The Existence and Density of Generalized Complexity Cores

RONALD V. BOOK AND DING-ZHU DU

University of California, Santa Barbara, California

Abstract. If \mathscr{C} is a class of sets and A is not in \mathscr{C} , then an infinite set H is a proper hard core for A with respect to \mathscr{C} , if $H \subseteq A$ and for every $C \in \mathscr{C}$ such that $C \subseteq A$, $C \cap H$ is finite. It is shown that if \mathscr{C} is a countable class of sets of strings that is closed under finite union and finite variation, then every infinite set not in \mathscr{C} has a proper hard core with respect to \mathscr{C} . In addition, the density of such generalized complexity cores is studied.

Categories and Subject Descriptors: F.1.3 [Computation by Abstract Devices]: Complexity Classes machine-independent complexity

General Terms: Theory, Verification

Additional Key Words and Phrases: Generalized complexity cores, proper hard cores

1. Introduction

Consider a recursive set A that is not in the class P of sets recognizable deterministically in polynomial time. For every algorithm α that recognizes A and for every polynomial p, there exist infinitely many inputs x such that α s running time on x exceeds p(|x|); this follows from the fact that A is not in the class P. Lynch [12] showed that this condition implies that there exists an infinite set X such that for every algorithm α that recognizes A and every polynomial p, α 's running time on x exceeds p(|x|) on all but finitely many $x \in X$. The set X is a complexity core for X. Thus, a complexity core is a set of "hard" instances of candidates for membership in A. Lynch's proof shows that for every recursive set A not in P there exists $X \subseteq A$ such that X is a complexity core for A; we call such a complexity core proper. Thus, a proper complexity core X of a set A is a subset of A that is intrinsically difficult to recognize, and every recursive set not in P has an infinite proper complexity core.

If A is not in P, then it is still possible for A to have an infinite subset B such that B is in P, that is, an infinite subset that is intrinsically easy; such a set B may be called a P-approximation of A. If a set not in P has no P-approximation, then this set is P-immune. It is clear that a set A is P-immune if and only if A is a proper

This research was supported in part by the National Science Foundation under grants DCR 83-12472 and CCR 86-11980.

Authors' present addresses: R. V. Book, Department of Mathematics, University of California, Santa Barbara, CA 93106; D.-Z. Du, Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, People's Republic of China.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission. © 1987 ACM 0004-5411/87/0700-0718 \$01.50

Journal of the Association for Computing Machinery, Vol. 34, No. 3, July 1987, pp. 718-730.

complexity core for A. If a set is P-immune and its complement is also P-immune, then the set is *bi-immune* for P. Balcázar and Schöning [1] showed that a set A is bi-immune for P if and only if Σ^* is a complexity core for A.

Immunity in computational complexity theory has been studied recently in the context of relativized complexity classes (see, for example, [9] or [22]). The use of that notion in the context of the present paper stems to earlier work of Berman [3] and of Ko and Moore [10] but it appears that Balcázar and Schöning [1] were the first to link the notion of immunity with that of complexity core. This latter work has led to a comprehensive study by Balcázar, Du, Isakowitz, Orponen, Russo, and Schöning [1, 5, 6, 14–19] of P-immune sets, complexity cores, the lattice of complexity cores of a set, sparse complexity cores, and other topics.

The purpose of the present paper is to extend some of the principal results of [1, 5, 6, 14–19] to a more general setting. Even et al. [7] have considered the existence of generalized complexity cores with respect to classes such as R, ZPP (= $R \cap co$ -R), NP, and NP $\cap co$ -NP. These generalized complexity cores are called *hard cores* with respect to the appropriate class. We follow this terminology but we define the basic notions in a machine-independent, measure-independent manner that allows us to capture the fundamental ideas in the previous studies as well as to present conceptually simple proofs (using only elementary methods).

The results are presented in three sections. The topic of the Section 2 is the existence of infinite proper hard cores. The basic result (Theorem 2.7) gives conditions under which a set not in a class \mathscr{C} must possess an infinite proper hard core with respect to \mathscr{C} . Conditions on the very general classes \mathscr{C} are given and shown to be necessary. In addition, the situation in which \mathscr{C} is a recursively enumerable class of recursive sets is considered and the existence of an infinite recursive proper hard core is established for recursive sets not in \mathscr{C} (Theorem 2.10). This generalizes the results of Lynch [12] and of Even et al. [7].

In Section 3 the density of proper hard cores is studied. Motivated by results of Orponen and Schöning [16] characterizing those sets not in P having nonsparse proper complexity cores, we establish a density theorem (Theorem 3.3) for the proper hard cores with respect to a recursively enumerable class of recursive sets.

Section 4 contains some remarks on related topics. An existence theorem for "partially immune" hard cores is established and the complexity of hard cores is discussed.

We assume that the reader is familiar with the properties of complexity classes such as P, NP, PSPACE, and the union PH of the classes in the polynomial time hierarchy. Unless explicitly noted to the contrary, all strings are taken over the alphabet $\Sigma = \{0, 1\}$. The set of all such strings is denoted by Σ^* . For a set A the complement of A is denoted \overline{A} . For a class \mathcal{C} of sets, co- \mathcal{C} denotes the class $\{\overline{C} \mid C \in \mathcal{C}\}$. For a string x, the length of x is denoted by |x|. The cardinality of a set S is denoted by ||S||. The set of natural numbers is denoted by \mathcal{N} . For a machine M, the set of strings accepted by M is denoted by L(M).

2. The Existence of Hard Cores

We begin by recalling Lynch's [12] definition of "complexity core."

A set H is a complexity core for a recursive set A if, for every polynomial p and every deterministic Turing machine M recognizing A, M's running time on x exceeds p(|x|) for all but finitely many $x \in H$. A complexity core H for A is proper if $H \subseteq A$. Du et al. [6] have shown that an infinite subset H is an infinite proper complexity core for A if and only if for every set $D \subseteq A$, $D \in P$ implies $D \cap H$ is finite. The following lemma improves this slightly.

LEMMA 2.1. An infinite set H is a complexity core for a recursive set A if and only if for every set $D \in P$ with $D \subseteq A$ or $D \subseteq \overline{A}$, $D \cap H$ is finite.

