Inverse NP Problems

Hubie Chen*

Abstract

One characterization of the class NP is as the class of all languages for which there exists a verifier, operating in polynomial time, with the following properties: for every member of the language, there exists a polynomially-sized proof causing the verifier to accept; and, for every non-member, there is no proof causing the verifier to accept. Relative to a particular verifier, every member x of the language induces a set of proofs, namely, the set of proofs causing the verifier to accept x as a member. This paper studies the complexity of deciding, given a set Π of proofs, whether or not there exists some x inducing Π (relative to a particular verifier). We call this decision problem the inverse problem for the verifier.

We develop a new notion of reduction which allows one to compare the complexity of inverse problems. Using this notion, we classify as coNP-complete the inverse problems for the "natural" verifiers of many NP-complete problems. We also show that the inverse complexity of a verifier for a language L cannot be predicted solely from the complexity of L, but rather, is highly dependent upon the choice of verifier used to accept L. In this context, a verifier with a Σ_2^p-complete inverse problem is exhibited, giving a new and natural example of a Σ_2^p-complete problem.

*Department of Computer Science, Cornell University, Ithaca NY 14853. Email: hubes@cs.cornell.edu.

1 Introduction

By now, the complexity class NP has become one of the most pervasive and successful notions in computer science. Indeed, the intimately connected notion of NP-completeness is described as "computer science's favorite paradigm, fad, punching bag, buzzword, alibi, and intellectual export" in a lively discussion on the influence and nature of NP-completeness [13]. There are many equivalent characterizations of the class NP. According to one such characterization, a language is in NP if there is an efficient (polynomial-time) *verifier* with the following properties: for every member of the language, there exists a polynomially-sized proof causing the verifier to accept (upon being given the member along with the proof); and, for every non-member, there is no proof causing the verifier to accept. Thus, with a verifier for a language in hand, deciding membership in the language amounts to deciding whether or not a given potential member has a proof or not.

Inverse problems. Relative to a particular verifier, every member x of the language induces a *set of proofs*, namely, the set of proofs causing the verifier to accept x as a member. Call a set of proofs arising in this way *induced*. This paper studies a simply stated and natural question: what is the complexity of deciding, given a set Π of proofs, whether or not Π is induced? We call this decision problem the *inverse problem* for the verifier.

As a first example, consider the natural verifier for 3-SAT, which accepts a 3-SAT formula and an assignment to the variables of the formula when the assignment satisfies the formula; call the inverse problem for this verifier INVERSE 3-SAT. INVERSE 3-SAT is then the question of deciding, given a set Π of assignments, if there is a 3-SAT formula with *exactly* Π as its set of satisfying assignments. It is fairly straightforward to show that INVERSE 3-SAT is in coNP: for any set of assignments Π, we can efficiently compute the "candidate formula" F containing all 3-clauses satisfied by all assignments in Π. It can be verified that if *any* 3-SAT formula has Π as set of satisfying assignments, then the candidate formula is such a formula. Intuitively, this is because we placed as many clauses as possible in the candidate formula F, so any 3-SAT formula including Π in its induced set can only be less constrained – that is, admit more satisfying assignments – than F. Thus, INVERSE 3-SAT is in coNP: to show that a set Π is not in INVERSE 3-SAT, it suffices to give an assignment outside of Π satisfied by the candidate formula (of Π).

In previous work, Kavvadias and Sideri [9] considered the *inverse satisfiability problem*, which includes as a specific case the INVERSE 3-SAT problem. (It is from them that we adopt the term *inverse problem*.) They examined the class of generalized boolean satisfiability problems on which Schaefer's dichotomy theorem [14] was proved – a class which includes Horn SAT and 2-SAT, two well-known tractable subclasses of SAT; as well as not-all-equal SAT and one-in-three SAT, two intractable variants of SAT. For any satisfiability problem from this generalized class, the inverse problem is always in coNP by an argument similar to that given above. The intriguing result they obtained is that the inverse satisfiability problem is intractable (coNP-complete) if and only if the original satisfiability problem is intractable (NP-complete)! Thus, it follows from their result that INVERSE 3-SAT is coNP-complete because 3-SAT is NP-complete, as well as that INVERSE 2-SAT is in P because 2-SAT is in P.

To give another example of an inverse problem, we begin with the language K-CLIQUE consisting of pairs $\langle G, k\rangle$ such that G is an undirected graph with a clique of size k. Consider the verifier for this language which accepts H as a proof for $\langle G, k\rangle$ if H is a size k subset of the vertex set of G that forms a clique in G. Call the inverse problem for this verifier INVERSE K-CLIQUE. Then, a well-formed input to the INVERSE K-CLIQUE problem is a set $\Pi = \{H_1, \ldots, H_m\}$, where each H_i is a set of size k; and the problem is to decide whether or not there is a graph G such that $\{H_1, \ldots, H_m\}$ is exactly the set of k-cliques in G. In analogy to the "candidate formula" for an instance of INVERSE 3-SAT, we can compute from an instance Π of INVERSE K-CLIQUE a "candidate graph" G with the property that if *any* graph has Π as its set of k-cliques, then G is such a graph: our candidate graph has vertex set $\cup_{i=1}^{m} H_i$, with an edge between two

vertices if the two vertices are both contained in H_i, for some i. Thus, INVERSE K-CLIQUE is in coNP: to show that Π is not in INVERSE K-CLIQUE, it suffices to demonstrate that the candidate graph for Π has a k-clique not in Π. It will be demonstrated in this paper that INVERSE K-CLIQUE is also coNP-complete.

These examples might seem to be evidence for the conjecture that the inverse problem for verifiers accepting NP-complete languages will always be coNP-complete. However, a fairly succinct argument demonstrates that this conjecture fails. Suppose we take the verifier for K-CLIQUE and define a new verifier which accepts only "padded" proofs of the form $\langle s, H\rangle$ where H is a clique and s is a string with, say, length equal to $|H|$. Then, the inverse problem becomes "easier": for every clique H accepted by the first verifier, exponentially many proofs of the form $\langle s, H\rangle$ are accepted by the second verifier. If padding of sufficient length is applied to the proofs of the verifier for K-CLIQUE, we obtain a verifier for K-CLIQUE with inverse complexity in P. Of course, such a padded verifier may seem to be quite artificial. For many NP-complete languages L, one may feel that there is a "natural" verifier for L (such as the verifiers described in the above examples), and we will show that the inverse problem for many such verifiers is coNP-complete. Nonetheless, we emphasize that for each language L in NP, there is more than one verifier[1] for L, and that the inverse problem is defined relative to a *verifier* for an NP language, and not just an NP language by itself.

Related work and contributions. Our study of inverse problems contributes to a line of research which aims to understand the properties of the natural verifiers for NP-complete languages. This research program has its origins in the work done on the counting class ♯P by Simon and Valiant [15, 17]; one important outcome of this work was the observation that many natural verifiers are related by *parsimonious reductions*, a stronger notion of many-one reduction that preserves the number of proofs. A notion of reduction even stronger than parsimonious reduction was studied by Lynch and Lipton [12]; they required that when a language member x many-one reduces to another language y, there be an efficiently computable function mapping proofs for x to proofs for y. Fischer, Hemaspaandra, and Torenvliet [7] went one step further to study what they called *witness-isomorphic reductions*, which require that there be efficiently computable isomorphisms between the proof sets of language members reduced to one another. Other work along this line includes [1], in which Agrawal and Biswas gave sufficient conditions for a verifier to accept a NP-complete language. These sufficient conditions allow one to derive many of the known NP-completeness results in a uniform manner. Our work complements these previous results by identifying an entirely new feature shared by many of the natural verifiers for NP-complete languages – namely, coNP-complete "inverse complexity."

