Optimal Advice

Lane A. Hemaspaandra*
Department of Computer Science
University of Rochester
Rochester, NY 14627

Leen Torenvliet†
Dept. of Math. and Comp. Sci.
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam

November, 1994; Corrected Typescript Prepared June, 1995

Abstract

Ko proved that the P-selective sets are in the advice class P/quadratic, and Hemaspaandra, Naik, Ogihara, and Selman showed that they are in PP/linear. We strengthen the latter result by establishing that the P-selective sets are in NP/linear ∩ coNP/linear. We show linear advice to be optimal.

1 Introduction

Selective sets are sets for which there is a “selector function,” usually a polynomial-time deterministic or nondeterministic function, that declares one if its (two) inputs to be logically no less likely than the other to belong to the given set.

Definition 1.1 [HNOSb] Let $\mathcal{F}$ be any class of functions (possibly multivalued or partial). A set $A$ is $\mathcal{F}$-selective if there is a function $f \in \mathcal{F}$ such that for every $x$ and $y$, it holds that

$$f(x, y) \subseteq \{x, y\}, \text{ and }$$

if $\{x, y\} \cap A \neq \emptyset$, then $f(x, y) \neq \emptyset$ and $f(x, y) \subseteq A$.

Let $\mathcal{F}$-sel denote the class of sets that are $\mathcal{F}$-selective.

The class that would be notated FP$_{\text{single-valued, total}}$-sel according to the definition above was defined directly by Selman in 1979 [Sel79]. Henceforward, we refer to these sets by his term, P-selective, and we denote the class of such sets as P-sel. Also of interest

*Supported in part by grants NSF-CCR-8957604, NSF-INT-9116781/JSPS-ENGR-207, and NSF-CCR-9322513. Work done in part while visiting the University of Amsterdam.
†Supported in part by HC&M grant ERB4050PL93-0516.
to us will be the NPSV-selective sets, which were first studied by Hemaspaandra, Naik, Ogihara, and Selman [HNOs], and the NPSV\textsubscript{total}-selective sets, which were first studied by Hemaspaandra et al. [HHO+93]. NPSV and NPSV\textsubscript{total} denote the standard notions of single-valued NP functions, as introduced by Book, Long, and Selman [BLS84,BLS85]. Selective sets have recently been the focus of quite a bit of research activity (see the survey [DHT94]).

All three types of selective sets just mentioned are known to be contained in polynomial-advice classes. (We define these formally later. Loosely put, $C/f(n)$ is the class of languages for which the membership of all length $n$ strings can be decided by a set from $C$ that is in addition given, on input $x$, $f(|x|)$ extra bits of “advice” that can depend only on $x$’s length, not its value.) In particular, Ko [Ko83] proved—via a tournament in which the “losing” player and all players that that player lost to are eliminated each round—that $\text{P-sel} \subseteq \text{P/quad}$. Relatively, it has also been proven (in [HNOSb,HHO+93], with the “quadratic” claims being implicit in their proofs) that $\text{NPSV-sel} \subseteq (\text{NP} \cap \text{coNP})/\text{poly} \cap \text{NP/quad}$ and $\text{NPSV}_{\text{total-sel}} \subseteq (\text{NP} \cap \text{coNP})/\text{quad}$.

