

Investigations Concerning the Structure of Complete Sets

Eric Allender

Abstract. This paper will discuss developments bearing on three related research directions where Somenath Biswas has made pioneering contributions:

- Isomorphism of Complete Sets
- Creative Sets
- Universal Relations

Some open questions in each of these directions will be highlighted.

Mathematics Subject Classification (2010). Primary 68Q15.

Keywords. Berman-Hartmanis Conjecture, Isomorphism, NP-Completeness, Creativity, Universality.

1. Introduction

How many NP-complete sets are there?

Although there is a trivial and uninteresting answer to this question (namely: there is a countably infinite number of NP-complete sets), there is a large body of work investigating the proposition that in actuality there is *precisely one* NP-complete set (modulo minor encoding details).

Let us clarify what is meant by “minor encoding details”: When we consider the set SAT of satisfiable Boolean formulae, it is irrelevant if we encode formulae using round parentheses () or square ones [], or if we write variables in italic font or in bold face. Any of these choices would lead to a reasonable encoding of SAT; they all yield encodings of SAT that are equivalent in some sense.

One way to attempt to formalize this notion of “equivalence” is to say that two sets A and B , $A, B \subseteq \{0, 1\}^*$, are p-isomorphic if there is a bijection f defined on $\{0, 1\}^*$ computable and invertible in polynomial time, such that $f(A) = B$. This approach leads to the famous Berman-Hartmanis conjecture [17], which asserts that all of the sets that are NP-complete under \leq_m^p reductions are p-isomorphic.

The isomorphism conjecture(s) will be discussed in more detail in Section 2. However, a bit of background about isomorphism of complete sets is necessary here,

in order to provide a coherent overview of the current paper. The Berman-Hartmanis conjecture arose, at least in part, because of a cultural inheritance from the study of computability theory. If we accept the rough idea that NP is analogous to the class of computably-enumerable sets, and polynomial time is analogous to the class of computable functions, then the Berman-Hartmanis conjecture is analogous to the Myhill Isomorphism Theorem in computability theory, which states: All of the sets that are complete for the class of computably-enumerable sets under \leq_m -reductions are computably-isomorphic to the Halting Problem. (For expositions of this work, see [32] or [33].)

Central to the proof of the Myhill Isomorphism Theorem is the notion of a *creative* set. We postpone until Section 3 the precise definition of “creativity”, but this is an appropriate time to mention that the name was coined by Emil Post [31], who was profoundly influenced by certain consequences of Gödel’s incompleteness theorems. Post believed that there was a link between the notion of “mathematical creativity” and the fact that there is a computable function that, given a set of consistent axioms for arithmetic, will produce a true statement that cannot be proved from those axioms. Post’s definition of creativity abstracts out this property of the set of theorems provable from a list of axioms.

In the setting of recursion theory, the creative sets turn out to exactly coincide with the sets that are complete for the class of computably-enumerable sets under \leq_m -reducibility, and this is useful in proving Myhill’s Isomorphism Theorem. Thus it was natural for researchers to try to define a resource-bounded analog of creativity. But it is not entirely clear what is the best way to define such an analog. Different definitions were presented by various authors [29, 34], but the definition of *NP-creative* sets by Agrawal and Biswas [10] provides several advantages over other definitions. (For instance, all NP-creative sets are NP-complete; this is not known for some other notions.) Just as all of the NP-complete sets in Garey and Johnson [22] are p-isomorphic to SAT, so also are they all NP-creative. Although Agrawal and Biswas refrain from conjecturing that all NP-complete sets are NP-creative, we may as well consider a “creativity” version of the Berman-Hartmanis conjecture:

The Creativity Hypothesis: The class of NP-creative sets coincides with the class of sets that are NP-complete under \leq_m^p reductions.

One might at first guess that, since SAT is NP-creative, then *everything* that is p-isomorphic to SAT would also be NP-creative – but Agrawal and Biswas showed that, if this is true, then the Creativity Hypothesis is true (and hence $P \neq NP$).

Thus, NP-creativity and p-isomorphism yield two possibly-different subclasses of the NP-complete sets, each of which captures a notion of “naturalness” (in the sense that all of the currently-known “natural” examples of NP-complete sets are both creative and p-isomorphic to SAT). Neither the Creativity Hypothesis nor the Berman-Hartmanis Conjecture is known to imply the other. Section 3 discusses creativity in more detail.

