Natural Proofs Explained

Chris Calabro

August 20, 2004

Abstract

The current inability to prove

$$P \neq NP \tag{1}$$

may come from any of the following:

1. It may be independent of ZFC set theory.
2. Current techniques may be too weak to prove it.
3. It may be false.

To overcome 3 would seem to require developing a polytime algorithm for some NP-complete problem. The possibility of overcoming 1 will be discussed in section 1. But the main purpose of this essay is to consider how to overcome 2.

1 Independence from ZFC

Gödel's incompleteness theorem guarantees that in any "sufficiently rich" [1] theory, e.g. number theory or set theory, there are sentences independent of the axioms. The proof constructs a sentence that essentially means, "The Turing machine that gets a description of itself and searches for a proof of whether or not it halts and does the opposite indeed halts." If there were an explicit formal proof that this machine halts or not, then it would certainly find such a proof and we would have a contradiction; so there can be no such formal proof. Of course the proof that *we* just gave that such a machine does not halt lies outside of the formal proof system.

Such a diagonalization argument almost certainly would not suffice to prove $P \neq NP$ since Baker, Gill, Solovay [1] proved that any relativizing proof technique could not settle (1) and diagonalization arguments (nearly always) relativize. More specifically, they demonstrated 2 oracles A, B such that

$$P^A = NP^A$$
$$P^B \neq NP^B,$$

[1] what "sufficiently rich" means is not important since the only cases we care about here are number theory and set theory, but roughly speaking, it means that the language is capable of expressing recursion.

and so if each step of a proof holds relative to any oracle, then the proof cannot settle (1).

Of course, this doesn't stop diagonalization from proving that (1) is independent of ZFC. But if diagonalization is too weak, then it is hard to imagine any technique that will prove independence.

2 Natural proofs

Just as Baker, Gill, Solovay showed that relativizing techniques were too weak to settle (1), Razborov and Rudich [2] showed that another very large class of proof techniques, called "natural proofs", were too weak to prove (1). We now introduce some definitions to explain this idea.

2.1 Notation

The big-oh-tilde notation is a minor extension of the big-oh notation, only "fuzzier" as it ignores $\lg n$ factors. Also, the $\exists!$ notation is a minor extension of $\exists$, meaning "there is a unique". So it should not be too offensive to fuzzify the quantifiers $\forall, \exists$ with a tilde to produce $\widetilde{\forall}, \widetilde{\widetilde{\exists}}$, which we use to mean "for sufficiently large" (or "ultimately"), and "for infinitely many" (or "infinitely often"). More formally,

$$\widetilde{\forall} n\ \phi(n) \Leftrightarrow \exists n_0\ \forall n \geq n_0\ \phi(n)$$
$$\widetilde{\widetilde{\exists}} n\ \phi(n) \Leftrightarrow |\{n \mid \phi(n)\}| = \infty,$$

where the quantification is over ω. The appendix explores the basic properties of $\widetilde{\forall}, \widetilde{\widetilde{\exists}}$.

2.2 Definitions

Let $F_n = 2^{2^n}$ be the set of Boolean functions mapping n-bit strings to 1 bit and let $F = \cup_n F_n$. If $f \in F_n$, then the *truth table* of f is $\langle f \rangle = f(0)f(1)\cdots f(2^n - 1)$ is a string of 2^n bits. A *Boolean function family* is a function $f : 2^* \to 2$. We denote by f_n the restriction $f|2^n$. The *language* of f is $L_f = f^{-1}(1)$.

A *combinatoric property* is a set $C \subseteq F$. Notice that this is *not* a set of function families since no element of C has domain 2^*. A Boolean function family f *has property* C iff $\widetilde{\forall} n\ f_n \in C$. Similarly, a function $f \in F_n$ *has property* C iff $f \in C$.

Let us say that a set of languages $\Gamma \subseteq \mathcal{P}(2^*)$ is *asymptotic* iff

$$\forall L \in \Gamma, L' \subseteq 2^*, |L'| < \infty\ L \triangle L' \in \Gamma.$$

In other words, changing the membership of a finite number of strings does not change whether a language is in the class. Every complexity class I can think of is asymptotic, even i.o.AvgP, which is not closed under polytime reductions. So let us say that Γ is a *complexity class* iff Γ is asymptotic.