Is quadratic advice necessary to accept selectivity classes? Hemaspaandra et al. [HNOSa], via an approach quite different from Ko’s, proved that if one allows the advice interpreter to be quite powerful then linear advice (indeed $n + 1$ bits of advice) suffices: $\text{P-sel} \subseteq \text{PP/lin}$.\footnote{One cannot find this result explicitly stated in [HNOSa], but it is implicit in that paper’s proof that “If $\text{P} = \text{PP}$ then each $\text{P}$-selective set is many-one equivalent to a ‘standard left cut [of a real number]’. ’ That is, we are not claiming that the result follows from their just-quoted theorem itself; as in their construction the many-one reduction to the “standard left cut of a real number” has quadratic output length. Rather, the idea behind their proof is simply that for a $\text{P}$-selective set $A$ having $\text{P}$-selector function $f$ that $\text{w.l.o.g.}$ satisfies $(\forall x, y)[f(x, y) = f(y, x)]$ it holds that, for all $n$, $A^{−n}$ is exactly the set of length $n$ strings $z$ such that the number of length $n$ strings $w$ other than $z$ for which $f(z, w) = w$ is less than $|A^{−n}|$. So clearly $\text{P-sel} \subseteq \text{PP/lin}$. Indeed, in the notation we will define later, $\text{P-sel} \subseteq \text{PP}/\{2^n + 1\} \subseteq \text{PP}/n + 1$, where the advice is just the census value, $|A^{−n}|$. We note in passing that this implies the existence of certain deterministic advice-finding algorithms for $\text{P}$-selective sets.}

As historical background, we mention that lower bounds on advice date at least back to Kannan’s 1982 paper [Kan82]. Kübler and Thierauf [KT94] have, in the context of discussing the complexity of advice, also studied limiting to a logarithmic number of bits

\textbf{Theorem 1.2} If $A$ is a $\text{P}$-selective set, then $A$ has an advice-finding algorithm in $\#\text{P}^A$, and indeed that algorithm can be implemented deterministically in $\text{DTIME}^A[2^{O(n)}]$.

One can contrast this runtime with the natural deterministic advice-finding algorithm implicit in Ko’s construction, which would run in $\text{DTIME}^A[2^{O(n^2)}]$, and the natural deterministic implementation of the NP/linear algorithm of Corollary 3.9 of this paper, which would run in $\text{DTIME}^A[2^{\Theta(n \log^2 n)}]$ (in both cases, we use $\Theta$ to refer to the particular, concrete, brute-force algorithm—other as yet undiscovered approaches might yield different times).
the amount of advice. A recent paper of Homer and Mocas [HM93] obtains a lower bound for the advice complexity of exponential time that is optimal in some oracle world; in contrast, our result is unconditionally optimal.

Why should one care about determining the optimal number of bits of advice? One reason might be that, as noted in the previous footnote, in many settings it will hold that the shorter the advice, the easier the brute force procedure to find a correct advice string. More importantly, one long-term reason is that one hopes that contrasts between the number of bits of advice (the \( f(n) \) of \( C/f(n) \)) that suffice with respect to various strengths of advice interpreters (the \( C \) of \( C/f(n) \)) will yield insight in the powers of, and the differences between, complexity classes \( C \). For example, the facts that \( P\text{-sel} \subseteq P/\text{quadratic} \) and \( P\text{-sel} \subseteq \text{NP}/\text{linear} \) frame a potential difference between the powers of \( P \) and \( \text{NP} \). Certainly, from our results it follows that it would be hard to prove that \( P\text{-sel} \) lacks linear-sized advice that can be interpreted in deterministic polynomial time (that is, to prove \( P\text{-sel} \not\subseteq P/\text{linear} \)), as any such proof would establish that \( P \neq \text{NP} \).

2 Definitions: k-ary Advice

All our sets are over the alphabet \( \Sigma = \{0, 1\} \). For any set \( A \), let \( A^{=k} \) denote the length \( k \) strings in \( A \). Let \( \langle a, b \rangle \) be a standard logspace computable and invertible pairing function. Let \( \mathbb{Z}^{\geq 0} \) denote \( \{0, 1, 2, \ldots\} \) and \( \mathbb{Z}^{\geq 1} \) denote \( \{1, 2, \ldots\} \). All uses of \( \log \) and \( \log^* \) in this paper assume that the logarithms are of base 2. For any class \( C \), let \( \text{co-}C \) denote \( \{A | \overline{A} \in C \} \). By convention, we use \( \text{coNP} \) as a shorthand for \( \text{co-NP} \). We write \( x \leq_{\text{lex}} y \) if \( x = y \) or \( y \) is greater than \( x \) with respect to the standard lexicographical ordering.