The final section of this paper highlights one additional direction in which Somenath Biswas has pushed, in order to give additional insight into the structure underlying completeness. Although the Berman-Hartmanis conjecture focuses on

the NP-complete *sets*, let us not forget that much of the practical interest in NP-completeness derives from the desire to find *witnesses* for membership in an NP-complete set. That is, at a fundamental level, it is not a *set*, such as HAMILTONIAN-CIRCUIT, that is of primary interest, but rather the corresponding *relation* consisting of pairs (G, C) such that C is a Hamiltonian cycle in the graph G .

Is it possible that the (string, witness) relations for every NP-complete set are all “the same” in some sense? Note that it is not at all obvious how to formulate this sense of “sameness”. For instance, if there is a polynomial-time relation $W(x, y)$ consisting of witnesses y for string x , then there is a relation $W'(x, z)$ such that $W'(x, z)$ is true if and only if $W(x, y)$ holds, where y is the string that results by deleting every second symbol of z . These two relations both serve as witness relations for the same set in NP, but they do give different numbers of witnesses for the same string, and thus they fail to be “the same” on a fairly basic level. And yet, they do contain exactly the same information, in some intuitive sense. Agrawal and Biswas succeeded [9] in giving a useful definition of “universal relations”, in order to capture the sense in which the defining relations for all known NP-complete sets seem to be “the same”. More recently, Chaudhary, Sinha, and Biswas have adapted this notion for nondeterministic logspace [19]. This topic is explored in Section 4.

2. The Isomorphism Conjecture(s)

An outstanding survey of recent developments related to isomorphisms of complete sets is now available [5], and the reader is urged to consult that source for a more complete introduction to the topic and an in-depth discussion of the current state of the field. The discussion here will focus on describing aspects of the topic that are (1) related to the work of Somenath Biswas, or (2) related to some open questions or developments that are not mentioned in [5].

The winds of public opinion have blown back and forth, regarding the Berman-Hartmanis Conjecture. It appears to have initially been viewed as fairly plausible. One of the first published accounts questioning whether the isomorphism conjecture is true appears in the work of Joseph and Young [29]. They defined a class of NP-complete sets they named the *k-creative sets* (which will also be discussed in Section 3), and they explicitly conjectured that some *k-creative sets* are not p-isomorphic to SAT. In particular, for any one-one length-increasing function f computable in polynomial time, they defined a set K_f , and they pointed out that, if f is suitably hard to invert, then it is hard to see how K_f can be p-isomorphic to SAT. Kurtz, Mahaney, and Royer subsequently elaborated on this intuition, and formulated the *Encrypted Complete Set Conjecture*, which states that there is a one-one, length-increasing, one-way function f such that SAT and $f(\text{SAT})$ are not p-isomorphic.

Several papers were then written, all of which tended to buttress support for the Encrypted Complete Set Conjecture (all of which are discussed in the survey [5]). But then attention shifted to some interesting classes of *restricted* \leq_m^p -reductions; we will discuss some of these developments in more detail below – but the general trend of these investigations has been to weaken our confidence in the Encrypted Complete Set Conjecture. More recently, there has been a productive

series of investigations of *more powerful* classes of reductions, notably including m -reductions computed in P/poly [12, 3] and in $NP \cap \text{coNP}$ [23] (this latter class of reductions is known as SNP-reductions). As a consequence, we now know that some fairly-plausible hypotheses imply that all sets complete for NP under P/poly -reductions and SNP-reductions are P/poly -isomorphic and SNP-isomorphic, respectively. Combined with the results about restricted \leq_m^P reductions that will be discussed below, the picture that emerges is that NP -complete sets are either provably isomorphic or at least are reasonably-likely to be isomorphic, both when one considers reductions strictly *less* powerful and *more* powerful than \leq_m^P reductions. It remains to be seen whether any of these lessons ultimately shed much light on the case of \leq_m^P reductions themselves.

2.1. Restricted Reductions

It is debatable whether \leq_m^P reductions really constitute the most important class of reductions. There is a rich structure of complexity classes within P , and \leq_m^P -reducibility is essentially useless in elucidating this structure. This was the motivation for Jones et al. to introduce logspace reductions [28] – but even in that pioneering work, it was realized that a higher-precision tool was necessary, in order to investigate the structure of logspace, which is why Jones introduced what he called *log-bounded rudimentary reductions* [28]. This was several years before the modern study of circuit complexity got under way, and it took a while before it was noticed that log-bounded rudimentary reductions actually correspond to many-one reductions computed by uniform AC^0 circuits (that is, constant-depth polynomial-size families of circuits of AND and OR gates) [16]. Ultimately, AC^0 reductions have proved to be the most useful notion of reducibility for investigating subclasses of P , surpassing both NC^1 reducibility [20] and 1-L reducibility (discussed below). However, AC^0 reducibility posed greater challenges initially, and thus progress was made first with 1-L reducibility.