Let $\Gamma, \Lambda \subseteq \mathcal{P}(2^*)$ be complexity classes. A lower bound proof for a Boolean function family f always has the form "f has the combinatoric property C and any family g with the property C fails to be in Λ." Any lower bound proof can be made to have this form since we could always take $C = \{f_n \mid n \in \omega\}$ (To see this, suppose indirectly that f has property C, $L_f \notin \Lambda$, yet $\exists$ function family g with property C and $L_g \in \Lambda$. then $\widetilde{\forall} n\ g_n \in C$, which implies that $L' = L_f \triangle L_g$ is finite, contradicting that Λ is asymptotic.), but in common cases C is quite large since its purpose is to ignore irrelevant features of f and focus only on some specific property. E.g. to lower bound the circuit complexity of parity, it is sufficient to consider the property of parity that it has many isolated solutions, and clearly many families have that property.

A property C is Γ-*natural* iff $\exists C^* \subseteq C$ such that, letting $C_n^* = C^* \cap F_n$,

constructivity: $\{\langle f \rangle \mid f \in C^*\} \in \Gamma$, i.e. deciding whether a truth table corresponds to a function with property C^* is computable in Γ. Notice that for $f \in F_n$, $|\langle f \rangle| = 2^n$, and so if $\Gamma = \mathrm{P}$, then this condition would be met provided we could decide whether $f \in C^*$ in time $2^{O(n)}$.

largeness: $|C_n^*| \geq 2^{-O(n)} |F_n|$, i.e. we demand that C_n^* has size that is an inverse poly (in the size of truth tables) fraction of F_n so that a random function in F_n has a nonnegligible probability of having property C^*.

Next, C *is useful against* Λ iff

usefulness: Every Boolean function family f with property C has $L_f \notin \Lambda$.

We can also say that a *language* $L \subseteq 2^*$ is constructible, large, or useful iff the characteristic function

$$\chi_L(x) = \begin{cases} 1 & \text{if } x \in L \\ 0 & \text{else} \end{cases}$$

is constructible, large, or useful.

A Γ-*natural proof against* Λ is then a proof showing that some family f has the Γ-natural property C and that C is useful against Λ, the conclusion of which is that $L_f \notin \Lambda$.

2.3 What's so natural about natural proofs?

If we take, e.g., $\Gamma = \mathrm{P/poly}$, then nearly any lower bound proof is Γ-natural since the kinds of combinatoric properties that we use in proofs can usually be exhaustively checked by looking at the truth table, and certainly we gain even more checking power if we allow our checks to use nonuniform circuits. E.g. every language in NP is P-constructible.

The largeness condition only requires that the property hold with nonnegligible probability for a random function. Whether this is "natural" is debatable. Possibly a lower bound proof could use a truly rare combinatoric property that the function family in question nonetheless has.

All current lower bound proofs are natural for some relatively weak class such as P.

2.4 Why are natural proofs weak?

Natural proofs are unlikely to yield strong lower bounds because of the following very surprising duality: if a combinatoric property C is Γ-natural against Λ, then there are no pseudo-random generators in a certain complexity class (related to Λ) with a certain hardness (related to Γ). [2]

For example, if C is P/poly-natural against P/poly then there would be no P/poly [3] pseudo-random generator with exponential hardness. Razborov and Rudich [2] prove many other properties of natural proofs, but we will content ourselves with the above. The proof below is the same as theirs, but hopefully clearer.

Define the *hardness* of a pseudo-random generator (PRG) $G : 2^n \to 2^{2n}$ as

$$\begin{aligned} H(G) = \min\{S > 0 \mid \exists \text{ circuit } c \text{ of size} \leq S \text{ with} \\ |Pr_{x \in_u 2^n}(c(G(x)) = 1) - Pr_{y \in_u 2^{2n}}(c(y) = 1)| \geq S^{-1}\}. \end{aligned}$$

Also, say a *PRG family* is a function family $G : 2^* \to 2^*$ where $G_n : 2^n \to 2^{2n}$.

Theorem 1. *Let $C \subseteq F$ be* P/poly*-natural against* P/poly*. Then every PRG family G computable in* P/poly *has $\widetilde{\exists} k \; H(G_k) \leq 2^{k^{o(1)}}$.* [4]

In particular, if there is a P/poly exponentially hard PRG family, then there are no P/poly-natural proofs against P/poly.