This paper is also a contribution to the study of *structure identification* [5], where the goal is to decide, given a set of data points, whether or not the set is meaningfully structured in a way that admits a description of desirable type. More broadly, we believe that the study of inverse complexity sheds new light on the multitude of *natural properties of classical combinatorial structures* yielding NP-complete decision problems.

We note that inverse complexity has consequences for another task involving computationally difficult problems – the task of generating multiple solutions to a search problem, which is of both theoretical and practical interest (see, for example, [8, 11, 10, 6]). Here, a relevant problem is to decide whether or not all proofs of a language member have (at some point in time) been generated – precisely, to decide, given a language member x and a non-empty set of proofs Π of x, whether or not Π is the full set of proofs induced by x. Hardness of the inverse problem implies hardness of this decision problem, under a mild assumption.

Results and paper organization. The notation and conventions used throughout the paper are fixed in Section 2. In this section, we also define the notion of a *candidate function* for a verifier, examples of which include the "candidate formula" and "candidate graph" procedures given above.

[1] In fact, there are infinitely many.

In Section 3, we list a number of verifiers for well-known NP languages (all having candidate functions), and prove that INVERSE K-CLIQUE is coNP-complete. We also formulate a sufficient condition for an inverse problem to be in P, and give an example application of the condition by showing that INVERSE BIPARTITE MATCHING is in P.

We then develop a notion of reduction (Section 4) which allows us to compare the complexity of inverse problems with relative ease. In particular, we define what we call a π-*reduction* between verifiers with candidate functions, and show that the existence of a π-*function* implies the existence of a π-reduction. By giving a sequence of π-functions, we leverage the hardness of INVERSE K-CLIQUE and INVERSE 3-SAT to show the coNP-completeness of many other inverse problems on our list.

We have already hinted that the "inverse complexity" of a language depends strongly on the verifier used to accept the language. In Section 5, we formally demonstrate and study this dependence. We give a series of results that allow one to manipulate inverse complexity. We also show that there exists a verifier with a Σ_2^p-complete inverse problem, giving a new and natural example of a Σ_2^p-complete problem. The section culminates in a theorem which shows that for all natural NP-complete languages L and *any* language A in NP, there is a verifier for L with inverse complexity A.

Due to space limitations, proofs are given in the appendix.

2 Preliminaries

We assume familiarity with basic notions of complexity theory, such as the complexity classes P, NP, coNP, Σ_2^p, and Π_2^p, described for instance in [3]. We write $A \leq_m^p B$ if there is a many-one polynomial time reduction from A to B, and $A \equiv_m^p B$ if $A \leq_m^p B$ and $B \leq_m^p A$. The power set of a set S is denoted by $\mathcal{P}(S)$. Throughout this paper, Σ denotes a fixed finite alphabet. When $S \subseteq \Sigma^*$ is a set of strings, $\|S\|$ denotes $\|S\| = \sum_{x \in S} |x|$, that is, the sum of the lengths of the strings in S. (We assume that such a set $S \subseteq \Sigma^*$ is represented by a string of length linear in $\|S\|$.)

Definition 2.1 *A relation $R \subseteq \Sigma^* \times \Sigma^*$ is a* verifier *if membership in R is decidable in polynomial time and there is a polynomial p such that (for all $x, \pi \in \Sigma^*$) $R(x, \pi)$ implies that $|\pi| \leq p(|x|)$. When R is a verifier, we let $R(x)$ denote the set of proofs for $x \in \Sigma^*$, that is, $\{\pi \in \Sigma^* : R(x, \pi)\}$; and, we let $L(R)$ denote the set $\{x \in \Sigma^* : \exists \pi \in \Sigma^*$ such that $R(x, \pi)\}$, which we call the language associated with or accepted by R.*

We now define the inverse problem corresponding to a verifier, which is the problem of focal interest in this paper.

Definition 2.2 *Suppose that R is a verifier. Let R^{-1} denote the language*

$$\{\Pi \subseteq \Sigma^* : \exists x \in L(R) \text{ such that } R(x) = \Pi\}.$$

We call this language the inverse problem for R.

When R is a verifier, we will refer to the members of the language $L(R)$ as theorems; when x is a theorem, the elements of $R(x)$ are said to be its proofs. Using this terminology, the inverse problem is the question of deciding, given a set of proofs Π, whether or not there is a theorem x with exactly Π as its set of proofs. A natural initial consideration is whether or not there is an upper bound on the complexity of R^{-1}. We can obtain such an upper bound when we place a restriction which we call fairness on R. All verifiers considered in this paper obey this restriction.

Definition 2.3 *A verifier R is a* fair verifier *if there is a polynomial q such that for all $x \in L(R)$, there exists $x' \in L(R)$ where $|x'| \leq q(\|R(x)\|)$ and $R(x') = R(x)$.*

Put differently, R is fair if for all sets of proofs Π, when there exists a theorem x with *exactly* Π as its set of proofs (that is, $\Pi = R(x)$), then there exists such an x with length bounded above by a polynomial in $\|\Pi\|$. Suppose R is a fair verifier. In deciding whether or not a set of proofs Π is in R^{-1}, we need only consider as potential theorems the x which are bounded above in length by a polynomial in $\|\Pi\|$, the length of the representation of Π. Moreover, by the definition of a verifier, a proof of x (that is, an element of $R(x)$) has length polynomial in $|x|$. These two facts lead to the following observation.

Observation 2.4 *If R is a fair verifier, then R^{-1} is in Σ_2^p.*

We note that without the restriction of fairness, the complexity of R^{-1} can be non-recursive – in fact, as high as the halting problem for general Turing machines.[2]

Definition 2.5 *A non-empty set $\Pi \subseteq \Sigma^*$ is* well-formed *(relative to an NP verifier R) if there exists $x \in \Sigma^*$ such that $\Pi \subseteq R(x)$.*

The inverse problem is only "interesting" on well-formed sets, since a set that is not well-formed cannot be in R^{-1}. Checking well-formedness of a set can be done in polynomial time for all verifiers considered in this paper.

We now introduce the notion of a *candidate function* for a verifier R. This notion can be thought of as a generalization of the "candidate formula" and "candidate graph" procedures described in the introduction. A candidate function efficiently maps a set of proofs Π to a "candidate theorem" having the property that if any theorem has Π as its set of proofs, then the candidate theorem is such a theorem. In addition, the candidate theorem for Π is optimistic in the sense that all proofs in Π are proofs of it.

Definition 2.6 *Let R be a verifier. A polynomial time computable function $C : \mathcal{P}(\Sigma^*) \to \Sigma^*$ is a* candidate function *for R if the following two conditions hold.*

1. *For all well-formed $\Pi \subseteq \Sigma^*$, all proofs in Π are proofs for $C(\Pi)$: $\Pi \subseteq R(C(\Pi))$.*
2. *For all well-formed $\Pi \subseteq \Sigma^*$, if there exists an $x \in L(R)$ such that $\Pi = R(x)$, then $\Pi = R(C(\Pi))$.*

Note that when $C(\Pi)$ has exactly Π as its set of proofs, there is no requirement that $C(\Pi)$ be the unique theorem with Π as its set of proofs. Indeed, there will generally not be such a unique theorem.

When a verifier has a candidate function, we can give an even better upper bound on the complexity of its inverse problem than that of Observation 2.4. (Notice that every verifier with a candidate function is fair.)

Observation 2.7 *If R is a verifier with a candidate function, then R^{-1} is in coNP.*

This is because we no longer have to search over all theorems x to find a match for Π, but rather can simply compute the candidate theorem in polynomial time, and then check whether or not its set of proofs is Π.