2.2. 1-L Reductions

1-L reductions are functions computed by logspace-bounded Turing machines that make a single pass over their input tape (from left to right). They were introduced by Hartmanis, Immerman, and Mahaney [24, 25] for many of the same reasons that had led Jones to introduce log-bounded rudimentary reductions. 1-L reductions offered the advantages of being significantly more convenient and intuitive (since the original formulation of log-bounded rudimentary reductions lacked the intuitive appeal of the AC^0 formulation). In this brief overview, we will avoid giving more detailed definitions of 1-L reductions, but it is appropriate to note that there are some differences in the formulations as presented in [24] and [25], and that some of these formulations result in a class of reductions that is not closed under composition (see [13]).

1-L reductions are easy to invert, and this fact, combined with some diagonalization techniques, enabled a proof that all sets complete for $PSPACE$ under 1-L reductions are p -isomorphic [13]. They are not isomorphic under 1-L isomorphisms [18], but they are complete under isomorphisms computable in nondeterministic logspace [26].

Agrawal and Biswas [11] succeeded in showing that, even for classes as small as deterministic logspace (and indeed, for any class that is closed under logspace reductions that produce output at most linearly-longer than the input) the sets complete under 1-L reductions are complete under one-one, length-increasing, polynomial-time invertible reductions. (Thus by [17] all such sets are p-isomorphic.) Finally, Agrawal proved that all such sets are isomorphic via reductions computed by one-way nondeterministic logspace (1-NL) machines [1] (and the same paper proves an analog of the Berman-Hartmanis conjecture for 1-NL reductions).

At this point, study of the structure of sets complete under 1-L reductions effectively stopped.¹ The major open questions had been solved. But this was merely a prelude to a much more exciting and significant development in the history of work on the isomorphism problem, focusing on reductions that are computable by constant-depth circuits. Indeed, although there are problems (such as the PARITY problem) that are computable by 1-L machines but are not computed by AC^0 circuits [21], Agrawal had shown that, for essentially all complexity classes of interest, all sets complete under 1-L reductions *are* complete under reductions computable in AC^0 [1]. And whereas all functions computable by 1-L machines are easy to invert, this is not the case for AC^0 . Thus, by considering AC^0 reductions, the research community was moving on to a richer class of complete sets, and was confronting some of the essential issues raised by the Encrypted Complete Set Conjecture.

2.3. Constant-Depth Circuits

Attention was first focused on AC^0 isomorphisms by considering a very restricted class of AC^0 reductions: projections (which are reductions computed by circuits with no gates, other than negation gates). Sets complete for NP (and other classes) under uniform projections are isomorphic under uniform AC^0 isomorphisms [15].

The logical next step was to work on extending this result from projections to NC^0 functions (that is, functions computed by constant-depth circuits with fan-in $O(1)$, so that each output bit depends on only $O(1)$ input bits – as contrasted with projections, where each output bit depends on either zero or one input bit). As part of this investigation, it was also discovered that, at least for the class NC^1 , the sets complete under AC^0 reductions are also complete under NC^0 reductions, thereby obtaining the first theorem showing that the sets complete under AC^0 reductions are all AC^0 isomorphic [6]. Subsequently, the authors were joined by Rudich, in showing that this holds not only for NC^1 , but also for NP and for most other complexity classes of interest [8].

These initial AC^0 isomorphism theorems were proved only in the non-uniform setting. After a series of intermediate steps improving the uniformity condition [7, 2], Agrawal succeeded in overcoming some daunting technical difficulties, in presenting a Dlogtime-Uniform version of the isomorphism theorem [4], which stands as one of the crowning achievements of the study of the structure of complete sets. This work not only shows that a natural re-phrasing of the Berman-Hartmanis Conjecture (in terms of AC^0 reductions and isomorphisms) is true, but also gives a convincing setting where the Encrypted Complete Set Conjecture fails (since even

¹This is not to suggest that work on 1-L computation stopped. Indeed, much of the very large body of work on *streaming algorithms* consists of the study of 1-L computation.

when f is an AC^0 function that provably cannot be inverted in AC^0 , $f(\text{SAT})$ is still AC^0 -isomorphic to SAT).