Karp and Lipton [KL80] introduced the notion of advice classes as follows.

**Definition 2.1 [KL80]**

1. Let \( f \) be a function from \( \mathbb{Z}^{\geq 0} \) to \( \mathbb{Z}^{\geq 0} \). Let \( C \) be any collection of subsets of \( \{0, 1\}^* \). Define \( \frac{C}{f(n)} = \{A | (\exists B \in C) (\forall h : \mathbb{Z}^{\geq 0} \rightarrow \{0, 1\}^*) (\forall n) [f(n) = |h(n)|] (\forall x \in \{0, 1\}^*) [x \in A \iff \langle x, h(|x|) \rangle \in B] \}. \)

2. Let \( \mathcal{F} \) be any class of functions from \( \mathbb{Z}^{\geq 0} \) to \( \mathbb{Z}^{\geq 0} \). Define \( \frac{C}{\mathcal{F}} = \{A | (\exists f \in \mathcal{F}) [A \in \frac{C}{f}] \}. \)

Let \( \text{poly}, \text{linear}, \text{and quadratic} \) respectively denote the classes of polynomially bounded, linearly bounded, and quadratically bounded functions.

Note that Karp and Lipton’s above definition of \( \frac{C}{f(n)} \) requires all potential advice strings for an input of some length \( n \) to be of equal length. Thus their notion—in contrast with that found in some other papers that by an advice bound of \( f(n) \) mean that the advice is some string of length less than or equal to \( f(n) \)—exactly captures in \( \frac{C}{f(n)} \) the notion of \( f(n) \) bits of advice (though both definitions yield the same notion of, e.g., \( P/\text{poly} \), due to the unioning over all polynomials).
Karp and Lipton, in their seminal paper, not only suggested (and started) the study of advice, but also emphasized the importance of the length of advice, in particular by studying both logarithmic advice and polynomial advice. In order to make the amount of advice needed for a certain set as precise as possible, we refine their advice notion. They allow advice strings that take on one of a number of values that must be a power of two; we allow advice to take on one of a number of values not limited to powers of two. Thus, we introduce the notion of “\(k\)-ary advice.”

**Definition 2.2 \(k\)-ary advice**

1. Let \(h : \mathbb{Z}^\geq 0 \rightarrow \mathbb{Z}^\geq 1\). Let \(g\) be a function from \(\mathbb{Z}^\geq 0\) to \(\mathbb{Z}^\geq 1\). Assume natural numbers have their standard encoding over binary strings. Let \(\mathcal{C}\) be any collection of subsets of \(\{0, 1\}^*\). Define

\[
\mathcal{C}/\{g(n)\} = \{A \mid (\exists B \in \mathcal{C})(\exists h : \mathbb{Z}^\geq 0 \rightarrow \mathbb{Z}^\geq 1) \left[ (\forall n)[h(n) \in \{1, \cdots, g(n)\}] \text{ and } (\forall x \in \{0, 1\}^*)[x \in A \iff \langle x, h(x) \rangle \in B] \right] \}.
\]

2. Let \(\mathcal{G}\) be any class of functions from \(\mathbb{Z}^\geq 0\) to \(\mathbb{Z}^\geq 1\). Let \(\mathcal{C}\) be any collection of subsets of \(\{0, 1\}^*\). Define

\[
\mathcal{C}/\mathcal{G} = \{A \mid (\exists g \in \mathcal{G})[A \in \mathcal{C}/g] \}.
\]

**Lemma 2.3** For any class \(\mathcal{C}\) closed under composition with logspace functions, and for any \(f(n) : \mathbb{Z}^\geq 0 \rightarrow \mathbb{Z}^\geq 1\), \(\mathcal{C}/\{2^{f(n)}\} = \mathcal{C}/f(n)\).

We note that refining from powers of two to a \(k\)-ary token is not a new idea. Cai and Furst, in the setting of bottleneck machines ([CF91], see also the underlying paper Barrington [Bar89]) wanted to precisely describe the point at which bottleneck machines captured PSPACE, and to do this they defined a machine model in which they could speak of passing, e.g., a 5-ary token (rather than, e.g., the cruder “3 bits”).