Definition 2.8 *Let R be a verifier. Define the* exhaustive proof problem *for R to be the language*

$$\text{EXHAUSTIVE}(R) = \{\langle x, \Pi\rangle \in \Sigma^* \times \mathcal{P}(\Sigma^*) : R(x) = \Pi, \Pi \neq \emptyset\}$$

[2]For any recursively enumerable language L (such as the halting problem), there exists a verifier R such that $R^{-1} = L$. To see this, let M be a Turing machine accepting L. Define the verifier $R(x, \pi)$ to be true when x is of the form $\langle 0^n, \pi\rangle$ and M accepts when simulated for time n on input π: then, $R^{-1} = L$. Since R^{-1} is always recursively enumerable, we obtain that a language is recursively enumerable if and only if it is equal to R^{-1} for some verifier R.

In other words, the exhaustive proof problem is to determine, given an x and a *nonempty* set of proofs Π for x, whether or not Π contains all proofs of x (relative to R). This is the case when π is a proof of x if and only if π is in Π (for all π). To verify this one needs only examine π of length polynomial in x (by definition of a verifier), leading to the following observation.

Observation 2.9 *If R is a verifier, then* EXHAUSTIVE(R) *is in coNP.*

Notice that without the restriction that the given set of proofs must be nonempty, the exhaustive proof problem for the verifier R of any NP-complete language L would trivially be coNP-complete. In this case, the reduction mapping a string x to the pair $\langle x, \emptyset \rangle$ would reduce co-L to the exhaustive proof problem for R. Although this trivial reduction fails for the given definition of the exhaustive proof problem, we can nonetheless establish hardness of this problem by using the hardness of the corresponding inverse problem.

Lemma 2.10 *If R is a verifier with candidate function, then $R^{-1} \leq_m^p$* EXHAUSTIVE(R).

3 Inverse Problems

In this section, we give a list of inverse problems considered in the paper, along with some initial complexity classifications. Before giving the list, we specify some conventions and notation.

The set $\{1, \ldots, n\}$ is denoted by $[n]$. When G is a graph, $V(G)$ and $E(G)$ denote its vertex and edge sets, respectively. Unless specified otherwise, graphs in this paper are assumed to be undirected, and for undirected graphs G, all elements of $E(G)$ are assumed to be size two subsets of $V(G)$. If $g : D \to R$ is a function and $S \subseteq D$, $g_{|S}$ denotes the restriction of g to S. All sets in the descriptions below are assumed to be finite.

3.1 Examples of Inverse Problems

We now give a list of example inverse problems. All are specified by fair verifiers with candidate functions. For each verifier, we give the well-formed sets, as well as a candidate function. We leave it to the reader to verify, for each verifier R, that the well-formed sets of R are as described, as well as that the given function C is indeed a candidate function for R. We define the candidate functions only on the well-formed sets.

1. INVERSE 3-SAT

   Verifier: $R(x, \pi)$ if x is a set of 3-clauses over the variable set $\{v_1, \ldots, v_n\}$, and $\pi : \{v_1, \ldots, v_n\} \to \{T, F\}$ satisfies x. (A 3-clause over the variable set V is a set of literals with variables from V where no variable is repeated.)

   Well-formed sets: Π such that there exists $n \geq 1$ where all elements of Π are mappings $\{v_1, \ldots, v_n\} \to \{T, F\}$.

   Candidate function: $C(\Pi) = \{c : c \text{ is a 3-clause over } \{v_1, \ldots, v_n\} \text{ satisfied by all } \pi \in \Pi\}$.

2. INVERSE CIRCUIT-SAT

   Verifier: $R(x, \pi)$ if x is a boolean circuit with inputs from the variable set $\{v_1, \ldots, v_n\}$, and the assignment $\pi : \{v_1, \ldots, v_n\} \to \{T, F\}$ causes the circuit to evaluate to true.

   Well-formed sets: Π such that there exists $n \geq 1$ where all elements of Π are mappings $\{v_1, \ldots, v_n\} \to \{T, F\}$.

   Candidate function: $C(\Pi) = \bigvee_{\pi \in \Pi} r_\pi$, where $r_\pi = \bigwedge_{i=1}^{n} l_i^\pi$ and $l_i^\pi = v_i$ if $\pi(v_i) = T$; $\neg v_i$ if $\pi(v_i) = F$.

3. INVERSE K-CLIQUE

   Verifier: $R(x, \pi)$ if $x = \langle G, k \rangle$ where G is a graph, $1 \leq k \leq |V(G)|$, and $\pi = H$ where $H \subseteq V(G)$ is a clique of size k in G.

   Well-formed sets: Π such that there exists $k \geq 1$ where all elements $\pi \in \Pi$ are sets of size k.

   Candidate function: $C(\Pi) = \langle G, k \rangle$ where G has vertex set $\bigcup_{H \in \Pi} H$ and edge set $\bigcup_{H \in \Pi} K(H)$, where $K(H)$ is the set of all size two subsets of H.

4. INVERSE BIPARTITE MATCHING

   Verifier: $R(x, \pi)$ if $x = \langle A, E \rangle$ where $E \subseteq A \times A$, $\pi \subseteq E$, $|\pi| = |A|$, and no two elements of π agree in any coordinate.

   Well-formed sets: Π such that there exists a set A where for all $\pi \in \Pi$, $\pi \subseteq A \times A$, $|\pi| = |A|$, and no two elements of π agree in any coordinate.

   Candidate function: $C(\Pi) = \langle A, E \rangle$ where $E = \bigcup_{\pi \in \Pi} \pi$.

5. INVERSE VERTEX COVER

   Verifier: $R(x, \pi)$ if $x = \langle G', l \rangle$ where G' is a graph; $1 \leq l \leq |V(G)|$; and $\pi = \langle l, R \rangle$, where $R \subseteq V(G')$ is of size l and is a vertex cover of G': every edge of G' contains a vertex in R.

   Well-formed sets: Π such that there exists $l \geq 1$ where all elements of Π are of the form $\langle l, R \rangle$, with R a set of size l.

   Candidate function: $C(\Pi) = \langle G', l \rangle$, where G' has vertex set $V(G') = \bigcup_{\langle l, R \rangle \in \Pi} R$ and edge set $\{\{u, v\} \subseteq V(G') : \forall \langle l, R \rangle \in \Pi, R \cap \{u, v\} \neq \emptyset\}$.

6. INVERSE EXACT COVER

   Verifier: $R(x, \pi)$ if $x = \langle B, \mathcal{U} \rangle$ where for all $U \in \mathcal{U}$, $U \subseteq B$, and $\pi \subseteq \mathcal{U}$ forms a disjoint union of B.

   Well-formed sets: Π such that there exists a set B where for all $\pi \in \Pi$, $\pi \subseteq \mathcal{U}$ forms a disjoint union of B.

   Candidate function: $C(\Pi) = \langle B, \mathcal{U} \rangle$ where $\mathcal{U} = \{U : \text{ exists } \pi \in \Pi \text{ such that } U \in \pi\}$.

7. INVERSE KNAPSACK

   Verifier: $R(x, \pi)$ if $x = \langle A, b \rangle$ where A is a multiset over $\mathbb{Z}$, $b \in \mathbb{Z}$, $\pi \subseteq A$, and $\sum_{a \in \pi} a = b$.

   Well-formed sets: Π such that there exists $b \in \mathbb{Z}$ where $\sum_{a \in \pi} a = b$ and π is a multiset over $\mathbb{Z}$, for all $\pi \in \Pi$.

   Candidate function: $C(\Pi) = \langle A, b \rangle$ where A is a multiset with elements from $\bigcup_{\pi \in \Pi} \pi$, where the frequency of an element is equal to the maximum number of times it appears in any $\pi \in \Pi$.

8. INVERSE STEINER TREE

   Verifier: $R(x, \pi)$ if $x = \langle G, k \rangle$ where G is a graph with edge weights from $\mathbb{Z}$, $k \in \mathbb{N}$ and $\pi = \langle k, T \rangle$ where T is a subtree of G (with the same weights on its edges as in G) having total edge weight $\leq k$.

