On Lower Bounds of the Closeness Between Complexity Classes*

Bin Fu

Computer Science Department, Beijing Computer Institute, 100044 Beijing, People's Republic of China, and Laboratory of Cognitive Science, University of Science and Technology of China, Beijing, People's Republic of China

Abstract. We show that if an NP-m-hard set is the union of a set in $P_{bit}(\text{Sparse})$ and the set $A$, then $NP \subseteq P_{det}(A)$. Since co-NP, R, and FewP are closed under $\leq^p_{det}$-reductions, so no NP-m-hard set is the union of a set in $P_{bit}(\text{Sparse})$ and a set in co-NP (resp. R, FewP) unless $NP = \text{co-NP}$ (resp. $NP = R$, $NP = \text{FewP}$). We also study the distance between many important complexity classes. Let $A$, $B$ be two sets. The distance function $\text{dist}_{A,B}$ is defined as in [SI] such that $\text{dist}_{A,B}(n)$ is the number of strings of length $\leq n$ in $A \triangle B = (A - B) \cup (B - A)$. We show that there exists a set $A$ in $P_{(k+1)-tt}(\text{Sparse})$ such that, for every $B$ in $P_{k-tt}(\text{Sparse})$, $\text{dist}_{A,B}$ has an exponential lower bound.

1. Introduction

The relationship between the complexity classes P and NP is the central topic in structural complexity theory. It is widely believed that $P \neq \text{NP}$, so it may be impossible to find a polynomial-time algorithm to solve an NP-complete problem. People often try to construct polynomial algorithms to approximate the NP-complete problems. How much can an NP-complete set be approximated by a set in P is an interesting question. There may exist many different ways to measure the closeness of approximations. In this paper we use the density of the symmetric difference between two sets to characterize the closeness of approximation.

Yesha [Y] first considered the lower bound of symmetric difference between an NP-hard set and a set in P. He showed that for any NP-hard set $H$ and any

* This research was supported in part by HTP 863.
set \( A \) in \( P \), \( \text{dist}_{A,H}(n) \) is not \( O(\log(\log n)) \) unless \( P = NP \). Schöning [S1] showed that paddable NP-complete sets have no polynomial distance with the sets in \( P \) unless \( P = NP \). Recently as a corollary of Ogiwara and Watanabe’s result: \( \text{NP} \subseteq \text{PTIME} \Rightarrow P = NP \) [OW], no NP-hard set has a polynomial distance with some set in \( P \) unless \( P = NP \).

How about the distance between NP-hard sets with the sets of some other complexity classes beyond \( P \) (for example, co-NP, ZPP, BPP, \( \text{NP} \cap \text{co-NP} \))? In this paper we investigate the distances between some important complexity classes. We show that, for every NP-hard set \( H \) and any \( A \subseteq H \) in co-NP, \( \text{dist}_{A,H} \) has no polynomial upper bound unless \( \text{NP} = \text{co-NP} \). We also show that, for every \( \text{NEXP} \)-hard set \( H \) and every \( A \) in \( \text{EXP} \), \( \text{dist}_{A,H} \) has no subexponential upper bound unless \( \text{NEXP} = \text{EXP} \), etc.

The distances between some nonuniform complexity classes are also studied in this paper. Book and Ko [BK] separated \( \text{P}^{(k+1)} \text{tt}(\text{Sparse}) \) from \( \text{P}^{k} \text{tt}(\text{Sparse}) \). We use the notion of distance to give a deeper characterization of the difference between \( \text{P}^{(k+1)} \text{tt}(\text{Sparse}) \) and \( \text{P}^{k} \text{tt}(\text{Sparse}) \). We show that there exists a set \( A \) in \( \text{P}^{(k+1)} \text{tt}(\text{Sparse}) \) such that, for every \( B \) in \( \text{P}^{k} \text{tt}(\text{Sparse}) \), \( \text{dist}_{A,B}(n) \) has lower bound \( 2^{n/2^{k+2}} \) for sufficiently large \( n \).

2. Preliminaries

Throughout this paper we fix our alphabets \( \Sigma = \{0, 1\} \). By “string” we mean an element of \( \Sigma^* \). For a string \( x \) in \( \Sigma^* \), \( |x| \) denotes the length of \( x \). We consider a standard canonical order on \( \Sigma^* \). For any strings \( x \) and \( y \), \( x \) is canonically smaller than \( y \) (write \( x < y \)) if either \( |x| < |y| \) or \( |x| = |y| \) and there exists some \( k \), \( 1 \leq k \leq |x| \), such that, \( \forall i : 1 \leq i < k \) \( x_i = y_i \) and \( x_k = 0 \) and \( y_k = 1 \), where \( x_i \) is the \( i \)th alphabet of the string \( x \). Let \( S \subseteq \Sigma^* \), the cardinality of \( S \) is denoted by \( ||S|| \), and set \( S^=n (S^<n) \) consists of all words of length \( = n \) (\( < n \)) in \( S \). Especially, we let \( \Sigma^n = \{x|x \in \Sigma^* \text{ and } |x| = n\} \) and \( \Sigma^<=n = \{x|x \in \Sigma^* \text{ and } |x| \leq n\} \). For every string \( u \in \Sigma^* \) and \( A \subseteq \Sigma^* \), \( uA = \{ux|x \in A\} \). Let \( x, y \in \Sigma^* \), we define the interval \( [x, y] \): \( [x, y] = \{z|x \leq z \leq y \text{ and } z \in \Sigma^*\} \), and \( x \) is called the left point of interval \( [x, y] \). Let \( [a, b] \) and \( [c, d] \) be two disjoint intervals, we say \( [a, b] \leq [c, d] \) if \( b \leq c \). We use \( \lambda \) to denote the null string. \( N \) represents the set \( \{0, 1, 2, \ldots\} \).

We use the pairing function \( \langle \cdot, \cdot \rangle: \Sigma^* \times \Sigma^* \to \Sigma^* \). It is convenient to assume, for any \( x, y \) in \( \Sigma^* \), \( |\langle x, y \rangle| \leq 2(|x| + |y|) \). The function order: \( \Sigma^* \to N \), \( \text{order}(x) = i \) if the string \( x \) is the \( i \)th string among all strings of length \( |x| \) by canonical order.

Our computation model is the Turing machine. Now we involve the following complexity classes:

\[ \text{P: } \text{The class of languages accepted by deterministic Turing machines in polynomial time.} \]
\[ \text{PF: } \text{The class of polynomial-time computable functions.} \]
\[ \text{NP: } \text{The class of languages accepted by nondeterministic polynomial-time Turing machines.} \]