PROOF. Let H be a complexity core of A. If $D \subseteq A$ and $D \in P$, construct a procedure that behaves as follows:

input x;

if $x \in D$ then accept

else if $x \in A$ then accept else reject.

This procedure recognizes A. Since $D \in P$, there exists a polynomial p such that this procedure accepts every $x \in D$ in time p(|x|). Since H is a complexity core of A, this means that $H \cap D$ is finite. A similar proof can be used in case $D \subseteq \overline{A}$.

For the converse, suppose that H is not a complexity core of A. Thus, there is a machine M that recognizes A and a polynomial p such that M's running time on x is less than p(|x|) for infinitely many $x \in H$. Define $D = \{x \mid M \text{ accepts } x \text{ in time } p(|x|)\}$ and $E = \{x \mid M \text{ rejects } x \text{ in time } p(|x|)\}$. Then, $D \in P$, $E \in P$, $D \subseteq A$, and $E \subseteq \overline{A}$. Clearly, either $D \cap H$ or $E \cap H$ is infinite. \Box

This result provides motivation for our general definition.

Definition 2.2. Let \mathscr{C} be a class of sets. For any set A, let \mathscr{C}_A denote $\{C \in \mathscr{C} \mid C \subseteq A\}$. A set H is a hard core for A with respect to \mathscr{C} if for every $C \in \mathscr{C}_A$, $C \cap H$ is finite. If, in addition, H is a subset of A, then H is a proper hard core.

Before developing any properties of hard cores, we interpret this definition of some well-known complexity classes.

PROPOSITION 2.3. Let n > 0, let A be an infinite set, and let $H \subseteq A$. The set H is an infinite proper hard core for A with respect to Σ_n^P if and only if for every polynomial q and every nondeterministic oracle machine M that recognizes A relative to a set in Σ_{n-1}^P , say, $B \in \Sigma_{n-1}^P$ and L(M, B) = A, the length of the shortest accepting computation on M relative to B on x is less than or equal to q(|x|) for only finitely many $x \in H$.

PROOF. Suppose there exist a polynomial p, a nondeterministic oracle machine M, and a set $B \in \sum_{n=1}^{P}$ such that L(M, B) = A, and for infinitely many $x \in H$, there is an accepting computation of M on x relative to B of length at most q(|x|). Let C be the set of all such x. Clearly, $C \subseteq H \subseteq A$, $C \in \sum_{n=1}^{P}$, and C is infinite. Thus, $C \cap H$ is infinite so that H is not a proper hard core for A relative to $\sum_{n=1}^{P} C$.

Now suppose that H is not an infinite proper hard core for A relative to Σ_n^p . Then for some $C \in \Sigma_n^p$, $C \subseteq A$, and $C \cap H$ is infinite. Let M_1 be a nondeterministic polynomial time-bounded oracle machine and B_1 a set in Σ_{n-1}^p such that $L(M_1, B_1) = C$; let p be a polynomial that bounds M_1 's running time. Let M_2 be a nondeterministic oracle machine that recognizes A relative to some set $B_2 \in \Sigma_{n-1}^p$. Construct M_3 such that on input x, M_3 simulates a computation of M_1 on x relative to B_1 for p(|x|) steps. If this computation is accepting, then M_3 halts and accepts; otherwise, M_3 simulates a computation of M_2 on x relative to B_2 . Since $L(M_1, B_1) = C$, $C \subseteq A$, and $L(M_2, B_2) = A$, we see that $L(M_3, B_1 \oplus B_2) = A$; also, $B_1 \oplus B_2 \in \Sigma_{n-1}^p$. Since $C \cap H$ is infinite, there are infinitely many $x \in H$ such that the length of the shortest accepting computation of M_3 on x relative to $B_1 \oplus B_2$ is at most p(|x|). \Box Similar arguments yield the somewhat weaker facts regarding other classes.

The class R (random polynomial time) is the class of sets L such that for some nondeterministic polynomial time-bounded machine M, L(M) = L and for all $x \in L(M)$.

pr(there is an accepting computation of M on x with at most q(|x|) steps) $\geq \frac{1}{2}$.

The class ZPP is $R \cap co-R$.

PROPOSITION 2.4. Let A and H be infinite sets. Suppose that for every polynomial q and every nondeterministic Turing machine M recognizing A,

pr(there is an accepting computation of M on x with at most q(|x|) steps) $\geq \frac{1}{2}$

for only finitely many $x \in H$. Then H is an infinite hard core for A with respect to R.

Let M be a nondeterministic oracle machine that is endowed with designated accepted states and with designated rejecting states. We say that M is *strong* if, relative to each oracle set, for each input string x, there is either at least one accepting computation of M on x or at least one rejecting computation of M on x, but not both accepting and rejecting computations on x.

PROPOSITION 2.5. Let A be an infinite set and let H be an infinite subset of A. Suppose that for every polynomial q and every strong nondeterministic Turing machine M recognizing A, the minimum running time of M to decide whether x is in A is bounded above by q(|x|) for only finitely many $x \in H$. Then H is a proper hard core for A with respect to $NP \cap co-NP$.

PROPOSITION 2.6. Let A be an infinite set and let H be an infinite subset of A. Suppose that for every strong nondeterministic Turing machine M recognizing A,

pr(there is an accepting computation of N on x with at most q(|x|) steps) $\geq \frac{1}{2}$

for only finitely many $x \in H$. Then H is a proper hard core for A with respect to ZPP.

Now we establish a general existence theorem.

THEOREM 2.7. Let \mathscr{C} be a countable class of sets. An infinite set A has an infinite proper hard core with respect to \mathscr{C} if and only if A is not a finite union of a finite set and some sets in \mathscr{C} .

PROOF. Let A be an infinite set that is not a finite union of a finite set and some sets in \mathscr{C} . Let \mathscr{C}_A denote $\{C \in \mathscr{C} \mid C \subseteq A\}$. Let B be the union of all sets in \mathscr{C}_A . If A - B is infinite, then A - B is an infinite proper hard core of A with respect to \mathscr{C} , so assume A - B is infinite. Since \mathscr{C} is countable, \mathscr{C}_A is countable so that \mathscr{C}_A can be enumerated as C_1, C_2, \ldots . Let $D_k = \bigcup \{C_i \mid 1 \le i \le k\}$. Then, (i) for every k, $D_k \subseteq D_{k+1}$, and (ii) for every $C \in \mathscr{C}_A$, there exists k such that $C \subseteq D_k$.