   Well-formed sets: Π such that there exists $k \in \mathbb{N}$ where all elements of Π are of the form $\langle k, T \rangle$ for some tree T with edge weights; and, for all $\langle k, T_1 \rangle, \langle k, T_2 \rangle \in \Pi$, any edge appearing in both T_1 and T_2 has the same weight in each tree.

   Candidate function: $C(\Pi) = \langle G, k \rangle$ where $G = \bigcup_{\langle k, T \rangle \in \Pi} T$.

9. INVERSE PARTITION

   Verifier: $R(x, \pi)$ if x is a multiset over $\mathbb{Z}$ and $\pi = \{D, E\}$ where D and E are multisets forming a disjoint union of x.

   Well-formed sets: Π such that for all $\pi_1, \pi_2 \in \Pi$, $\pi_1 = \{D_1, E_1\}$ and $\pi_2 = \{D_2, E_2\}$ are sets of size 2, each containing two multisets over $\mathbb{Z}$ where $D_1 \cup E_1 = D_2 \cup E_2$, and $\sum_{d \in D_1} d = \sum_{e \in E_1} e = \sum_{d \in D_2} d = \sum_{e \in E_2} e$.

   Candidate function: $C(\Pi) = D \cup E$ for any $\pi = \{D, E\}$.

We denote the exhaustive proof problem corresponding to these verifiers by EXHAUSTIVE 3-SAT, EXHAUSTIVE CIRCUIT-SAT, etc.

3.2 Initial Complexity Results

We now establish the complexity of a few of the inverse problems from our list. The hardness results given here will serve as starting points for deriving further hardness results. First, we note a theorem established in previous work by Kavvadias and Sideri.

Theorem 3.1 *[9]* INVERSE 3-SAT *is coNP-complete.*

Next, we establish that the inverse clique problem is coNP-hard. This is done by reducing from the clique problem itself. The idea is to take a graph G (an instance of the clique problem) and "expose" its edges by creating a clique for each edge of G. The resulting candidate graph corresponding to the set of cliques Π has, for each clique in the original graph G, a clique not in Π.

Theorem 3.2 INVERSE K-CLIQUE *is coNP-complete.*

It generally seems to be the case that the inverse problem for a natural verifier accepting an NP-complete language is coNP-complete: in the next section, we give many examples where this is true. However, there is at least one exception to this apparent rule of thumb.

Observation 3.3 INVERSE CIRCUIT-SAT *is in P.*

This is because *every* well-formed set Π has a circuit with exactly Π as the satisfying assignments (such a circuit is given by $C(\Pi)$ where C is the candidate function for the CIRCUIT-SAT verifier). One might contrast this observation with Theorem 3.1; intuitively, the difference is that the class of boolean circuits is capable of expressing all possible sets of satisfying assignments, whereas the sets of satisfying assignments expressible by 3-SAT formulas is restricted.

We can place the complexity of other inverse problems inside P by making use of a link between inverse problems and the notion of *output polynomial time*. (This notion has been discussed in previous work, for example [8].)

Definition 3.4 *Say that a verifier R has an* output polynomial time algorithm *if given x, the set $R(x)$ can be computed in time polynomial in $|x| + \|R(x)\|$.*

When R has an output polynomial time algorithm, both the exhaustive proof problem and inverse problem for R become easy. The following lemma is implicit in [9].

Lemma 3.5 *Suppose that R is a verifier with an output polynomial time algorithm. Then,* EXHAUSTIVE(R) *is in P. If in addition R has a candidate function, then R^{-1} is in P.*

This lemma allows us to establish the tractability of INVERSE BIPARTITE MATCHING, as there is an output polynomial time algorithm for BIPARTITE MATCHING.[3]

Theorem 3.6 INVERSE BIPARTITE MATCHING *is in P.*

4 π-reductions

We now develop a notion of reduction between pairs of verifiers with candidate functions called π-reduction, which allows us to compare the complexity of the respective inverse problems. When there is a π-reduction from one verifier to another, the inverse problem of the first reduces in polynomial time to the inverse problem of the second. In addition, we show that the existence of a particular type of function, which we call a π-function, is sufficient for a π-reduction to exist; giving such π-functions will be our main tool for showing hardness of inverse problems. Throughout this section, we assume that R and S are verifiers with candidate functions C_R and C_S, respectively.

Definition 4.1 *A polynomial time computable function $g : \mathcal{P}(\Sigma^*) \to \mathcal{P}(\Sigma^*)$ is a* π-reduction *from R to S if the following three conditions hold.*

1. *For all well-formed $\Pi \subseteq \Sigma^*$, g preserves the number of proofs: $|\Pi| = |g(\Pi)|$.*
2. *For all well-formed $\Pi \subseteq \Sigma^*$, the candidate theorems for Π and $g(\Pi)$ have the same number of proofs: $|R(C_R(\Pi))| = |S(C_S(g(\Pi)))|$.*
3. *If $\Pi \subseteq \Sigma^*$ is not well-formed, then $g(\Pi)$ is not well-formed.*

Lemma 4.2 *If g is a π-reduction from R to S, then $R^{-1} \leq_m^p S^{-1}$ via g.*

Definition 4.3 *Suppose $f : \Sigma^* \times \Sigma^* \to \Sigma^*$ is a partial function computable in polynomial time, and let $f_x : \Sigma^* \to \Sigma^*$ denote the function defined by $f_x(\pi) = f(x, \pi)$. Let us say that f is a* π-function *from R to S (relative to C_R and C_S) if for all well-formed $\Pi \subseteq \Sigma^*$, when x and y are set as $x = C_R(\Pi)$ and $y = C_S(f_x(\Pi))$, f_x is a bijection between $R(x)$ and $S(y)$.*

Lemma 4.4 *Suppose that there exists a π-function from R to S (relative to C_R and C_S). Then there exists a π-reduction from R to S, and $R^{-1} \leq_m^p S^{-1}$.*

Having developed a notion of reduction between inverse problems, we can classify the complexity of many natural inverse problems with relative ease.

Theorem 4.5 INVERSE EXACT COVER, INVERSE VERTEX COVER, INVERSE KNAPSACK, INVERSE STEINER TREE, *and* INVERSE PARTITION *are all coNP-complete.*

All of these problems are in coNP since the corresponding verifiers have candidate functions. We establish that they are all coNP-hard by starting with the fact that INVERSE 3-SAT and INVERSE K-CLIQUE are coNP-hard (Theorems 3.1 and 3.2), and then giving a sequence of π-functions. We note that parsimonious reductions h between R and S such that there is an efficient mapping from proofs of $x \in L(R)$ to proofs of $h(x)$ are often (but not always) of great help in developing π-functions.

[3] Although it would be tempting to conjecture that there is a correspondence between the complexity of the counting problem corresponding to a verifier R (given x, compute $|R(x)|$) and the inverse problem R^{-1}, a general correspondence seems unlikely in light of Theorem 3.1, Theorem 3.2, Observation 3.3 and Theorem 3.6. The counting versions of 3-SAT, K-CLIQUE, CIRCUIT-SAT, and BIPARTITE MATCHING are all $\sharp$P-complete [17]. Moreover, the decision problems for 3-SAT and CIRCUIT-SAT are both NP-complete, so even attempting to predict the inverse complexity based on *both* the counting complexity and decision complexity seems difficult.