## 3 Results

We now turn to our results.

**Theorem 3.1** \(P\text{-sel} \subseteq \text{NP}/\{2^n + 1\} \cap \text{coNP}/\{2^n + 1\}\).

Via Lemma 2.3, we immediately have the following corollary.

**Corollary 3.2** \(P\text{-sel} \subseteq \text{NP}/n + 1 \cap \text{coNP}/n + 1\).

**Proof of Theorem 3.1** Let \(A \in P\text{-sel}\). Since \(P\text{-sel}\) is closed under complementation, it suffices to show that \(A \in \text{NP}/\{2^n + 1\}\). Let \(D_m\) denote the class of all simple (that is, having no self-loops) graphs \(G\) on nodes \(\{1, 2, \cdots, m\}\) such that for each \(i \neq j\) exactly one of \((i,j)\) or \((j,i)\) belongs to \(G\). That is, \(D_m\) is the class of graphs obtained by directing the edges of the complete graph \(K_m\). Given an \(\ell\) element collection \(B\) (for simplicity, name
the elements by \( B = \{1, 2, \ldots, \ell\} \), the action of a P-selector \( f \) implicitly specifies a graph \( G \in D_t \), via the rule (for \( i \neq j \)) that \((i, j) \in G\) if and only if \( f(\min(i, j), \max(i, j)) = j \), and otherwise \((j, i) \in G \).

Ko [Ko83] let \( B \) correspond to the set of elements of length \( n \) in a given P-selective set. He proved \( \text{P-sel} \subseteq \text{P}/\text{quadratic} \) essentially by noting that for each \( G \in D_{2^n} \) (and thus for \( G \in D_k, k < 2^n \)) there exists a set of \( n \) nodes from which every node can be reached by a path of length 0 or 1. We first note (for completeness—we suspect this to be a folk theorem) that for each \( G \in D_{2^n} \) there exists some node from which every node can be reached via a path of length at most \( n \). To see this, consider the following multi-state “knock-out” tournament. In the first round, 1 is compared to 2, 3 is compared to 4, and so on. From each pair, the “loser” (that is, the one with the edge pointing away from it) remains in the tournament and the winner is retired. In each subsequent round, the remaining players are again paired off and the process is repeated. Note that after \( n \) rounds one player survives, and this “super-losing” player reaches every other player via a directed path of length at most \( n \).

We can now state our \( \text{NP}/\{2^n + 1\} \) procedure for any given P-selective set \( A \) having selector function \( f \). We first specify the advice for length \( n \). If \( A \cap \Sigma^n = \emptyset \), use advice value \( 2^n + 1 \). Otherwise, play the previously described tournament among \( A \cap \Sigma^n \). (Note that there are at most \( 2^n \) elements in this intersection. Thus, even though each tournament round may have one element that fails to be paired off, and thus that is promoted for free, the tournament will produce a single super-loser in at most \( n \) rounds—indeed, in \( \lceil \log_2 |A \cap \Sigma^n| \rceil \) rounds.) If the lexicographically \( i \)-th length \( n \) string is the super-loser, use advice value \( i \). The NP machine, given advice value \( j \) and input \( x \), rejects if \( j = 2^n + 1 \), and otherwise accepts if and only if there exists a directed path with (respect to the P-selector \( f \)) of length at most \( n \) from the lexicographically \( j \)-th length \( n \) string, call it \( s_{j,n} \), to \( x \) (that is, if either \( s_{j,n} = x \) or there exist \( s_{j,n} = v_0, v_1, \ldots, v_{\ell} = x, \ell \leq n \), such that \((\forall i : 1 \leq i \leq \ell) [f(v_{i-1}, v_i), \max(i, v_i)] = v_i \)).