We claim that for infinitely many k, $D_k \neq D_{k+1}$. For otherwise, there exists j such that for every $k \ge j$, $D_k = D_j$ so that $A = (A - B) \cup B = (A - B) \cup D_j$, contradicting the hypothesis that A is not the finite union of some finite set and some sets in \mathscr{C} .

For each k such that $D_k \neq D_{k+1}$, choose $a_k \in D_{k+1} - D_k$ and let H be the set of all such chosen a_k s, that is, $H = \{a_k \mid D_k \neq D_{k+1}\}$. Then H is infinite and for every $k, H \cap D_k$ is finite. By (ii), this means that for every $C \in \mathscr{C}_A, H \cap C$ is finite. Thus, H is an infinite proper hard core for A with respect to \mathscr{C} .

To see the converse, suppose that $A = F \cup C_1 \cup \cdots \cup C_k$ where F is a finite set and $C_i \in \mathscr{C}$ for i = 1, ..., k. Let H be an infinite subset of A. Since $H = H \cap A =$ $(H \cap F) \cup (H \cap C_1) \cup \cdots \cup (H \cap C_k)$, there is at least one infinite set among $H \cap C_1, ..., H \cap C_k$. Hence, H cannot be a proper hard core for A. \Box

We say that there is a hard-core theorem for \mathcal{C} , if every infinite set A that does not belong to \mathcal{C} has a proper infinite hard core with respect to \mathcal{C} .

THEOREM 2.8. Let \mathscr{C} be a countable class of sets. There is a hard-core theorem for \mathscr{C} if \mathscr{C} satisfies the following conditions:

- (a) if C_1 and C_2 are infinite sets in \mathcal{C} , then so is $C_1 \cup C_2$;
- (b) if C is an infinite set in \mathscr{C} and F is a finite set, then $C \cup F \in \mathscr{C}$.

PROOF. By Theorem 2.7, there is a hard-core theorem for \mathscr{C} if every set that is not in \mathscr{C} is not a union of a finite set and some sets in \mathscr{C} ; this is equivalent to the hypothesis that both (a) and (b) hold. \Box

COROLLARY 2.9. Let \mathscr{C} be a countable class of sets that is closed under finite union and finite variation. Let A be an infinite set. Suppose that for every infinite set $H \subseteq A$, there exists $C \in \mathscr{C}_A$ such that $H \cap C$ is infinite. Then $A \in \mathscr{C}$.

Note that Theorem 2.7 is very general. Nothing is said about either A or H being recursive or about \mathcal{C} being a complexity class. Nothing is said about H being obtained effectively for A or of \mathcal{C} being effectively closed under union or under finite variation.

As applications of Theorem 2.7, consider the following examples:

- (i) Every nonarithmetic set A has a hard core with respect to the class of arithmetic sets.
- (ii) Let PH be the union of the polynomial-time hierarchy, $PH = \bigcup_{n\geq 0} \Sigma_n^P$. Suppose that PSPACE \neq PH. Every A in PSPACE – PH has a hard core with respect to PH.
- (iii) Every set A that is not regular has a hard core with respect to the class of regular sets.

Even et al. [7] have established two theorems asserting the existence of complexity cores for complexity classes specified by recursive predicates. They require that each class be closed under finite variation and that each class satisfy the parallel computation property. Since the parallel computation property implies closure under union, Theorem 2.7 applies to these classes. Furthermore, one can extend the arguments of Du et al. [6] to show that these classes are such that the proper hard cores described in [7] satisfy Definition 2.2.

Schöning [20] has considered a decomposition of the class NP: for every $n \ge 0$, define $L_n^P = \{A \in NP \mid \sum_{n=1}^{P} (A) \subseteq \sum_{n=1}^{P} \}$ and $H_n^P = \{A \in NP \mid \sum_{n=1}^{P} \subseteq \sum_{n=1}^{P} (A)\}$, and define $LH = \bigcup_{n\ge 0} L_n^P$ and $HH = \bigcup_{n\ge 0} H_n$. It turns out that $H_0^P = \{A \mid A \text{ is } \le_{m=1}^{P} \text{-complete}$ for NP}, $L_0^P = P$, and $L_1^P = NP \cap \text{ co-NP}$. Note that each L_i^P and H_i^P is closed under finite variation as are LH and HH. Although L_0^P and L_1^P are closed under union, it is not known whether any L_i^P , $i \ge 2$, or any H_j^P is closed under union. Thus, it is not known whether Theorem 2.6 can be applied to the classes L_i^P , $i \ge 2$, or the classes H_j^P . However, Theorem 2.7 can be applied to the classes LH and HH. Further, if the polynomial-time hierarchy is infinite, then there exists $A \in NP$ such that A has a hard core with respect to LH and A has a hard core with respect to HH. Similar facts hold for the decomposition of PH described by Balcázar et al. [2].

The proof of Theorem 2.7 essentially requires that the class \mathscr{C} be countable, be closed under finite union, and be closed under finite variation, that is, if \mathscr{C} is countable and \mathscr{D} is the closure of \mathscr{C} under finite union and under finite variation, then an infinite set A not in \mathscr{D} has an infinite proper hard core with respect to \mathscr{D} if and only if it has an infinite proper hard core with respect to \mathscr{C} . In Corollary 2.8 these requirements are explicitly stated. Each of these conditions is necessary as seen by the following arguments.

(i) The hypothesis of countability of the class under consideration is necessary. To see this, recall that a set is "sparse" if there is a polynomial p that for all n bounds the number of strings in the set having length at most n. It is clear that the collection of all sparse sets is uncountable but is closed under both union and finite variation. Note that every infinite subset of a proper hard core is a proper hard core itself, and every infinite set has an infinite sparse subset; thus, if a set has a proper hard core, then it has an infinite proper hard core that is sparse. Thus, if \mathscr{C} is a class that contains every sparse set, then no set A not in \mathscr{C} can have a proper hard core with respect to \mathscr{C} . Examples of such classes are the class of sets with polynomial size circuits [4, 11] and the class of sets with small generators [21, 23].