Corollary 4.6 EXHAUSTIVE 3-SAT, EXHAUSTIVE CIRCUIT-SAT, EXHAUSTIVE K-CLIQUE, EXHAUSTIVE EXACT COVER, EXHAUSTIVE VERTEX COVER, EXHAUSTIVE KNAPSACK, EXHAUSTIVE STEINER TREE, *and* EXHAUSTIVE PARTITION *are all coNP-complete.*[4]

5 How hard can inverse problems be?

In this section, we demonstrate formally that the inverse complexity R^{-1} of a verifier R cannot be predicted solely from $L(R)$. That is, for a fixed language L, the "inverse complexity" of a verifier accepting L can vary dramatically depending on the choice of verifier. This justifies our analysis of *particular* verifiers performed in the previous section.

We give a series of results allowing one to manipulate the inverse complexity. We also show that there is a fair verifier with a Σ_2^p-complete inverse problem, giving a new and natural example of a Σ_2^p-complete problem. We then use these results to show that for a "natural" NP-complete language A, for any B that is in NP or Σ_2^p-complete, there is a verifier accepting A with inverse complexity B.

Define the join of two languages A and B to be $A \otimes B = \{0x : x \in A\} \cup \{1x : x \in B\}$. Call a language L *nontrivial* if $L \neq \emptyset$ and $L \neq \Sigma^*$. Call a verifier R *nontrivial* if both $L(R)$ and R^{-1} are nontrivial. Throughout this section, we assume that all languages and verifiers discussed are nontrivial.

Our first lemma shows that the inverse complexity can be dropped down to P while holding still the language accepted.

Lemma 5.1 *For all $A \in NP$, there exists a fair verifier R such that $L(R) = A$ and $R^{-1} \in P$.*

The next lemma shows that inverse problems of all varied complexity levels (in NP) exist.

Lemma 5.2 *For all $B \in NP$, there exists a fair verifier S such that $S^{-1} \equiv_m^p B$.*

One might conjecture that this lemma is true for all $B \in \Sigma_2^p$. A first result along these lines is the following theorem, establishing that the lemma is true when B is Σ_2^p-complete: that is, the maximum complexity (Observation 2.4) of the inverse problem is attainable.

Theorem 5.3 *There exists a fair verifier R such that R^{-1} is Σ_2^p-complete.*

Our next theorem allows one to take two verifiers R, S and cook up a new one accepting essentially the same language as R, but with inverse complexity equal to the join of R^{-1} and S^{-1}.

Theorem 5.4 *For any fair verifiers R, S, there exists a fair verifier T such that $L(T) \equiv_m^p L(R)$ and $T^{-1} \equiv_m^p R^{-1} \otimes S^{-1}$.*

Collecting together the above results, we have the following corollary.

Corollary 5.5 *For all $A \in NP$, if $B \in NP$ or B is Σ_2^p-complete, then there exists a fair verifier T such that $L(T) \equiv_m^p A$ and $T^{-1} \equiv_m^p B$.*

One may conjecture that this is true for all $B \in \Sigma_2^p$; this would follow from an extension of Lemma 5.2 to all $B \in \Sigma_2^p$.

[4]For some (but not all) of the listed problems, we are aware of elementary proofs of coNP-completeness where the reduction is from the complement of the original language. For instance, co-K-CLIQUE reduces to EXHAUSTIVE K-CLIQUE by mapping $\langle G, k\rangle$ to $\langle\langle G \cup C, k\rangle, V(C)\rangle$, where C is a clique of size k with vertex set disjoint from that of G.

Conjecture 5.6 *For all infinite nontrivial $A \in NP$ and $B \in \Sigma_2^p$, there exists a fair verifier T such that $L(T) \equiv_m^p A$ and $T^{-1} \equiv_m^p B$.*

Indeed, one may conjecture something stronger: for any $A \in NP$, there exists such a verifier accepting A itself, and not just a language many-one equivalent to A.

Conjecture 5.7 *For all infinite nontrivial $A \in NP$ and $B \in \Sigma_2^p$, there exists a fair verifier T such that $L(T) = A$ and $T^{-1} \equiv_m^p B$.*

Berman and Hartmanis observed in 1977 that isomorphisms exist between a number of NP-complete languages [4], and conjectured that all NP-complete languages are isomorphic. Define a *natural* NP-complete language to be any language isomorphic to one of the languages considered in their paper. In the direction of Conjecture 5.7, we have the following theorem.

Theorem 5.8 *For all natural NP-complete languages A, if $B \in NP$ or B is Σ_2^p-complete, then there exists a fair verifier T such that $L(T) = A$ and $T^{-1} \equiv_m^p B$.*

6 Conclusions

In this paper, we formalized the inverse problem for an arbitrary NP verifier, which includes as particular cases the inverse satisfiability problems studied by Kavvadias and Sideri [9]. We developed a notion of reduction for inverse problems and used it to classify the inverse complexity of many natural verifiers. We also formally demonstrated that the "inverse complexity" of a verifier for language L cannot be predicted from L, but rather, exhibits a strong dependence on the choice of verifier used to accept L.

There are some interesting directions for future work. Can the intuitive claim that natural verifiers for NP-complete problems tend to have coNP-complete inverse problems, be formalized and proved? (Obviously, a formalized notion of "natural verifier" having the property that all natural verifiers have coNP-complete inverse problems would have to exclude the verifier for CIRCUIT-SAT given in this paper, by Observation 3.3.) Along these lines, it would be of interest to classify the inverse complexity of the natural verifiers for GRAPH 3-COLORABILITY and HAMILTONIAN CIRCUIT. In addition, one could study inverse problems for other complexity classes definable by nondeterministic machines, such as NL and Σ_2^p.

Acknowledgements. The author wishes to thank Andrew Blumberg, Carla Gomes, Dexter Kozen, Riccardo Pucella, Tim Roughgarden, and Bart Selman for useful discussions and comments. The author was supported by a NSF Graduate Research Fellowship.

References

[1] M. Agrawal and S. Biswas. Universal relations. In *Proc. 7th Structure in Complexity Theory Conference*, pages 207–220, 1992.

[2] J. L. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity II*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, Berlin, 1990.

[3] J. L. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity I*. Texts in Theoretical Computer Science – An EATCS series. Springer-Verlag, Berlin, 2nd edition, 1995.

[4] L. Berman and J. Hartmanis. On isomorphism and density of NP and other complete sets. *SIAM Journal on Computing*, 6:305–322, 1977.

[5] R. Dechter and J. Pearl. Structure identification in relational data. *Artificial Intelligence*, 58:237–270, 1992.

[6] Thomas Eiter and Kazuhisa Makino. On computing all abductive explanations. In *AAAI/IAAI*, pages 62–67, 2002.

[7] S. Fischer, L. Hemaspaandra, and L. Torenvliet. Witness-isomorphic reductions and local search. *Complexity, Logic and Recursion Theory, Lecture Notes in Pure and Applied Mathematics, A. Sorbi (ed.)*, pages 207–223, 1997.

[8] D. S. Johnson, C. H. Papadimitriou, and M. Yannakakis. On generating all maximal independent sets. *Information Processing Letters*, 27:119–123, 1988.

[9] D. Kavvadias and M. Sideri. The inverse satisfiability problem. *SIAM Journal on Computing*, 28(1):152–163, 1998.

[10] Dimitris J. Kavvadias, Martha Sideri, and Elias C. Stavropoulos. Generating all maximal models of a Boolean expression. *Information Processing Letters*, 74(3–4):157–162, 2000.

[11] A. Kwan, E. P. K. Tsang, and J. E. Borrett. Phase transition in finding multiple solutions in constraint satisfaction problems. In *Workshop on Studying and Solving Really Hard Problems, First International Conference on Principles and Practice of Constraint Programming*, pages 119–126, 1995.

[12] N. Lynch and R. Lipton. On structure preserving reductions. *SIAM J. Comp.*, 7(2):119–125, 1978.

[13] C. Papadimitriou. NP-completeness: A retrospective. In *Proceedings of the 24th International Colloquium on Automata, Languages and Programming*, volume 1256 of *Lecture Notes in Computer Science*, pages 2–6. Springer, 1997.