2.4. Open Questions

Again, please refer to [5] for several interesting open questions. Here are a few additional questions relating to isomorphisms, that are not discussed there.

Two important problems that are not believed to be NP-complete are Factoring and the Minimum Circuit Size Problem:

$\text{FACT} = \{(x, i, b) : \text{the } i\text{th bit of the prime factorization of } x \text{ is } b\}$.

$\text{MCSP} = \{(\chi_f, s) : \chi_f \text{ denotes a string of length } 2^m \text{ (for some } m) \text{ that is the truth-table of a Boolean function } f \text{ on } m \text{ variables and } s \text{ denotes an integer such that } f \text{ can be computed by a Boolean circuit of size at most } s.\}$

(In the definition of FACT , the prime factorization is presented as $p_1^{e_1}, \dots, p_k^{e_k}$, where each exponent $e_i > 0$, and $p_i < p_{i+1}$, so that each number has a unique prime factorization.)

I suspect that FACT is probably not complete for *any* reasonable complexity class under AC^0 reductions. In an earlier survey [14] I outlined a possible approach toward proving that this is the case. Namely, I noted that it would suffice to show that there is no one-one length-increasing NC^0 reduction from $\text{FACT} \times \{0, 1\}^*$ to FACT (or no isomorphism between these sets, computable and invertible in depth-three AC^0). All of my attempts to construct such a reduction have involved multiplication in some form, and this is not computable in AC^0 . Perhaps, I suggested, one could show that multiplication is inherent, in computing such a reduction. Now, however, after some illuminating discussions with Michal Koucký, I no longer think that this is a promising approach. One way to build a padding function would be to map the pair $((x, i, b), y)$ to the triple (z, i, b) , where $z = xy'$, where y' has binary representation $10^\ell y_1 0^\ell y_2 0^\ell \dots y_{n-1} 0^\ell y_n 0^\ell z'$ where ℓ is suitably large, and where z' has $\log^{O(1)} n$ bits. If y' is prime, then it will be the largest prime factor of z , and thus the initial part of the prime factorizations of x and of z will be the same. The product xy' can be computed in AC^0 , because of the padding by 0^ℓ and because z' is small. It is reasonably likely that a value of z' exists so that y' will be prime, although number theorists have not yet established that this holds. It would be *very* hard to show that *no* such z' can be found in uniform AC^0 . Thus it is reasonably likely that a padding function for (a suitable encoding of) FACT does exist in AC^0 . This does not guarantee that such a padding function can be found in NC^0 , but it does illustrate some of the difficulties of pursuing this approach.

Kabanets and Cai have presented some arguments, suggesting that MCSP is not complete for NP under \leq_m^P reductions [30]. Can one obtain even stronger evidence, suggesting that MCSP is not p-isomorphic to SAT ? It is certainly not clear that MCSP should have a padding function (i.e., a polynomial-time computable and invertible function f mapping $\text{MCSP} \times \{0, 1\}^*$ onto MCSP). It is even harder to see how to construct a padding function if one fixes the circuit size s to be something exponentially large, but still much smaller than 2^m , such as this set:

MCSP2 = $\{\chi_f : \chi_f \text{ denotes a string of length } 2^m \text{ that is the truth-table of a Boolean function } f \text{ on } m \text{ variables such that } f \text{ can be computed by a Boolean circuit of size at most } 2^{m/2}.\}$

As Kabanets and Cai observe [30], if MCSP2 has a padding function computed in polynomial time, then BPP=P. The connection between the paddability of MCSP2 and the BPP vs P problem arises through the easy observation that any set C isomorphic to SAT has P-printable sets contained both in C and in \overline{C} , combined with the following equivalence (where the first condition listed is the well-studied Impagliazzo-Wigderson derandomization hypothesis [27]):

- There is a set $A \in \mathbb{E}$ that requires circuits of size greater than $2^{n/2}$ for all large n iff
- There is a P-printable set B contained in the complement of MCSP2 of the form $B = \{\chi_f : \chi_f \text{ denotes a string of length } 2^m \text{ that is the truth-table of } A^{=m}\}.$

The observations above do *not* provide much evidence against MCSP2 being isomorphic to SAT; rather, they merely indicate that it will not be easy to prove that it *is* isomorphic to SAT. What unlikely consequences would follow if MCSP2 (or MCSP) turned out to be isomorphic to SAT?

3. Creative Sets

A set is defined to be *creative* if its complement is *productive* (and this holds for all of the variants of “creativity” and “productivity” that have been considered). Thus, in order to discuss creative sets, we must first define productive sets.