(ii) C. Wrathall (personal communication) has observed that closure under union is necessary. For, let \mathscr{C} be the class of deterministic context-free languages so that \mathscr{C} is closed under finite variation but not under finite union. Let $A = \{a^n b^n \mid n \ge 1\} \cup \{a^n b^{2n} \mid n \ge 1\}$, so that A is not in \mathscr{C} . For any infinite $H \subseteq A$, either $H \cap \{a^n b^n \mid n \ge 1\}$ is infinite or $H \cap \{a^n b^{2n} \mid n \ge 1\}$ is infinite. Since both $\{a^n b^n \mid n \ge 1\}$ and $\{a^n b^{2n} \mid n \ge 1\}$ are in \mathscr{C} , H is not a proper hard core of A.

(iii) Let \mathscr{C} be the class of all subsets of $\{00\}^*$ such that $C \in \mathscr{C}$ implies that $\{00\}^* - C$ is finite. It is clear that \mathscr{C} is countable and is closed under union, but \mathscr{C} is not closed under finite variation. Let $A = \{0, 1\} \cup \{00\}^*$ so that A is not in \mathscr{C} . Note that for every infinite subset B of A and every set $C \in \mathscr{C}$, $B \cap C$ is infinite. Hence, A has no proper hard core with respect to \mathscr{C} .

Up to this point we have only considered the existence of proper cores. In general the cores need not be recursive sets. Now we turn to the problem of showing the existence of infinite proper hard cores that are recursive sets.

THEOREM 2.10. Let \mathscr{C} be a recursively enumerable class of recursive sets that is closed under finite union and finite variation. Any infinite recursive set not in \mathscr{C} has an infinite recursive proper hard core with respect to \mathscr{C} .

PROOF. Let a_1, a_2, \ldots be some recursive enumeration of Σ^* , say an enumeration in lexicographic order. Let C_0, C_1, \ldots be an effective enumeration of \mathcal{C} . Let A be an infinite recursive set not in \mathcal{C} .

We construct a set H in stages.

```
Stage 0. m := 0.
```

```
Stage n + 1
```

```
for uncanceled i \le m do

if a_n \in \overline{A} and a_n \in C_i then cancel i

else

if a_n \in A and for all uncanceled i \le m, a_n \in \overline{C_i}

then H := H \cup \{a_n\} and m := m + 1

else m := m + 1.
```

It is clear that H is recursive and $H \subseteq A$. Let us show that H is infinite. If H is finite, there exist m_0 and n_0 such that $n \ge n_0$ and $a_n \in A$ imply $a_n \in C_i$ for some uncanceled $i \le m_0$. Now if any $i \le m_0$ with $C_i \in \mathscr{C}_A$, the index *i* will eventually be canceled so, without loss of generality, we can assume that n_0 is such that in step $n \ge n_0$, for any uncanceled $i \le m_0$, $C_i \in \mathscr{C}_A$. Let $U = \{C_i \mid i \le m_0 \text{ and } i \text{ is uncanceled}\}$. Then, $U \subseteq A$ and $||A - U|| \le n_0$. Since \mathscr{C} is closed under finite union and under finite variation, $A \in \mathscr{C}$, contradicting the hypothesis.

Since H is infinite, m goes to infinity as n goes to infinity. Hence, for any $C_i \in \mathscr{C}_A, C_i \cap H$ is finite.

Thus, H is an infinite proper hard core for A with respect to \mathscr{C} and H is a recursive set. \Box

3. Density of Hard Cores

Here we consider the density of hard cores. In Section 2 we noted that if set A has a proper hard core H with respect to a class \mathscr{C} , then every infinite sparse subset of H is also a proper hard core for A. Under what conditions do there exist proper hard cores that are not sparse? To study questions such as this, it is appropriate to review some definitions.

For any set A, the census of A is defined as $census_A(n) = || \{x \in A \mid |x| \le n\} ||$. A set A is sparse if for some polynomial p, $census_A(n) \le p(n)$ for all n.

For a deterministic Turing machine M and a function f on the natural numbers, the set of *f*-hard inputs for M is defined to be $H(M, f) = \{x \in \Sigma^* \mid \text{the running time of M on } x \text{ is greater than } f(|x|)\}.$

Observe that a set C is a hard core for a set A with respect to P (i.e., a complexity core for A in Lynch's terms) if and only if for every machine M recognizing A and every polynomial p, C - H(M, p) is finite. Do there exist sets not in P that have nonsparse hard cores with respect to P? This question was answered in the affirmative by Balcázar and Schöning [1]. Later, Orponen and Schöning [16] characterized the class of all such sets. In order to explain that characterization, the definitions of certain classes must be reviewed.

The class APT (almost polynomial time recognizable sets) is $\{L(M) \mid M \text{ halts on all inputs and for some polynomial } p, H(M, p) \text{ is sparse}\}$. The class WAPT (weakly almost polynomial time recognizable sets) is $\{L(M) \mid M \text{ halts on all inputs and for some polynomial } p \text{ and infinitely many } n, || H(M, p) \le n || \le p(n)\}$. The class 1-APT is $\{L(M) \mid M \text{ halts on all inputs and for some polynomial } p, H(M, p) \cap L(M) \text{ is sparse}\}$. The class 1-WAPT is $\{L(M) \mid M \text{ halts on all inputs and for some polynomial } p \text{ and infinitely many } n, || (H(M, p) \cap L(M)) \le n || \le p(n)\}$.

The classes APT and WAPT were introduced by Meyer and Paterson [13]. The classes 1-APT and 1-WAPT are the asymmetric versions of these classes and were introduced by Orponen and Schöning [16] in order to study proper complexity cores. Orponen and Schöning developed the following characterization theorem: A recursive set has only sparse complexity cores if and only if it is in APT. We establish a general theorem about the density of hard cores that yield some of the known results about complexity cores as corollaries.

Notation 3.1. Let \mathscr{C} be a class of sets and let A be an infinite set not in \mathscr{C} . For any set B, let $\Gamma_A^{\mathscr{C}}(B) = \{H \subseteq B \mid H \text{ is a proper hard core of } A \text{ with respect to } \mathscr{C}\}$ and let $\Upsilon_A^{\mathscr{C}}(B) = \{H \in \Gamma_A^{\mathscr{C}}(B) \mid H \text{ is recursive}\}$. Let $\Gamma_A^{\mathscr{C}} = \Gamma_A^{\mathscr{C}}(A)$ and $\Upsilon_A^{\mathscr{C}} = \Upsilon_A^{\mathscr{C}}(A)$.