Thus, in comparison with Ko’s result that \( \text{P-sel} \subseteq \text{P}/\text{quadratic} \), we have proven that less advice suffices given a more powerful acceptance mechanism: \( \text{P-sel} \subseteq \text{NP}/\text{linear} \cap \text{coNP}/\text{linear} \). In comparison with Hemaspaandra et al.’s [HNOSa] result that \( \text{P-sel} \subseteq \text{PP}/\text{linear} \), we have shown that a potentially weaker (as \( \text{NP} \subseteq \text{PP} \) and the inclusion is strict unless, for example, \( \text{NP} = \text{coNP} \)) interpreter can make do with linear advice. Our result is incomparable (that is, neither seems to imply the other) with an interesting result of Burtschick and Lindner [BL93] stating that \( R_{\text{coNP}}^{\text{P}}(\text{P-sel}) \subseteq \text{E}/\text{linear} \).

Of course, it would be interesting to prove that \( \text{P-sel} \not\subseteq \text{P}/\text{linear} \) (as otherwise the possibility remains open that Ko’s result and ours can be simultaneously improved via establishing that \( \text{P-sel} \subseteq \text{P}/\text{linear} \)). However, Corollary 3.2 creates a potential difficulty in proving this “interesting” result: such a proof would also prove \( \text{P} \neq \text{NP} \). We state this as Corollary 3.3 below.

**Corollary 3.3** If \( \text{P-sel} \not\subseteq \text{P}/\text{linear} \), then \( \text{P} \neq \text{NP} \).
We now prove that the amount of advice used in Theorem 3.1, a \(2^n + 1\)-ary token, is optimal, even with respect to powerful advice interpreters. The proof seeks to wedge into a set as much information as possible (via diagonalizing against smaller amounts of information and all possible interpreters), while still maintaining the P-selectivity of the set, via a variant on a gap construction. Ladner [Lad75] pioneered the technique of putting large gaps in sets and then exploiting the gaps, and our proof is inspired by a construction that, in the context of determining the closure properties of the P-selective sets, uses gaps to establish P-selectivity [HJ].

**Theorem 3.4** Let \(f(n)\) be any recursive function. \(\text{P-se} \not\in \text{DTIME}[f(n)]/\{2^n\}\).

**Corollary 3.5** Let \(f(n)\) be any recursive function. \(\text{P-se} \not\in \text{DTIME}[f(n)]/n\).

**Proof of Theorem 3.4** Note that for any recursive function \(f(n) : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 1}\) there is a recursive function \(g(n)\) such that (a) for each \(n\), \(g(n) \geq f(n)\) and \(g(n) \geq 2\), and (b) \(g(n)\) is nondecreasing, and (c) for some Turing machine \(M\) computing \(g(n)\), it holds for every \(n\) that \(\text{RUNTIME}_M(n) \leq 2^{g(n)}\). If \(f\) did not originally have the above properties, henceforward replace it with \(g\).

Let \(\alpha_1 = 1\). For \(i \geq 1\), let \(\alpha_{i+1} = \min\{m | 2^{2^{2^{g(\alpha_i) - 1}}} \leq m\}\). Let \(Q = \{\alpha_1, \alpha_2, \ldots\}\). We will define a set \(A\) such that:

1. \(A \in \text{DTIME}[2^{g(n)}]\), and
2. \(x \in A \Rightarrow |x| \in Q\), and
3. \((|x| = |y|\) and \(x \leq_{\text{lex}} y\) and \(x \in A\) \(\Rightarrow y \in A\).

Any set \(A\) satisfying these three conditions is P-selective. Why? Suppose we are given strings \(x\) and \(y\). If either \(|x|\) or \(|y|\) is not in \(Q\), by (2) our task is trivial. If both lengths are in \(Q\) and \(|x| \neq |y|\), then by (1) and the definition of the \(\alpha_i\) the membership status of the shorter of \(x\) and \(y\) can be computed in time polynomial in \(\max(|x|, |y|)\). If both lengths are in \(Q\) and \(|x| = |y|\), then output, in light of (3), the lexicographically largest string in \(\{x, y\}\). Thus, we have just given a polynomial-time selector function for \(A\). We now turn to the construction of an \(A\) that meets these conditions, and that also satisfies the other requirement of the theorem.

