisco, 1979.

[Huy82] Thiet-Dung Huynh. The Complexity of Semilinear Sets. *Journal of Information Processing and Cybernetics*, 18(6):291–338, 1982.

[HW87] David Haussler and Emo Welzl. Epsilon-nets and simplex range queries. *Discrete and Computational Geometry*, 2:127–151, 1987.

[HW97] Edith Hemaspaandra and Gerd Wechsung. The Minimization Problem for Boolean Formulas. In *38th Annual Symposium on Foundations of Computer Science*, pages 575–584, Miami Beach, Florida, 20–22 October 1997. IEEE.

[KKR+97] Evangelos Kranakis, Danny Krizanc, Berthold Ruf, Jorge Urrutia, and Gerhard Woeginger. The VC-dimension of set systems defined by graphs. *Discrete Applied Mathematics and Combinatorial Operations Research and Computer Science*, 77:237–257, 1997.

[KL95a] Ker-I Ko and Chih-Long Lin. On the complexity of min-max optimization problems and their approximation. In Ding-Zhu Du and Panos M. Pardalos, editors, *Minimax and Applications*, pages 219–239. Kluwer Academic Publishers, Boston, 1995.

[KL95b] Ker-I Ko and Chih-Long Lin. On the longest circuit in an alterable digraph. *Journal of Global Optimization*, 7:279–295, 1995.

[KT91] Ker-I Ko and Wen-Guey Tzeng. Three Σ_2^p-complete Problems in Computational Learning Theory. *Computational Complexity*, 1:269–310, 1991.

[Lin95] Chih-Long Lin. Optimizing TRIEs for Ordered Pattern Matching is Π_2^p-Complete. In *Proceedings of the 10th Annual Conference on Structure in Complexity Theory (SCTC '95)*, pages 238–245, Los Alamitos, CA, USA, June 1995. IEEE Computer Society Press.

[MS72] Albert R. Meyer and Larry J. Stockmeyer. The equivalence problem for regular expressions with squaring requires exponential space. In *13th Annual Symposium on Switching and Automata Theory*, pages 125–129, The University of Maryland, 25–27 October 1972. IEEE.

[MS94] Alain J Mayer and Larry J. Stockmeyer. The complexity of word problems - this time with interleaving. *Information and Computation*, 115:293–311, 1994.

[MU01] Elchanan Mossel and Christopher Umans. On the complexity of approximating the VC dimension. In *Proceedings of the 16th Annual Conference on Computational Complexity*, pages 220–225. IEEE Computer Society Press, 2001.

[Pap94] Christos H. Papadimitriou. *Computational Complexity.* Addison-Wesley, New York, 1994.

[PY93] Christos H. Papadimitriou and Mihalis Yannakakis. On Limited Nondeterminism and the Complexity of the V-C Dimension. In *Proceedings of the 8th Annual Conference on Structure in Complexity Theory (SCTC '93)*, pages 12–18, San Diego, CA, USA, May 1993. IEEE Computer Society Press.

[Rut86] Vladislav Rutenburg. Complexity of generalized graph coloring. In Jozef Gruska, Branislav Rovan, and Juraj Wiedermann, editors, *Mathematical Foundations of Computer Science, 12th Symposium*, volume 233 of *lncs*, pages 573–581, Bratislava, Czechoslovakia, 25–29 August 1986. Springer.

[Sch99a] Marcus Schaefer. Graph ramsey theory and the polynomial hierarchy. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 592–601. ACM, 1999.

[Sch99b] Marcus Schäfer. Deciding the vapnik-cervonenkis dimension is Σ_3^P-complete. *Journal of Computer and System Sciences*, 58:177–182, 1999.

[Sch00] Marcus Schaefer. Deciding the vapnik-cervonenkis dimension is Σ_3^P-complete, ii. Technical Report TR00-006, DePaul University, October 2000.

[Sto76] Larry J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):1–22, October 1976.

[Uma98] Christopher Umans. The Minimum Equivalent DNF Problem and Shortest Implicants. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 556–563, 1998.

[Wag86] Klaus W. Wagner. The Complexity of Combinatorial Problems with Succinct Input Representation. *Acta Informatica*, 23(3):325–356, 1986.

[Wra76] C. Wrathall. Complete sets and the polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):23–33, October 1976.