THEOREM 3.2. Let \mathscr{C} be a countable class that is closed under finite variation and under finite union. Let $\{f_k\}_{k\geq 0}$ be a nondecreasing sequence of functions on the natural numbers, that is, for all n and k, $f_k(n) \leq f_{k+1}(n)$. Let A be an infinite set not in \mathscr{C} and let $B \subseteq A$. The following are equivalent:

- (a) for every $H \in \Gamma^{\mathscr{C}}_{A}(B)$, there exists k such that $\operatorname{census}_{H}(n) \leq f_{k}(n)$ for all sufficiently large n;
- (b) either
 - (i) $\mathscr{C}_A = \emptyset$ and there exists k such that $\operatorname{census}_B(n) \le f_k(n)$ for all sufficiently large n; or
 - (ii) there exist $C \in \mathscr{C}_A$ and k such that $census_{B-C}(n) \le f_k(n)$ for all sufficiently large n.

PROOF. Clearly, each of (b-i) and (b-ii) implies (a).

Suppose that (a) is true and $\mathscr{C}_A = \emptyset$. Then *B* is a hard core for *A* with respect to \mathscr{C} and there exists *k* such that census_{*B*}(*n*) $\leq f_k(n)$ for all sufficiently large *n* so that (a) implies (b-i). Suppose that $\mathscr{C}_A \neq \emptyset$. Assume contrary to (b-ii) that such a *C* and *k* do not exist, that is, for all $C \in \mathscr{C}_A$ and all *k*, census_{*B*-*C*}(*n*) $> f_k(n)$ for infinitely many *n*. Consider any sequence of sets $D_i \in \mathscr{C}$ such that (i) $D_1 \subseteq D_2 \subseteq \ldots$, (ii) for every $C \in \mathscr{C}$ there exists *k* such that $C \subseteq D_k$, and (iii) $\cup \{D_k \mid k > 0\} = A$. Then for any *k* and *i*, census_{*B*-*Dk*}(*n*) $> f_i(n)$ for infinitely many *n*. Thus, for each k > 0 there exist n_k with $n_1 < n_2 < \cdots$ such that census_{*B*-*Dk*}(n_k) $> f_k(n_k)$. Define $H = \bigcup_{k \ge 1} \{x \in B - D_k \mid |x| \le n_k\}$. Then for every *k*, $H \cap D_k$ is finite. Hence, *H* is a hard core of *A* with respect to \mathscr{C} and $H \subseteq B \subseteq A$. Further, for all *k*, census_{*H*}(n_k) \ge census_{*B*-*Dk*}(n_k) $> f_k(n_k)$. Hence, for any k_0 , $k > k_0$ implies census_{*H*}(n_k) $> f_k(n_k) \ge f_{k_0}(n_k)$. Thus, (a) does not hold. This means that if (a) is true and $\mathscr{C}_A \neq \emptyset$, then (b-ii) holds. \Box

Consider the situation that the appropriate sets are recursive.

THEOREM 3.3. Let \mathscr{C} be a recursively enumerable class of recursive sets that is closed under finite variation and under finite union. Let $\{f_k\}_{k\geq 0}$ be a nondecreasing sequence of recursive functions on the natural numbers, that is, for all n and k, $f_k(n) \leq f_{k+1}(n)$. Let A be an infinite recursive set not in \mathscr{C} and B a recursive subset of A. The following are equivalent:

- (a) for every $H \in \Upsilon_A(B)$, there exists k such that $\operatorname{census}_H(n) \leq f_k(n)$ for all sufficiently large n;
- (b) either
 - (i) $\mathcal{C}_A = \emptyset$ and there exists k such that $census_B(n) \le f_k(n)$ for all sufficiently large n, or
 - (ii) there exist $C \in \mathscr{C}_A$ and k such that $census_{B-C}(n) \leq f_k(n)$ for all sufficiently large n.

PROOF. Clearly, each of (b-i) and (b-ii) implies (a).

Suppose that (a) is true and $\mathscr{C}_A = \emptyset$. Then *B* is a hard core for *A* with respect to \mathscr{C} and there exists *k* such that $\operatorname{census}_B(n) \leq f_k(n)$ for all sufficiently large *n* so that (a) implies (b-i). Suppose that (a) is true and that $\mathscr{C}_A \neq \emptyset$. Assume contrary to (b-ii) that such a *C* and *k* do not exist, that is, for all $C \in \mathscr{C}_A$ and all *k*, $\operatorname{census}_{B-C}(n) > f_k(n)$ for infinitely many *n*.

Let C_1, C_2, \ldots be an enumeration of \mathscr{C} . Consider the following construction of a set H.

Stage 0. m := 0. Stage n + 1. for uncanceled $i \le m$ do if $a_n \in \overline{A}$ and $a_n \in C_i$ then cancel i; if $a_n \in B$ and for all uncanceled $i \le m, a_n \in \overline{C}_i$ then $H := H \cup \{a_n\}$; if census_H($|a_n|$) $\ge f_k(|a_n|)$ then k := k + 1 and m := m + 1; end.

It is clear that H is recursive and $H \subseteq A$. Let us show that m goes to infinity. If m does does not go to infinity, neither does k so that there exist n', m', and k' such that n > n', $a_n \in B$, and $a_n \in \overline{C}_i$ for all uncanceled $i \le m'$ imply $a_n \in H$, and n > n' implies census_H($|a_n|$) $< f_{k'}(|a_n|)$. Now for any $i \le m'$ with C_i not in $\mathcal{C}_A = \{C \in \mathcal{C} \mid C \subseteq A\}$, the index i will eventually be canceled so without loss of generality, we can assume that n' is such that for every n > n', for any $i \le m'$ that is uncanceled at stage n, $C_i \in \mathcal{C}_A$. Let $U = \bigcup \{C_i \mid i \le m' \text{ and } i \text{ is}$ uncanceled}. Then $U \subseteq A$ and census_{B-U}(n) $< f_{k'}(n)$ for all $n > |a_{n'}|$. Since \mathcal{C} is closed under finite union, $U \in \mathcal{C}$. Thus, $U \in \mathcal{C}_A$ and for all sufficiently large n, census_{B-U}(n) $< f_{k'}(n)$, contradicting the hypothesis that (b-ii) is false.