Let \(\{M_i\}_{i \geq 1}\) by an enumeration of machines, and let this enumeration have the property that \((\bigcup_k L(M_k)) \supseteq \text{DTIME}[f(n)]\), and let it also have the property that, for each \(M_i\) in the enumeration, the function \(\text{RUNTIME}_{M_i}(n) = O(2^{f(n)})\). By the assumption at the start of the proof that computing \(f\) does not take "too much" longer than the value of \(f\), such an enumeration exists. Furthermore, let us assume, without loss of generality, that each machine in the list appears infinitely often in the list. That is, if \(M = M_i\) (here, equality refers to the actual program) for some \(i\), then there is an infinite set of distinct integers \(J\) such that for each \(j \in J\) we have \(M = M_j\). This assumption allows us to avoid having to explicitly "slow down" the enumeration. If \(n \not\in Q\), \(A^n = \emptyset\) is empty.
If \( n \in Q \) (say \( n = \alpha_i \)), then do the following. Keep a clock, and if at any point the simulation has taken \( \sqrt{n} \cdot 2^{f(n)} \) steps, cut off this stage. The clock overhead can easily be kept sufficiently low that, with this cutoff, both clock overhead time and simulation time as a total remain less than \( n^3 2^{f(n)} \). This clock-cutoff could happen, for example, if some machine \( M_i \) in our enumeration has a very large (relative to \( i \)) constant as part of its \( O(2^{f(n)}) \) runtime. However, as we have assumed each machine appears infinitely often in the enumeration, it follows from the fact that \( \sqrt{n} = \omega(1) \) that each machine will be cut off in only a finite number of its incarnations. For our purposes this suffices, as the incarnations that are not cut off will carry out the desired diagonalization against the machine.

Recall that \( n = \alpha_i \). Run \( M_i(x, y) \) for each \( x \) and \( y \) such that \( |x| = n \) and \( y \in \{1, 2, \cdots, 2^n\} \). In particular, for each \( \hat{y} \in \{1, 2, \cdots, 2^n\} \), consider \( M_i(x, \hat{y}) \) with \( x \) varying over all strings of length \( n \). If the acceptances of these \( 2^n \) runs with \( \hat{y} \) as the second component do not form a right cut of the length \( n \) strings,\(^2\) let \( z_{\hat{y}} = \text{undefined} \). If the acceptances of these \( 2^n \) runs do form a right cut (we take the two border cases of all acceptances and all rejections to be valid instances of right cuts), then let

\[
  z_{\hat{y}} = \begin{cases} 
    \min\{u \mid |u| = n \text{ and } M_i(u, \hat{y}) \text{ accepts}\} & \text{if } (\exists z) \{|z| = n \text{ and } M_i(z, \hat{y}) \text{ accepts}\} \\
    0^{1+n} & \text{otherwise.}
  \end{cases}
\]

Since there are only \( 2^n \) strings \( \hat{y} \) of length \( n \), at most \( 2^n \) right cuts are formed. So, in particular, \( F - B \neq \emptyset \), where \( F = \Sigma^n \cup \{0^{1+n}\} \) and \( B = \{w \mid |w| = n \text{ and } (\exists v) \{|v| \in \{1, 2, \cdots, 2^n\} \text{ and } z_v = w\}\} \). Let \( k_i \) (recall \( n = \alpha_i \)) be the lexicographically smallest element of \( F - B \).

We can now define \( A = n \) (for \( n = \alpha_i \); recall that \( A \) is empty for \( n \not\in Q \)):

\[
  A = n = \{v \mid |v| = \alpha_i \text{ and } k_i \leq_{\text{lex}} v\}.
\]

Note that by our construction, \( A \) satisfies (1), (2), and (3). However, as we argued, each machine will eventually be diagonalized against, and so for each \( M_j \in \{M_i\}_{i \geq 1} \) there is an \( \ell \) such that \( M_j = M_\ell \) and at stage \( \ell \) it was established that \( A \not\in \{L(M_i)\}/\{2^n\} \), where \( \{L(M_i)\} \) denotes the class containing only the set \( L(M_i) \).