A set A is *productive over a class of languages* \mathcal{C} if there is a function (a so-called “*productive function*”) witnessing that $A \notin \mathcal{C}$, in some sense. In order to make this definition precise, we must be explicit what notion of Turing machine indices I we are using to represent elements of \mathcal{C} , and what class of functions F we will allow as productive functions. Thus we can say that A is (F, I) -productive if there is a function $f \in F$ such that, for every $i \in I$, $f(i) \in A$ if and only if $f(i)$ is not in the language accepted by machine i . That is, given i , f finds an input on which A differs from the i -th element of \mathcal{C} .

Agrawal and Biswas define a set A to be NP-creative if its complement is (polynomial-time, I)-productive, where I is an indexing of nondeterministic polynomial-time Turing machines, where machine M_i has the property that, on all inputs, it runs in time $|i|$ [10]. It is far from obvious that this is an appropriate definition, since these machines all run in $O(1)$ time! Thus, in particular, $\{L(M_i) : i \in I\}$ is *not* equal to NP, and doesn’t even contain all of the sets in, say, AC^0 ! Nevertheless, Agrawal and Biswas are able to demonstrate that this definition yields at least as many sets as an earlier notion of creativity (the “ k -creative” sets of [29], which were re-dubbed “ k -completely-creative” sets by Wang [34] to distinguish them from another creativity notion he introduced), and they also show that all NP-creative sets are NP-complete (in contrast to the situation for the “ k -creative” sets of [34], which are neither known to include sets such as SAT, nor to be contained in the class of

NP-complete sets). Figure 1 indicates the inclusion relations among these various classes of “creative” sets for NP.

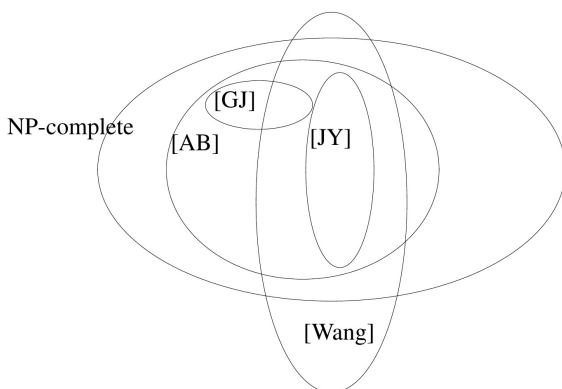


FIGURE 1. Diagram, showing (likely) inclusions among classes of “creative” sets for NP. The region labeled “[GJ]” indicates the list of “natural” NP-complete problems catalogued in [22]. It is not known to contain any of the k -creative sets defined by Joseph and Young [29], indicated by the region labeled [JY]. This same class was called “ k -completely-creative” by Wang [34], who also introduced another class of k -creative sets, indicated by the region labeled [Wang]; it is not known whether all of those sets are NP-complete. The region labeled [AB] indicates the NP-creative sets of Agrawal and Biswas [10].

However, when these creativeness definitions are adapted to larger complexity classes (such as EXP), they all coincide exactly with the class of sets complete under \leq_m^p reductions. (The issue boils down to a question of whether the complexity class \mathcal{C} can diagonalize over the class F of productive functions. This is true when $\mathcal{C} = \text{EXP}$, but is not known to be true for $\mathcal{C} = \text{NP}$.)

Even though the definition of NP-creative sets is less intuitive than the definition of the class of sets that are p-isomorphic to SAT, Agrawal and Biswas make a convincing argument that all “natural” NP-complete sets (including all of the NP-complete sets listed in [22]) are NP-creative. Thus there is some merit in investigating the “Creativity Hypothesis” mentioned in the introduction – the hypothesis that all NP-complete sets are NP-creative – as an alternative to the Berman-Hartmanis conjecture. Proving that either of these conditions hold would entail proving $P \neq \text{NP}$. (In the case of the Creativity Hypothesis, Agrawal and Biswas show that any NP-creative set is complete for NP under reductions that are “exponentially honest”,

in the sense that, for some constant c , $2^{c|f(x)|} > |x|$ for all x [10]. Thus, in particular, no finite set can be NP-creative.) It is particularly interesting that Agrawal and Biswas show that, if all of the sets that are p-isomorphic to SAT are NP-creative, then the Creativity Hypothesis holds.