Since *m* goes to infinity, *H* is infinite and *n* goes to infinity. Hence, for any $C_i \in \mathscr{C}_A, C_i \cap H$ is finite. Thus $H \in \Upsilon^{\mathscr{C}}_A(B)$. By (a) and the choice of $\{f_k\}_{k\geq 0}$ as an increasing sequence, there exist k'' and n'' such that k > k'' and n > n'' imply census_H $(n) \leq f_k(n)$. It follows that *m* cannot go to infinity, a contradiction. \Box

As corollaries we obtain a number of results due to Orponen and Schöning [16].

COROLLARY 3.4. A recursive set A is in 1-APT if and only if for every $H \in \Gamma_A^P$, $H \in 1$ -APT.

PROOF. Let $\mathscr{C} = P$, $f_k(n) = n^k + k$ for all n and k, and B = A. Note that $\emptyset \in P$ so that $\mathscr{C}_A \neq \emptyset$. Now apply Theorem 3.2. \Box

Interpreting Corollary 3.4, we have the following fact.

COROLLARY 3.5. A recursive set A has a nonsparse proper (recursive) complexity core if and only if A is not in 1-APT.

Since APT = 1-APT \cap co-1-APT, Corollary 3.5 yields the following important result.

COROLLARY 3.6. A recursive set A has a nonsparse (recursive) complexity core if and only if A is not in APT.

Next we generalize a result of Du et al. [6], which becomes a useful lemma in the present development.

COROLLARY 3.7. Let \mathscr{C} be a countable class that is closed under finite union and finite variation. Let A be an infinite set not in \mathscr{C} and let $B \subseteq A$. The following are equivalent:

- (a) for all $H \in \Gamma_A^{\mathscr{C}}$, $B \cap H$ is finite;
- (b) either
 - (i) $\mathscr{C}_A = \emptyset$ and B is finite, or
 - (ii) there exists $C \in \mathscr{C}_A$ such that B C is finite.

PROOF. That each of (b-i) and (b-ii) implies (a) is trivial. Suppose that (a) is true. For every k let $f_k(n) = k$. For any $H \in \Gamma_A^{\mathscr{C}}(B)$, H is finite since $H \subseteq B$ and $H \cap B$ is finite. Thus, there exists k such that $\operatorname{census}_H(n) \leq f_k(n)$ for all n. Using Theorem 3.2, note that \mathscr{C}_A being empty implies that there exist k such that $\operatorname{census}_B(n) \leq f_k(n) = k$ for all sufficiently large n. Hence, B is finite. Also using Theorem 3.2, \mathscr{C}_A not empty implies that there exist $C \in \mathscr{C}_A$ and k such that $\operatorname{census}_{B-C}(n) \leq f_k(n) = k$ for all sufficiently large n so that B - C is finite. \Box

Let A and B be two sets. Define $A \approx B$ if A - B and B - A are finite. Clearly, \approx is an equivalence relation on sets. We write \tilde{A} for the equivalence class of A. If Ω is a class of sets, we write either $\tilde{\Omega}$ or $\approx(\Omega)$ for $\{\tilde{A} \mid A \in \Omega\}$. We sometimes denote the equivalence class \tilde{A} by $\approx(A)$.

Orponen [14, 15] has shown that $\approx (\Gamma_A^P) = \approx (\Gamma_B^P)$ if and only if $A - B \in P$ and $B - A \in P$. We generalize this result to other classes in the following way.

COROLLARY 3.8. Let \mathscr{C} be as in Theorem 3.3. The class \mathscr{C} is closed under relative complement (i.e., if $C_1, C_2 \in \mathscr{C}$, then $C_1 - C_2 \in \mathscr{C}$) if and only if for every two sets A, B, the following statements are equivalent:

(a) $\approx (\Upsilon_A^{\mathscr{C}}) = \approx (\Upsilon_B^{\mathscr{C}});$

(b) both A - B and B - A are in \mathscr{C} .

PROOF. First, suppose that \mathscr{C} is closed under relative complement. For any sets A, B, if both A - B and B - A are in \mathscr{C} , then it follows trivially that $\approx(\Upsilon_A^{\mathscr{C}}) = \approx(\Upsilon_B^{\mathscr{C}})$. Thus, suppose that for two sets A and B, $\approx(\Upsilon_A^{\mathscr{C}}) = \approx(\Upsilon_B^{\mathscr{C}})$. Let $C \in \mathscr{C}$. Since \mathscr{C} is closed under relative complement, $\mathscr{O} = C - C$ is in \mathscr{C} so that for every set D, $\mathscr{O} \in \mathscr{C}_D$. Now $\approx(\Upsilon_A^{\mathscr{C}}) = \approx(\Upsilon_B^{\mathscr{C}})$ implies that $(A - B) \cap H$ is finite for every $H \in \Upsilon_A^{\mathscr{C}}$. By Corollary 3.7 this means that there exists $C_1 \in \mathscr{C}_A$ such that $(C_1 - (A - B)) \cap H \subseteq C_1 \cap H$ and $C_1 \cap H$ is finite. Hence, for any $H \in \Upsilon_A^{\mathscr{C}}$, $(C_1 - (A - B)) \cap H$ is finite. Again applying Corollary 3.6, this means that there exists $C_2 \in \mathscr{C}_B$ such that $C_1 - (A - B) \subseteq C_2$. But $(A - B) \cap C_2 = \mathscr{O}$ so that $A - B = C_1 - C_2 \in \mathscr{C}$ since \mathscr{C} is closed under relative complement. Similarly, $B - A \in \mathscr{C}$.