We now turn briefly to optimal advice as it relates to nondeterministic selectivity. Note that though the statement of the first part of the following result looks like a generalization of Theorem 3.1, its proof shows that the two statements are in fact equivalent.
Proof of Proposition 3.6 (1) It is well-known that $NP^{NP \cap \text{coNP}} = NP$ and $\text{coNP}^{NP \cap \text{coNP}} = \text{coNP}$, and by essentially the same proof it is known [HHN+93] that $FP^{NP \cap \text{coNP}}_{\text{single-valued, total}} = \text{NPSV}_{\text{total}}$. Part (1) follows immediately from these facts and that fact that Theorem 3.1 relativizes (as, in particular, it holds relativized by each set $I \in NP \cap \text{coNP}$). Alternatively, this part is a corollary of Part (2) and the closure under complementation of NPSV$_{\text{total} \cdot \text{sel}}$.

(2) It is not hard to see by inspection that the proof of Theorem 3.1 (except for the part about closure under complementation and thus the coNP result) goes through (with the appropriate modifications—use the nondeterministic function everywhere) if the selector function itself is an NPS function. This is true because, though the function may in general be partial, by Definition 1.1 the function certainly is defined whenever both its inputs are in the set, and such is indeed the case in each match in the knockout tournament used in the proof of Theorem 3.1. (Indeed, even NPMV-selectivity suffices, as we can in the two-output case eliminate either element.)

Theorem 3.4 and Corollary 3.5 provide lower bounds showing that the advice lengths of Proposition 3.6 and Corollary 3.7 are optimal.

We conclude this section with some technical remarks on slightly stronger results implicit in our proofs, and on related claims. First, careful inspection of the proof of Theorem 3.1 reveals that in fact that proof as stated establishes:

$$P \cdot \text{sel} \subseteq \text{TIME-ND}[\text{poly, quadratic}]/\{2^n + 1\} \cap \text{co-TIME-ND}[\text{poly, quadratic}]/\{2^n + 1\}.$$  

Note that TIME-ND has been studied before; it is a level of the limited nondeterminism hierarchy for NP as introduced by Fischer and Kintala in [KF77] (this should not be confused with the other limited nondeterminism hierarchy for NP the same authors introduced in [KF80]). We now show how to save nondeterminism while retaining linear advice size. In particular, consider the following hybrid of Ko's P/poly tournament technique and our nondeterministic interpreter approach. In the proof of Theorem 3.1, start by conducting a tournament (via the technique of Ko described at that start of that proof) that finds $n$ elements such that all length $n$ strings in the set can be reached from those elements by a path of length at most 1. Among those $n$ elements, play a similar tournament. Continue this process until we are down to one element. Note that this will take at most $\log^*(2^n)$ rounds.\(^3\) Use the element found—which has an $O(\log^* n)$ length path to each length $n$ element in the set—as our advice. Retain the NP advice interpreter described in the proof of Theorem 3.1, except modify it to look only for such short paths. Thus, we have established:

**Theorem 3.8** $P \cdot \text{sel} \subseteq \text{TIME-ND}[\text{poly, } n(1 + \log^* n)]/\{2^n + 1\}$.

\(^3\)The reader may be surprised by the use of $\log^*$ in light of the fact that we iterate Ko's technique, and Ko's technique in fact in one round selects a set of size at most $\log(n + 1)$ rather than a set of size at most $\log n$ (recall, all logarithms in this paper are base two). However, since for all $n \geq 2$ it holds that $\log(n + 1) \leq \log n$, our use of $\log^*$ is not problematic.
Corollary 3.9 \( P\text{-sel} \subseteq \text{TIME-NONDETERMINISM}[\text{poly}, n(1 + \log^* n)]/n + 1. \)

Additionally, note that tournaments are quite sensitive to the number of players. In particular, one can state the following, where \( \log^* (n) \) is taken to be a shorthand for \( \lceil \log^* (\max(1, n)) \rceil \), and \( \text{census}_A(n) = \text{def} \ |A \cap \Sigma^n| \), and in which, for simplicity, we speak in terms of bits rather than in terms of tokens.