The Creativity Hypothesis has not received much attention. Here are some questions that might yield some interesting insights:

- Are all of the sets that are complete for NP under AC^0 reductions NP-creative? How about the sets that are complete under first-order projections? Or the sets that are complete for NP under 1-L reductions?
- If one assumes that FACT or MCSP are NP-creative, can one derive stronger conclusions than if one merely assumes that these sets are NP-complete?
- Agrawal and Biswas have shown that the complement of any NP-creative set contains an infinite subset in NP. Consider a set such as $\{x : \text{the time-}n^2\text{-bounded Kolmogorov complexity of } x \text{ is greater than } |x|/2\}$. Would we expect this set to have an infinite NP-subset? (Actually, the answer is probably *Yes!* It is observed in Section 2.4 that, under the Impagliazzo-Wigderson derandomization hypothesis, this set even has a P-printable subset.) Can one derive strong and unlikely conclusions from the assumption that this set is the complement of an NP-creative set?

4. Universal Relations

All of the NP-complete sets that are p-isomorphic to SAT have a padding function, and even the NP-complete sets that are not known to be p-isomorphic to SAT (such as certain k -creative sets, and sets of the form $f(\text{SAT})$ where f is one-way) have “padding” functions if we drop the requirement of invertibility (i.e., a reduction from $A \times \Sigma^*$ to A that is one-one and length-increasing, but is not necessarily invertible). Similarly, all known NP-complete sets are disjunctive-self-reducible. (A set A is called “disjunctive-self-reducible” if there is a polynomial-time-computable function that takes a string x as input, and produces a list $y_1, \dots, y_{|x|^{O(1)}}$ as output, such that $x \in A$ iff $\exists i y_i \in A$.)

Agrawal and Biswas defined two operators on relations (which they name the *join* and *equivalence* operators) that are related to paddability (without invertibility) and disjunctive-self-reducibility, respectively (in the sense that if the witness relation for a set A has the given operator computable in polynomial time, then A is paddable or disjunctive-self-reducible, respectively). Remarkably, they were able to show that the relations with these two operators are *precisely* the relations from which any other NP-witness relation can be “recovered” in a fairly natural sense. (The details of these definitions will not be repeated here; see [9]).

One thing that I particularly like about [9] is their presentation of a new class of NP-complete sets. Let f be *any* one-one and size-increasing polynomial function. They define a relation R_f as follows: $(z, w) \in R_f$ if $|w| = 4|z|^3$ and one of the following three conditions hold:

1. $|z| = 1$ and $w \in \{0100, 0101, 0110, 0111, 1001, 1010, 1011\}$.

2. For some $r > 0$, $w = \#^r x_1 \# w_1 \# \# x_2 \# w_2 \# \# \dots \# \# x_n \# w_n$, where $f(1 \# x_1 \# x_2 \# \dots \# x_n) = z$ and for all $i \leq n$, $(x_i, w_i) \in R_f$.
3. For some $r > 0$, $w = \#^r x \# i_1 \# i_2 \# \dots \# i_n \# j_1 \# \dots \# j_n \# \# w'$, where $f(1 \# x \# i_1 \# \dots \# i_n \# j_1 \# \dots \# j_n) = z$ and for each $k \leq n$, bits number i_k and j_k of w' are the same.

Agrawal and Biswas show that $\{x : \exists y(x, y) \in R_f\}$ is NP-complete. I know of no *direct* way to see that this set is NP-complete; the proof of completeness presented by Agrawal and Biswas follows because the relation R_f is universal (because it has the required join and equivalence operators). It would be interesting to know if there is any example of a natural NP-complete problem, for which it is easier to prove NP-completeness by presenting the join and equivalence operators, than to present a traditional \leq_m^P reduction.

The theory of Probabilistically-Checkable Proofs tells us that problems in NP have witness relations with very special encoding structure. It would be interesting to know if this body of knowledge can be merged with the theory of universal relations, to obtain any new insights.

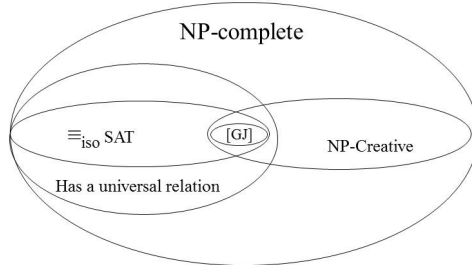


FIGURE 2. Diagram, showing inclusions among notions of “natural” NP-complete sets discussed in this survey.

Figures 2, 3, and 4 show inclusion relations among the different notions considered in this survey, all of which present ways to give a mathematically-precise definition that can serve as a proxy for the vague concept of what it means to be a “natural” NP-complete set:

- being p-isomorphic to SAT.
- being NP-creative.
- having a universal relation.