Second, suppose for every two sets $A, B, \approx(\Upsilon_A^{\mathscr{C}}) = \approx(\Upsilon_B^{\mathscr{C}})$ if and only if both A - B and B - A are in \mathscr{C} . We must show that \mathscr{C} is closed under relative complement. Let $C_1, C_2 \in \mathscr{C}$. Let $C_3 \supseteq C_1$ be such that $C_3 - C_1$ is finite. Then $C_3 \in \mathscr{C}$ and $\approx(\Upsilon_{C_1}^{\mathscr{C}}) = \approx(\Upsilon_{C_3}^{\mathscr{C}})$, so that $\mathscr{O} = C_1 - C_3$ and $C_3 - C_1$ are in \mathscr{C} . Now $C_1 = C_1 - \mathscr{O}$ and $\mathscr{O} = \mathscr{O} - C_1$ are in \mathscr{C} so $\approx(\Upsilon_{C_1}^{\mathscr{C}}) = \approx(\Upsilon_{\mathscr{O}}^{\mathscr{C}})$. Similarly, $\approx(\Gamma_{C_2}^{\mathscr{C}}) = \approx(\Gamma_{\mathscr{O}}^{\mathscr{C}})$ so that $\approx(\Gamma_{C_1}^{\mathscr{C}}) = \approx(\Gamma_{C_2}^{\mathscr{C}})$. Thus, $C_1 - C_2$ and $C_2 - C_1$ are in \mathscr{C} . \Box

COROLLARY 3.9. NP = co-NP if and only if for every two sets A, B, the following conditions are equivalent:

(a) $\approx (\Upsilon_A^{NP}) = \approx (\Upsilon_B^{NP});$

(b) both A - B and B - A are in NP.

Using arguments similar to those of the proof of Corollary 3.8, we have the following fact.

COROLLARY 3.10. For every two sets A, B, both A - B and B - A are in APT if and only if $\{H - S, S - H \mid H \in \Upsilon_A^P, S \text{ is sparse}\} = \{H - S, S - H \mid H \in \Upsilon_B^P, S \text{ is sparse}\}$.

Theorem 3.2 can be altered in the following way with the result again being true: uniformly replace "for all sufficiently large n" with "for infinitely many n." The

proof of the new result is essentially the same as the proof of Theorem 3.2; the details are left to the reader. This result yields the fact that a recursive set A is in 1-WAPT if and only if every proper recursive complexity core of A is in 1-WAPT.

The applications of the principal theorems on the density of hard cores have focused on the notion of sparse sets. But it is clear that one can apply these theorems to situations in which density bounds other than polynomials are considered, for example, subexponential bounds.

4. Additional Remarks

Recall that a set is P-immune if it has no infinite subset in P. Similarly, for a collection \mathscr{C} of sets, we can define a set to be \mathscr{C} -immune if it has no infinite subset in \mathscr{C} . From the results in Section 2 it is clear that if H is an infinite proper hard core with respect to \mathscr{C} for some set A not in \mathscr{C} , then H is \mathscr{C} -immune.

There is a variation on this notion that has arisen in the study of complexity measures for public-key cryptosystems [8]. If S is \mathscr{C} -immune, then the only subsets of S that can be in \mathscr{C} are the finite subsets. What we consider here is the situation where the notion of "sparse subset" replaces that of "finite subset."

Definition 4.1. Let \mathscr{C} be a collection of sets. A set is partially \mathscr{C} -immune if it has no nonsparse subset in \mathscr{C} .

Of course, we want partially \mathscr{C} -immune sets to be nonsparse just as we want \mathscr{C} -immune sets to be infinite. We establish an existence theorem for partially \mathscr{C} -immune sets where the proof depends on the density theorem for infinite proper hard cores (Theorem 3.2).

Notation 4.2. Let \mathscr{C}_1 and \mathscr{C}_2 be collections of sets. Let $\mathscr{C}_1 \wedge \mathscr{C}_2$ denote the collection $\{C_1 \cap C_2 \mid C_1 \in \mathscr{C}_1 \text{ and } C_2 \in \mathscr{C}_2\}$ and let $\mathscr{C}_1 \vee \mathscr{C}_2$ denote the collection $\{C_1 \cup C_2 \mid C_1 \in \mathscr{C}_1 \text{ and } C_2 \in \mathscr{C}_2\}$.

THEOREM 4.3. Let \mathscr{C} be a collection of recursive sets such that \emptyset is in \mathscr{C} and let \mathscr{S} be the collection of all recursive sparse sets. Let \mathscr{B} denote the smallest class containing every set in (($\mathscr{C} \land co-\mathscr{S}$) $\lor \mathscr{S}$) and is closed under finite union and finite variation. For any infinite recursive set A not in \mathscr{B} , there is an infinite subset H of A such that (i) H is nonsparse and (ii) H is partially \mathscr{C} -immune.

PROOF. For each integer k > 0, let $f_k(n) = n^k + k$. Recall that A is a subset of itself. Now $\mathscr{B}_A \neq \emptyset$ and there exist no $B \in \mathscr{B}_A$ such that for some k, census_{A-B}(n) $\leq f_k(n)$ for all but finitely many n. Thus, by applying Theorem 3.2 we see that there is an infinite subset H of A such that H is nonsparse and H is a proper hard core of A with respect to \mathscr{B} . Since $\mathscr{C} \land co - \mathscr{S} \subseteq \mathscr{B}$, H is also a proper hard core of A with respect to $\mathscr{C} \land co - \mathscr{S}$. Thus, for every $C \in \mathscr{C}$ and every $\overline{S} \in co - \mathscr{S}$, $(C \cap \overline{S}) \cap H$ is finite so that all but finitely many elements of $C \cap H$ are in S. This means that for every $C \in \mathscr{C}$, $C \cap H$ is sparse. \Box

In Theorem 4.3, one considers a nonsparse set A that is in a class \mathscr{D} but not in a class \mathscr{C} . This class \mathscr{S} is taken to be the class of sparse sets in \mathscr{D} . It is assumed that the smallest class containing all the sets in $((\mathscr{C} \land \text{co-}\mathscr{S}) \lor \mathscr{S})$ and closed under finite union and under finite variation is properly included in \mathscr{D} . When dealing with complexity classes it appears to be more convenient to assume that the Boolean closure of $\mathscr{C} \cup \mathscr{S}$ is properly included in \mathscr{D} . For example, let $\mathscr{C} = P$ and let \mathscr{S} be the collection of sparse sets in NP \cap co-NP. Assume that NP \cap co-NP is not the Boolean closure of $\mathscr{C} \cup \mathscr{S}$. If A is a set in (NP \cap co-NP) – Boolean closure ($\mathscr{C} \cap \mathscr{S}$), then there exists a proper hard core *H* of *A* that is nonsparse and partially P-immune. Thus, Theorem 4.3 and its proof yield some information about some specific cases regarding well-studied complexity classes.