[14] T. J. Schaefer. The complexity of satisfiability problems. In *Proc. 10th Annual ACM Symposium on Theory of Computing*, pages 216–226, 1978.

[15] J. Simon. On some central problems in computational complexity. Technical report, Cornell University, 1975.

[16] L. J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3:1–22, 1977.

[17] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.

A Proof of Lemma 2.10

Proof. By the reduction $f(\Pi) = \langle C_R(\Pi), \Pi \rangle$ (defined for well-formed Π). □

B Proofs of Section 3

B.1 Proof of Theorem 3.2

Proof. We reduce from co-K-CLIQUE, where K-CLIQUE is the language associated with the verifier R given in the description of INVERSE K-CLIQUE. Map an instance $\langle G, k \rangle$ of co-K-CLIQUE to an instance Π of INVERSE K-CLIQUE as follows. For each $e \in E(G)$, define V_e to be a vertex set of size $k-2$ such that all of the V_e and $V(G)$ are disjoint, that is, $|V(G) \cup \bigcup_{e \in E(G)} V_e| = |V(G)| + (k-2)|E(G)|$. Define $\Pi = \{e \cup V_e : e \in E(G)\}$.

Suppose $\langle G, k \rangle$ is not in co-K-CLIQUE. Then G contains a clique H of size k, and $H \in R(C_R(\Pi))$. However, $H \notin \Pi$, implying that $\Pi \neq R(C_R(\Pi))$, from which it follows that Π is not in INVERSE K-CLIQUE.

Suppose $\langle G, k \rangle$ is in co-K-CLIQUE. Then G has no clique of size k, and one can verify that the only cliques in the graph of $C(\Pi)$ are the cliques in Π, that is, $\Pi = R(C_R(\Pi))$. □

B.2 Proof of Lemma 3.5

Proof. By Lemma 2.10, it suffices to prove the first claim. Let A denote an output polynomial time algorithm for R, and let q be a polynomial upper bound for the running time of A. Given $\langle x, \Pi \rangle$, we decide membership in EXHAUSTIVE(R) in polynomial time as follows. Simulate $A(x)$, and terminate the simulation if it does not halt itself within time $q(|x| + \|\Pi\|)$. If $A(x)$ halts before the given time bound and outputs Π, then accept; otherwise, reject. □

B.3 Proof of Theorem 3.6

Proof. We give an output polynomial time algorithm which generates all matchings of a given bipartite graph. The algorithm is as follows: Pick a vertex v. For all edges e containing v, check to see if the graph $G - \{e\}$ has a bipartite matching; if so, call the algorithm recursively on $G - \{e\}$. If the algorithm is ever called on an empty graph, then output the set of edges removed up to that point, which constitutes a matching. It is straightforward to verify that this algorithm operates in output polynomial time. □

C Proofs of Section 4

C.1 Proof of Lemma 4.2

Proof. Assume $\Pi \subseteq \Sigma^*$ is well-formed, set $x = C_R(\Pi)$, and set $y = C_S(g(\Pi))$.

Suppose that Π is a yes-instance of R^{-1}, that is, $\Pi \in R^{-1}$. We have $R(x) = \Pi$. Suppose that $\pi \in g(\Pi)$; then $\pi \in S(C_S(g(\Pi))) = S(y)$, implying the containment $g(\Pi) \subseteq S(y)$. Since $|g(\Pi)| = |\Pi| = |R(x)| = |S(y)|$, we conclude that $g(\Pi) = S(y)$ and that $g(\Pi) \in S^{-1}$.

Suppose Π is a no-instance of R^{-1}, that is, $\Pi \notin R^{-1}$. Then $|\Pi| < |R(x)|$. We have the equalities $|\Pi| = |g(\Pi)|$ and $|R(x)| = |S(C_S(g(\Pi)))|$, so $|g(\Pi)| < |S(C_S(g(\Pi)))|$. It follows that $g(\Pi) \neq S(C_S(g(\Pi)))$, so by the last condition in the definition of a candidate function, $g(\Pi) \notin S^{-1}$. □

C.2 Proof of Lemma 4.4

Proof. Fix A to be a set that is not well-formed with respect to S. For all $\Pi \subseteq \Sigma^*$ define $g(\Pi)$ to be $f_{C_R(\Pi)}(\Pi)$ if Π is well-formed, and A otherwise. It is straightforward to verify that g is a π-reduction, so $R^{-1} \leq_m^p S^{-1}$ by Lemma 4.2. □

C.3 Proof of Theorem 4.5

We establish Theorem 4.5 by giving a sequence of theorems, all of which prove coNP-completeness of one of the given inverse problems.

Theorem C.1 INVERSE 3-SAT $\leq_m^p$ INVERSE EXACT COVER.

Before proving this theorem, we introduce some terminology, borrowed from [9]. Let Π be a set of mappings $\{v_1, \ldots, v_n\} \to \{T, F\}$. We say that a mapping $\sigma : \{v_1, \ldots, v_n\} \to \{T, F\}$ is *3-compatible* with Π if for any sequence of three positions $1 \leq i < j < k \leq n$, there exists a mapping $\pi \in \Pi$ that agrees with σ in these three positions: $\sigma(v_i) = \pi(v_i)$, $\sigma(v_j) = \pi(v_j)$, and $\sigma(v_k) = \pi(v_k)$.

Lemma C.2 *[9] Let Π be a set of mappings $\{v_1, \ldots, v_n\} \to \{T, F\}$, and let R be the verifier for* INVERSE 3-SAT. *The set $R(C_R(\Pi))$ consists of exactly those mappings $\sigma : \{v_1, \ldots, v_n\} \to \{T, F\}$ such that σ is 3-compatible with Π.*

Proof (Theorem C.1). Let R, S denote the verifiers for INVERSE 3-SAT and INVERSE EXACT COVER, respectively. By Lemma 4.4, it suffices to define a π-function f from R to S. Define $f(x, \pi)$ to contain $2n + |x|$ sets: for each $i \in [n]$, it contains the two sets

$$\{v_i\} \cup \{\langle v_i, \pi(v_i)\rangle\} \cup \{\langle v_i, c, \pi(v_i)\rangle : v_i \text{ is in clause } c, c \in x\}$$

and

$$\{\langle v_i, \neg\pi(v_i)\rangle\},$$

and for each clause $c \in x$ with variable set $\{v_i, v_j, v_k\}$, it contains

$$\{c, \langle v_i, c, \neg\pi(v_i)\rangle, \langle v_j, c, \neg\pi(v_j)\rangle, \langle v_k, c, \neg\pi(v_k)\rangle\}.$$

Note that the base set of the resulting proofs is

$$\{v_i : i \in [n]\} \cup \{\langle v_i, a\rangle : i \in [n], a \in \{T, F\}\} \cup \{c : c \in x\} \cup \{\langle v_i, c, a\} : i \in [n], c \in x, a \in \{T, F\}\}.$$

To verify that f is a π-function, we need to show that for all well-formed $\Pi \subseteq \Sigma^*$, if $x = C_R(\Pi)$ and $y = C_S(f_x(\Pi))$, then f_x is a bijection between $R(x)$ and $S(y)$.

We first show that f_x is injective. If $\pi_1 \neq \pi_2$ then there exists some variable v_i such that $\pi_1(v_i) \neq \pi_2(v_i)$. The coverings $f_x(\pi_1)$ and $f_x(\pi_2)$ cover the element v_i with different sets: one of the sets includes $\langle v_i, F\rangle$ but not $\langle v_i, T\rangle$, and the other includes $\langle v_i, T\rangle$ but not $\langle v_i, F\rangle$.