Do all NP-creative sets have universal relations? (Note in this regard that Agrawal and Biswas show that sets with universal relations are all complete for

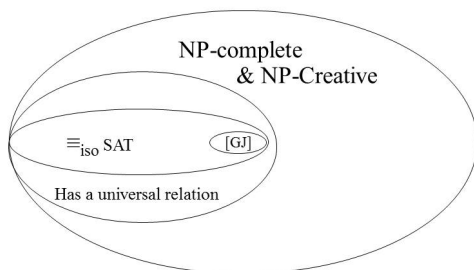


FIGURE 3. Diagram, showing inclusions among notions of “natural” NP-complete sets, if p-isomorphism preserves NP-Creativity.

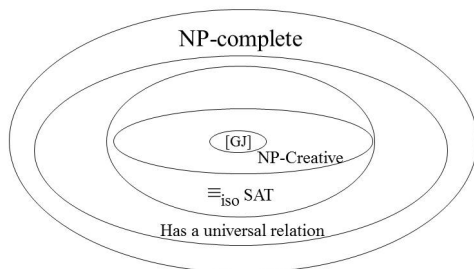


FIGURE 4. Diagram, showing inclusions among notions of “natural” NP-complete sets, if all NP-Creative sets are p-isomorphic.

NP under polynomially-honest reductions (i.e., reductions f where there is a polynomial p such that $p(|f(x)|) \geq |x|$ for all x [9]), whereas the NP-creative sets are only known to be complete under exponentially-honest reductions [10]. Thus

it might be better to ask first whether all NP-creative sets that are complete under polynomially-honest reductions have universal relations.)

Are the standard witness relations for MCSP and FACT universal? (Possibly this question is easy to answer ...) Note in this regard that Agrawal and Biswas show that the standard witness relation for Graph Isomorphism is not universal – as well as a (somewhat non-standard) witness relation for Simple Max Cut. They introduce a more general notion of universal relations, which they call “generalized universal relations”, and show that the Simple Max Cut witness relation is generalized universal. If one is able to show that the standard witness relations for MCSP and FACT are not universal, then perhaps one can show that they are also not generalized universal. This would provide some additional evidence that these problems are not NP-complete.

Are there any additional implications that one can prove, regarding the Berman-Hartmanis conjecture, the Creativity Hypothesis, and the question of whether all NP-complete sets have universal witness relations?

(There has been some additional work by other authors, regarding universal relations. The reader is referred to [19] for a discussion of this work.)

5. Conclusions

The notions of p-isomorphism, NP-creativity, and universality provide three ways to identify properties that are shared by all of the “natural” NP-complete sets. Although the work of Somenath Biswas and others has given us a body of interesting results regarding these notions, a number of intriguing open questions remain.

References

- [1] Manindra Agrawal. On the isomorphism conjecture for weak reducibilities. *Journal of Computer and System Sciences*, 53(2):267–282, 1996.
- [2] Manindra Agrawal. Towards uniform AC^0 - isomorphisms. In *Proc. IEEE Conf. on Computational Complexity*, pages 13–20, 2001.
- [3] Manindra Agrawal. Pseudo-random generators and structure of complete degrees. In *IEEE Conference on Computational Complexity*, pages 139–147, 2002.
- [4] Manindra Agrawal. The isomorphism conjecture for constant depth reductions. *Journal of Computer and System Sciences*, 77(1):3–13, 2011.
- [5] Manindra Agrawal. The isomorphism conjecture for NP. In S. Barry Cooper and Andrea Sorbi, editors, *Computability in Context: Computation and Logic in the Real World*, pages 19–48. World Scientific Press, 2011.
- [6] Manindra Agrawal and Eric Allender. An isomorphism theorem for circuit complexity. In *Proc. IEEE Conf. on Computational Complexity*, pages 2–11, 1996.
- [7] Manindra Agrawal, Eric Allender, Russell Impagliazzo, Toniann Pitassi, and Steven Rudich. Reducing the complexity of reductions. *Computational Complexity*, 10(2):117–138, 2001.