Although Theorem 4.3 shows the existence of partially immune sets in the context of hard cores, it does not provide insight into some of the specific problems that arise in the context of complexity measures for public-key cryptosystems. For example, one might wish that the set A itself were partially \mathcal{C} -immune; the theorem tells us nothing about this. In addition, there is no information given in Theorem 4.3 about the complexity of the set H and it is clear that such information would be very desirable.

Orponen and Schöning [16] have shown that for every recursive set not in P and every running time T that majorizes every polynomial, there is an infinite complexity core H for A such that H is recognizable by an algorithm with running time bounded above by T. Du [5] has shown that (assuming $P \neq NP$) every NP-complete set has an infinite proper hard core that can be recognized in exponential time. We would like to have a general theorem about the complexity of hard cores but this appears to be out of reach. What can be done is to consider "standard" complexity classes specified by machines with bounded computational resources such as time or space. For example, the following result can be obtained by using the methods of Orponen and Schöning [16].

THEOREM 4.4. Let \mathscr{F} be a recursively enumerable collection of recursive functions that are running times. Let $\mathscr{C}(\mathscr{F})$ be the class of sets recognized by algorithms that run within time bounds that are in \mathscr{F} . Let T be a running time that majorizes every function in \mathscr{F} . If A is a recursive set not in $\mathscr{C}(\mathscr{F})$, then A has an infinite proper hard core that can be recognized in time T.

Theorem 4.4 shows that the existence of a proper hard core whose complexity depends on the underlying class $\mathscr{C}(\mathscr{F})$ but not on the set for which it is a core. Thus, the complexity of a core seems to depend on the cost of "diagonalizing" out of the underlying class. But recall that every subset of a core is again a core so that this remark only pertains to cores obtained by certain procedures.

Finally, we note that one of the referees has observed that several results have counterparts in recursive function theory. In addition, it has been pointed out by others that some of the topics that have been studied as part of the investigation of the structure of complexity cores may be of interest to those who work in recursive function theory. In particular, the lattice-theoretic structure of the collection of cores of an arbitrary set has been classified by Orponen [14, 15]. Thus, in a sequel to the present paper, we develop a number of results about the structure of hard cores.

REFERENCES

- 1. BALCÁZAR, J. AND SCHÖNING, U. Bi-immunity for complexity classes. Math. Syst. Theory 18 (1985), 1-10.
- 2. BALCÁZAR, J., BOOK, R., AND SCHÖNING, U. Sparse set, lowness, and highness. SIAM J. Comput. 15 (1986), 739-747.
- 3. BERMAN, L. On the structure of complete sets: Almost-everywhere complexity and infinitely often speedup. In *Proceedings of the 17th IEEE Symposium on Foundations of Computer Science*. IEEE, New York, 1976, 76–80.
- 4. BERMAN, L., AND HARTMANIS, J. On isomorphism and density of NP and other complete sets. SIAM J. Comput. 6 (1977), 305-322.
- 5. DU, D.-Z. Generalized complexity cores and levelability of intractable sets. Ph.D. dissertation. Univ. of California, Santa Barbara, Calif., 1985.

- 6. DU, D.-Z., ISAKOWITZ, T., AND RUSSO, D. Structural properties of complexity cores, submitted for publication.
- 7. EVEN, S., SELMAN, A., AND YACOBI, Y. Hard core theorems for complexity classes. J. ACM 35, 1 (Jan. 1985), 205-217.
- GROLLMAN, J., AND SELMAN, A. Complexity measures of public-key cryptosystems. In Proceedings of the 15th IEEE Symposium on Foundations of Computer Science. IEEE, New York, 1984, pp. 495–503.
- 9. HOMER, S., AND MAASS, W. Oracle dependent properties of the lattice of NP sets. *Theoret.* Comput. Sci. 24 (1983), 279-289.
- 10. KO, K., AND MOORE, D. Completeness, approximation, and density. SIAM J. Comput. 10 (1981), 787-796.
- 11. KO, K., AND SCHÖNING, U. On circuit-size complexity and the low hierarchy in NP. SIAM J. Comput. 14 (1984), 41-51.
- 12. LYNCH, N. On reducibility to complex or sparse sets. J. ACM 22, 3 (July 1975), 341-345.
- 13. MEYER, A., AND PATERSON, M. With what frequency are apparently intractable problems difficult? Tech. Rep. TM-126, Massachusetts Institute of Technology, Cambridge, Mass., 1979.
- 14. ORPONEN, P. The structure of polynomial complexity cores. Ph.D. dissertation. Univ. of Helsinki, Helsinki, Finland, 1986.
- 15. ORPONEN, P. A classification of complexity core lattices. *Theoret. Comput. Sci.* 47 (1986), 121-130.
- ORPONEN, P., AND SCHÖNING, U. The density and complexity of polynomial cores for intractable sets. Inf. Control 70 (1986), 54-68.
- ORPONEN, P., RUSSO, D., AND SCHÖNING, U. Optimal approximations and polynomially levelable sets. SIAM J. Comput. 15 (1986), 399-408.
- RUSSO, D. Structural properties of complexity classes. Ph.D. dissertation. Univ. of California, Santa Barbara, Calif., 1985.
- 19. RUSSO, D., AND ORPONEN, P. On P-subset structures. Submitted for publication.
- 20. SCHÖNING, U. A low and a high hierarchy within NP. J. Comput. Syst. Sci. 27 (1983), 14-28.
- 21. SCHÖNING, U. A note on small generators. Theoret. Comput. Sci. 34 (1984), 337-341.
- 22. SCHÖNING, U., AND BOOK, R. Immunity, relativizations, and nondeterminism. SIAM J. Comput. 13 (1984), 329-337.
- YAP, C. Some consequences of non-uniform conditions on uniform classes. Theoret. Comput. Sci. 26 (1983), 287-300.

RECEIVED AUGUST 1984; REVISED DECEMBER 1986; ACCEPTED JANUARY 1987