We now show that f_x is surjective. Suppose that $\pi' \in S(y)$. Then for all $U \in \pi'$, $U \in f_x(\pi)$ for some $\pi \in \Pi$ (by the definitions of f_x and C_S). For each cover $f_x(\pi)$, the set covering v_i contains either $\langle v_i, F\rangle$ or $\langle v_i, T\rangle$ (for all $i \in [n]$). By the definition of C_S, the set in π' covering v_i must contain either $\langle v_i, F\rangle$ or $\langle v_i, T\rangle$; define $\tau(v_i)$ to be F or T accordingly.

We claim that $\tau \in R(x)$. By Lemma C.2, it suffices to show that τ is 3-compatible with Π. Suppose that $1 \leq i < j < k \leq n$. If there is no clause in x on variables v_i, v_j, and v_k, then by the definition of C_R

there is, for any $a_i, a_j, a_k \in \{T, F\}$, an assignment $\pi \in \Pi$ with $\pi(v_i) = a_i$, $\pi(v_j) = a_j$, and $\pi(v_k) = a_k$. On the other hand, if there is a clause c in x on variables v_i, v_j, v_k, then π' contains a set U_c covering c. By the initial observation (in this proof of surjectivity), $U_c \in f_x(\pi)$ for some $\pi \in \Pi$. It is easy to see that $\pi(v_i) = \tau(v_i)$, $\pi(v_j) = \tau(v_j)$, and $\pi(v_k) = \tau(v_k)$, establishing that τ is 3-compatible with Π.

A straightforward verification shows that $f_x(\tau) = \pi'$, concluding the proof that f_x is surjective. □

We proceed to give a sequence of reductions between the remaining inverse problems. For all of the reductions, the strategy is to give a π-frunction from the verifier of the first problem, denoted by R, to the verifier of the second. In the following proofs, we define a π-function f on pairs (x, π) such that $R(x, \pi)$ is true; verification that the defined functions are indeed π-functions is left to the reader.

Theorem C.3 INVERSE K-CLIQUE $\leq_m^p$ INVERSE VERTEX COVER.

Proof. Assume $x = \langle G, k\rangle$ and $\pi = H$. Define $f(x, \pi)$ to be $\langle l, R\rangle$ where $l = |V(G)| - k$ and $R = V(G) \setminus H$. □

Theorem C.4 INVERSE EXACT COVER $\leq_m^p$ INVERSE KNAPSACK.

Proof. Assume $x = \langle B,\mathcal{U}\rangle$ and $\pi \subseteq \mathcal{U}$. We assume that the sets B and $\mathcal{U}$ are totally ordered: $B = \{b_1, \ldots, b_{|B|}\}$ and $\mathcal{U} = \{U_1, \ldots, U_{|\mathcal{U}|}\}$. Set $d = |\mathcal{U}| + 1$ and $b = \sum_{t=0}^{|B|-1} d^t$. Define $g(j)$ to be $\sum_{t=0}^{|B|-1} d^t[b_{t+1} \in U_j]$. Define $f(x, \pi)$ to be the set of integers $\{g(j) : U_j \in \pi\}$. □

Theorem C.5 INVERSE EXACT COVER $\leq_m^p$ INVERSE STEINER TREE.

Proof. Assume $x = \langle B,\mathcal{U}\rangle$ and $\pi \subseteq \mathcal{U}$. Define $f(x, \pi) = \langle |B|, T\rangle$ where T has edges $\{v_0, U\}$ with weight $|U|$ for all $U \subseteq \pi$, and edges $\{U, b\}$ with weight 0 for all $b \in U \in \pi$. (The element v_0 is assumed to be outside of the set B.) □

Theorem C.6 INVERSE KNAPSACK $\leq_m^p$ INVERSE PARTITION.

Proof. Assume $x = \langle A, b\rangle$ and $\pi \subseteq A$. Define $f(x, \pi) = \{\{(\sum_{a\in A} a) + 1 - b\} \cup \pi, \{b + 1\} \cup (A \setminus \pi)\}$. □

C.4 Proof of Corollary 4.6

Proof. For all of the problems except EXHAUSTIVE CIRCUIT-SAT, this follows from Lemma 2.10 in conjunction with Theorems 3.1, 3.2 and 4.5. EXHAUSTIVE 3-SAT reduces to EXHAUSTIVE CIRCUIT-SAT by the identity reduction. □

D Proofs of Section 5

D.1 Proof of Lemma 5.1

Proof. Let q be a polynomial such that for all $x, \pi \in \Sigma^*$, $R(x, \pi)$ implies $|\pi| \leq q(|x|)$. Given x, there are $|\Sigma|^{q(|x|)}$ potential proofs of x, and so the set $R(x)$ can be generated in time $|\Sigma|^{q(|x|)}$ times a polynomial in $|x|$. Define a verifier R' such that $R'(x, \pi)$ is true when $\pi = \langle s\pi', x\rangle$ for some $s \in \Sigma^{q(|x|)}$ and $R(x, \pi')$ holds. The set $R'(x)$ has a closure property: for any proof $\langle s\pi', x\rangle$ in $R'(x)$, if $|t| = |s|$ then $\langle t\pi', x\rangle$ is also in $R'(x)$. To decide membership of Π in R'^{-1}, check to see if Π obeys the described closure property; if it does, run the algorithm generating $R(x)$ (which takes time polynomial in $\|\Pi\|$), and accept if the output is equal to Π. □

D.2 Proof of Lemma 5.2

Proof. Let R be a verifier such that $L(R) = B$. Define $R'(x, \pi)$ to be true if $\pi = \langle x, \pi' \rangle$, and $R(x, \pi')$ is true. Define $S(x, y)$ to be true if $R'(y, x)$ is true. It is straightforward to verify that S is fair; moreover, $x \in B$ if and only if $\{x\} \in S^{-1}$, so $S^{-1} \equiv_m^p B$. $\square$

D.3 Proof of Theorem 5.3

Let $\mathcal{V}$ be a set of variables. Call $\langle V, A \rangle$ a *3-pattern* of $\mathcal{V}$ if $V \subseteq \mathcal{V}$ is a subset of size 3, and $A : V \to \{T, F\}$ is a truth assignment to the variables of V. When P is a set of 3-patterns of $\mathcal{V}$, say that $f : \mathcal{V} \to \{T, F\}$ *satisfies* P if for all subsets $V \subseteq \mathcal{V}$ of size 3, $\langle V, f_{|V} \rangle \in P$.

Theorem D.1 *[9] Let S denote the verifier for* INVERSE 3-SAT*, and C_S denote the corresponding candidate function. There exists a polynomial time computable mapping σ from 3-patterns over a variable set to truth assignments (over a larger variable set) such that the following conditions hold.*

- *For all sets P of 3-patterns over the same variable set, $S(C_S(\sigma(P)))$ contains a truth assignment not in $\sigma(P)$ if and only if P is satisfiable.*
- *σ is injective.*
- *σ is invertible: there exists a polynomial time computable function τ such that for all 3-patterns p, $\tau(\sigma(p)) = p$.*

Proof (Theorem 5.3). Let $\forall\exists$3-SAT denote the language

$$\{\phi : \text{ for all } f : \mathcal{Y} \to \{T, F\} \text{ there exists } g : \mathcal{Z} \to \{T, F\} \text{ such that } \phi \text{ is satisfied by } f, g\}$$

where the ϕ are sets of 3-clauses over the variable set $\mathcal{Y} \cup \mathcal{Z}$.

We will define a fair verifier R such that $\phi \notin \forall\exists$3-SAT if and only if $M^\phi \in R^{-1}$, where M^ϕ is a set computable from ϕ in polynomial time. This suffices, as $\forall\exists$3-SAT is complete for Π_2^p [16], so the complement of $\forall\exists$3-SAT is complete for Σ_2^p.