Proposition 3.10 1. If \( A \in P\text{-sel} \) and \( h(n) \) is a polynomial-time computable function such that \( h(n) \geq \text{census}_A(n) \), then

\[
A \in \text{TIME-NONDETERMINISM}[\text{poly}, n \log^*(h(n))]/n + 1.
\]

2. If \( A \in P\text{-sel} \) and \( h(n) \) is a polynomial-time computable function such that \( h(n) \leq \text{census}_A(n) \), then

\[
A \in (\text{co-TIME-NONDETERMINISM}[\text{poly}, n \log^*(2^n - h(n))])/n + 1.
\]

Finally, note that Corollary 3.9 is incomparable to Ko’s result. Ko uses quadratically many advice bits and no nondeterminism. We use the optimal number (i.e., linearly many) advice bits but also use a certain amount of nondeterminism—\( n(1 + \log^* n) \). Is there an advice-for-nondeterminism trade-off curve between these two claims? It is not hard to see, via modifying the proof (i.e., the text discussion before the Corollary) of Corollary 3.9, that we indeed have a trade-off curve. For \( i \in \mathbb{Z}_{\geq 0} \), let

\[
\log^i (n) = \begin{cases} 
  n & \text{if } i = 0 \\
  \log(\log^i (n)) & \text{if } i \in \mathbb{Z}_{\geq 1}.
\end{cases}
\]

Theorem 3.11 Let \( f(n) \) be any integer-valued function that is exponential-time computable (i.e., is polynomial-time computable if \( n \) is input in unary) and that satisfies \((\forall n \in \mathbb{Z}_{\geq 0})[1 \leq f(n) \leq \log^*(2^n)]\). Then

\[
P\text{-sel} \subseteq \text{TIME-NONDETERMINISM}[\text{poly}, n(f(n) - 1)]/\left\{ \sum_{0 \leq j \leq \lceil \log^*(f(n)/2^n) \rceil} \left( \begin{array}{c} 2^n \\ j \end{array} \right) \right\}.
\]

Ko’s result is the \( f(n) = 1 \) case, and our Theorem 3.8 is the \( f(n) = \log^*(2^n) \) case.

4 Open Questions

The most pressing open question regarding the results of this paper is whether \( P\text{-sel} \subseteq \text{P}/\text{linear} \). We conjecture that \( P\text{-sel} \not\subseteq \text{P}/\text{linear} \), however Corollary 3.3 suggests that our conjecture may be hard to prove (though it does not suggest that our conjecture fails to hold). We also suggest as an interesting area the study, for classes in \( \text{P}/\text{poly} \) other than the \( P\)-selective sets, of their optimal amount of advice, and of the relationship between the
amount of advice and the complexity of the set required to interpret the advice. We note that it is not hard to see that if $A$ is sparse (i.e., for some polynomial $p(\cdot)$ it holds that $(\forall n) [census_A(n) \leq p(n)]$), then

$$A \in P / \sum_{i \in \{0, 1, \ldots, census_A(n)\}} \left( \frac{2^n}{i} \right),$$

and there exist sparse sets $B$ for which

$$B \not\in P / \max(1, -1 + \sum_{i \in \{0, 1, \ldots, census_A(n)\}} \left( \frac{2^n}{i} \right)).$$

Acknowledgments

We thank M. Cierniak, S. Fischer, and Y. Han for interesting conversations. We thank E. Hemaspaandra and A. Selman for discussing their work with us, A. Selman for pointing out that $P = PP \Rightarrow P\text{-sel} \subseteq P/linear$, and P. van Emde Boas for pointing out the intermediate points in the advice-nondeterminism trade-off.

References