- [8] Manindra Agrawal, Eric Allender, and Steven Rudich. Reductions in circuit complexity: An isomorphism theorem and a gap theorem. *Journal of Computer and System Sciences*, 57(2):127–143, 1998.
- [9] Manindra Agrawal and Somenath Biswas. Universal relations. In *Proc. IEEE Conf. on Structure in Complexity Theory*, pages 207–220, 1992.
- [10] Manindra Agrawal and Somenath Biswas. NP-creative sets: A new class of creative sets in NP. *Mathematical Systems Theory*, 29(5):487–505, 1996.
- [11] Manindra Agrawal and Somenath Biswas. Polynomial-time isomorphism of 1-L-complete sets. *Journal of Computer and System Sciences*, 53(2):155–160, 1996.
- [12] Manindra Agrawal and Osamu Watanabe. One-way functions and the Berman-Hartmanis conjecture. In *Proc. IEEE Conf. on Computational Complexity*, pages 194–202, 2009.
- [13] Eric Allender. Isomorphisms and 1-L reductions. *Journal of Computer and System Sciences*, 36(3):336–350, 1988.
- [14] Eric Allender. Some pointed questions concerning asymptotic lower bounds, and news from the isomorphism front. In G. Paun, G. Rozenberg, and A. Salomaa, editors, *Current Trends in Theoretical Computer Science: Entering the 21st Century*, pages 25–41. World Scientific Press, 2001.
- [15] Eric Allender, José L. Balcázar, and Neil Immerman. A first-order isomorphism theorem. *SIAM Journal on Computing*, 26(2):557–567, 1997.
- [16] Eric Allender and Vivek Gore. Rudimentary reductions revisited. *Inf. Process. Lett.*, 40(2):89–95, 1991.
- [17] L. Berman and J. Hartmanis. On isomorphism and density of NP and other complete sets. *SIAM Journal on Computing*, 6:305–322, 1977.
- [18] Hans-Jörg Burtschick and Albrecht Hoene. The degree structure of 1-L reductions. In Ivan M. Havel and Václav Koubek, editors, *MFCS*, volume 629 of *Lecture Notes in Computer Science*, pages 153–161. Springer, 1992.
- [19] Vinay Choudhary, Anand Kumar Sinha, and Somenath Biswas. Universality for nondeterministic logspace. In *Proc., 1st International Conference on Language and Automata Theory and Applications (LATA)*, pages 103–114, 2007.
- [20] Stephen A. Cook and Pierre McKenzie. Problems complete for deterministic logarithmic space. *J. Algorithms*, 8(3):385–394, 1987.
- [21] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.
- [22] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the theory of NP-completeness*. W.H. Freeman and Company, New York, 1979.
- [23] Ryan C. Harkins, John M. Hitchcock, and Aduri Pavan. Strong reductions and isomorphism of complete sets. In *Proc. Conf. on Found. of Software Technology and Theo. Comp. Sci. (FST&TCS)*, volume 4855 of *Lecture Notes in Computer Science*, pages 168–178. Springer, 2007.
- [24] Juris Hartmanis, Neil Immerman, and Stephen R. Mahaney. One-way log-tape reductions. In *Proc. IEEE Symp. on Found. of Comp. Sci. (FOCS)*, pages 65–72, 1978.
- [25] Juris Hartmanis and Stephen R. Mahaney. Languages simultaneously complete for one-way and two-way log-tape automata. *SIAM Journal on Computing*, 10(2):383–390, 1981.

- [26] Lane A. Hemachandra and Albrecht Hoene. Collapsing degrees via strong computation. *Journal of Computer and System Sciences*, 46(3):363–380, 1993.
- [27] R. Impagliazzo and A. Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proc. ACM Symp. on Theory of Computing (STOC)*, pages 220–229, 1997.
- [28] Neil D. Jones. Space-bounded reducibility among combinatorial problems. *Journal of Computer and System Sciences*, 11(1):68–85, 1975.
- [29] Deborah Joseph and Paul Young. Some remarks on witness functions for nonpolynomial and noncomplete sets in NP. *Theoretical Computer Science*, 39:225–237, 1985.
- [30] Valentine Kabanets and Jin-Yi Cai. Circuit minimization problem. In *Proc. ACM Symp. on Theory of Computing (STOC)*, pages 73–79, 2000.
- [31] Emil L. Post. Recursively enumerable sets of positive integers and their decision problems. *Bull. Am Math. Soc.*, 50:284–316, 1944.
- [32] Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967.
- [33] Robert I. Soare. *Recursively-Enumerable Sets and Degrees*. Springer-Verlag, 1987.
- [34] Jie Wang. On p-creative sets and p-completely creative sets. *Theoretical Computer Science*, 85(1):1–31, 1991.

Eric Allender
Dept. of Computer Science
Rutgers University
New Brunswick, NJ 08855, USA

e-mail: allender@cs.rutgers.edu