Let σ denote the function given by Theorem D.1. Define M^ϕ to be the set of strings of the form $\langle \phi, \langle Y, A_Y \rangle, \sigma(\langle Z, A_Z \rangle) \rangle$ such that $\langle Y, A_Y \rangle$ is a 3-pattern of $\mathcal{Y}$, $\langle Z, A_Z \rangle$ is a 3-pattern of $\mathcal{Z}$, and instantiating the variables $Y \cup Z$ in ϕ with the values given by A_Y and A_Z does not lead to a contradiction. (That is, for no clause are all of the literals mapped to false under such an instantiation.)

Let $f : \mathcal{Y} \to \{T, F\}$ be a truth assignment to the variable set $\mathcal{Y}$. Let us say that a 3-pattern $\langle Z, A_Z \rangle$ over variable set $\mathcal{Z}$ is *f-compatible (with respect to ϕ)* if for all subsets $Y \subseteq \mathcal{Y}$ of size 3, $\langle \phi, \langle Y, f_{|Y} \rangle, \sigma(\langle Z, A_Z \rangle) \rangle$ is in M^ϕ. Let B_ϕ^f be the set of all 3-patterns over $\mathcal{Z}$ that are f-compatible with respect to ϕ.

Claim. For all $f : \mathcal{Y} \to \{T, F\}$, B_ϕ^f is satisfied by $g : \mathcal{Z} \to \{T, F\}$ if and only if ϕ is satisfied by the mapping $h : \mathcal{Y} \cup \mathcal{Z} \to \{T, F\}$ defined by $h_{|\mathcal{Y}} = f$, $h_{|\mathcal{Z}} = g$.

For the forward direction of the claim, suppose that B_ϕ^f is satisfied by g. Let c be a clause of ϕ. Pick variable sets $Y \subseteq \mathcal{Y}$ and $Z \subseteq \mathcal{Z}$, each of size 3, such that $Y \cup Z$ covers the variable set of c. By definition of B_ϕ^f, $\langle Z, g_{|Z} \rangle$ is f-compatible (with respect to ϕ), so $\langle \phi, \langle Y, f_{|Y} \rangle, \sigma(\langle Z, g_{|Z} \rangle) \rangle \in M^\phi$. By the definition of M^ϕ and the injectivity of σ, we have that c is not falsified by instantiating the variables $Y \cup Z$ in ϕ with the values given by $f_{|Y}$ and $g_{|Z}$. Since $Y \cup Z$ covers the variables in c, the clause c is satisfied by h.

For the reverse direction of the claim, suppose that h satisfies ϕ. Fix a subset $Z \subseteq \mathcal{Z}$ of size 3. For all subsets $Y \subseteq \mathcal{Y}$ of size 3,

$$\langle \phi, \langle Y, f_{|Y} \rangle, \phi(\langle Z, g_{|Z} \rangle) \rangle = \langle \phi, \langle Y, h_{|Y} \rangle, \phi(\langle Z, h_{|Z} \rangle) \rangle \in M^\phi.$$

Thus $\langle Z, g_{|Z} \rangle$ is f-compatible and hence in B^f_ϕ.

We are now in a position to describe our verifier R. Let S denote the verifier for INVERSE 3-SAT, and C_S denote the corresponding candidate function. Define $R(x, \pi)$ to be true if both

1. x is of the form $\langle \phi, f : \mathcal{Y} \to \{T, F\} \rangle$ where ϕ is a 3-SAT formula over variable set $\mathcal{Y} \cup \mathcal{Z}$, and
2. $\pi \in M^\phi$ or $\pi = \langle \phi, \langle Y, f_{|Y} \rangle, \tau \rangle$ where $\tau \in S(C_S(\sigma(B^f_\phi)))$, and $Y \subseteq \mathcal{Y}$ is of size 3.

We show that $\phi \notin \forall\exists$3-SAT if and only if $M^\phi \in R^{-1}$.

Suppose $\phi \notin \forall\exists$3-SAT . Let $f : \mathcal{Y} \to \{T, F\}$ be a truth assignment such that $\phi[f]$ (which denotes ϕ with the variables in $\mathcal{Y}$ instantiated using f) is unsatisfiable. By the claim, B^f_ϕ is unsatisfiable, so by Theorem D.1, $S(C_S(\sigma(B^f_\phi))) = \sigma(B^f_\phi)$. Thus $M^\phi = R(\langle \phi, f \rangle)$, and $M^\phi \in R^{-1}$.

Suppose $M^\phi \in R^{-1}$. Then $M^\phi = R(\langle \phi, f \rangle)$ for some $f : \mathcal{Y} \to \{T, F\}$. For all $\tau \in S(C_S(\sigma(B^f_\phi)))$ and $Y \subseteq \mathcal{Y}$ of size 3, $\langle \phi, \langle Y, f_{|Y} \rangle, \tau \rangle \in M^\phi$, implying that $S(C_S(\sigma(B^f_\phi))) \subseteq \sigma(B^f_\phi)$. Thus $S(C_S(\sigma(B^f_\phi))) = \sigma(B^f_\phi)$, implying by Theorem D.1 that B^f_ϕ is unsatisfiable. By the claim, it follows that $\phi[f]$ is unsatisfiable, and we conclude that $\phi \notin \forall\exists$3-SAT . □

D.4 Proof of Theorem 5.4

Proof. Fix $x_0 \in L(R)$. Define $T(\langle x, y \rangle, \pi)$ to be true if $\pi = 0\pi'$ for some $\pi' \in R(x)$; or, if $x = x_0$ and $\pi = 1\pi'$ for some $\pi' \in S(y)$. Observe that $L(T) = \{\langle x, y \rangle : x \in L(R), y \in \Sigma^*\}$, so clearly $L(T) \equiv^p_m L(R)$.

$R^{-1} \leq^p_m T^{-1}$: By the reduction $f(\Pi) = \{0\pi : \pi \in \Pi\}$.
$S^{-1} \leq^p_m T^{-1}$: By the reduction $f(\Pi) = \{1\pi : \pi \in \Pi\} \cup \{0\pi : \pi \in R(x_0)\}$.
$T^{-1} \leq^p_m R^{-1} \otimes S^{-1}$: We define a reduction f based on three cases.

- If Π contains a string with 1 as its first character and the set of strings in Π beginning with 0 is exactly $\{0\pi : \pi \in R(x_0)\}$, $f(\Pi) = 1\Pi'$ where $\Pi' = \{\pi : 1\pi \in \Pi\}$.
- If Π contains no string with 1 as its first character, $f(\Pi) = 0\Pi'$ where $\Pi' = \{\pi : 0\pi \in \Pi\}$.
- Otherwise, $f(\Pi) = 0\Pi'$ where $\Pi' \notin R^{-1}$.

□

D.5 Proof of Corollary 5.5

Proof. Obtain R from Lemma 5.1. If B is in NP, obtain S from Lemma 5.2; if B is Σ^p_2-complete, obtain S from Theorem 5.3. Apply Theorem 5.4 to R and S. □

D.6 Proof of Theorem 5.8

Proof (sketch). Let R be the verifier for CIRCUIT-SAT. If B is in NP, obtain S from Lemma 5.2; if B is Σ^p_2-complete, obtain S from Theorem 5.3. Apply Theorem 5.4 to R and S to obtain T'. It can be verified that $L(T')$ is a p-cylinder (for the definition of this notion and related results, see for instance [2]). Hence, $L(T')$ is in the same isomorphism class as the languages considered in [4]. Let $i : \Sigma^* \to \Sigma^*$ be an isomorphism

between A and $L(T')$. Define $T(x, \pi)$ to be true if $T'(i(x), \pi)$ is true. For all $\Pi \subseteq \Sigma^*$, $\Pi \in T^{-1}$ if and only if $\Pi \in T'^{-1}$. $\square$