Proof. Let G be a PRG family computable in P/poly, i.e. $\exists$ poly p $\forall k \in \omega$, each output bit of G_k can be computed by a circuit of size $\leq p(k)$.

Let $0 < \epsilon < 1$. We will show that

$$\widetilde{\exists} k \; H(G_k) \leq 2^{O(k^{\epsilon})}, \tag{2}$$

which implies the conclusion. (The hidden variables inside the big-oh are, of course, quantified before k is.)

Let $C^* \subseteq C$ satisfy the constructivity and largeness conditions. We have a poly size circuit family c that can take as input a truth table and tell whether it has property C^*. We want to construct a statistical test that can take as input a single string X of size $2k$ and tell whether it came from G or from a uniformly random source. To do this we will use X to construct a pseudo-random function and analyze its truth table with c.

Let $n = n_\epsilon(k) = \lceil k^\epsilon \rceil$. Define $g_0, g_1 : 2^k \to 2^k$ as the least and most significant k bits of G_k, and for a string $y \in 2^n$, define $g_y = g_{y_n} \circ \cdots \circ g_{y_1}$. Note

[2] For this to hold in general I suspect that we need some extra hypotheses on Γ, Λ, such as closure under polytime reductions.

[3] [2] says that the PRG must be in P, but the proof seems to hold for P/poly as well.

[4] [2] says $\widetilde{\forall} k$, but such a strong conclusion is not necessary to conclude that G is weak; furthermore it would seem to necessitate a considerable strengthening of the hypothesis by changing the definition of "has the property C" within the definition of usefulness from $\widetilde{\forall} n \; f_n \in C$ to $\widetilde{\exists} n \; f_n \in C$. Such a change would also invalidate the claim that "any lower proof can be seen to have the combinatoric form".

that there is no conflict of notation here, even if $|y| = 1$. Define $f_k : 2^k \to F_n$ by $f_k(x)(y) = g_y(x) \mod 2$.

Since $\epsilon < 1$, $n \leq k$, and so f_k is the composition of at most k copies of (functions derived from) G_k. So f_k is computable by poly size circuits, the polynomial being independent of ϵ. If we fix a value $\tilde{x}_k \in 2^k$ for each k, then the function $f'_k = f_k(\tilde{x}_k)$ also has poly size circuits. (Although now the polynomial may depend on ϵ because it has input size n, which could be much smaller than k and by an amount dependent on ϵ.) Let f' be the family $\cup_k f'_k$. Then $L_{f'} \in$ P/poly.

f' does not have property C since C is useful against P/poly. Letting $\tilde{X}$ denote the Cartesian product $2^0 \times 2^1 \times 2^2 \times \cdots$, we have

$$\forall \tilde{x} \in \tilde{X} \ \tilde{\exists} k \ f_k(\tilde{x}_k) \notin C,$$

which implies, by theorem 3 in the appendix, that

$$\tilde{\exists} k \ \forall \tilde{x} \in \tilde{X} \ f_k(\tilde{x}_k) \notin C.$$

So, letting $x \in_u 2^n$, we have $\tilde{\exists} k \ Pr_x(f_k(x) \in C^*) = 0$, and by constructivity we can test for the condition $f_k(x) \in C^*$ using a circuit of size $2^{O(n)}$. By the largeness condition, letting $h \in_u F_n$, we have $Pr_h(h \in C^*) \geq 2^{-O(n)}$. So

$$|Pr(f_k(x) \in C^*) - Pr(h \in C^*)| \geq 2^{-O(n)}, \tag{3}$$

and so C^* provides a statistical test to distinguish $f_k(x)$ from a random function. We now use this test and the *hybrid argument* of Goldwasser and Micali [3] to break G.

Let T be the binary tree of height n with internal nodes $v_1, \ldots, v_{2^n - 1}$, with leaves the strings in 2^n, and so that the indexes of the internal nodes increase from left to right, bottom to top. (so level i consists of $v_{1+\sum_{j=i+1}^{n-1} 2^j}, \ldots, v_{\sum_{j=i}^{n-1} 2^j}$) Let T_i be the union of the leaves and $\{v_1, \ldots, v_i\}$. If $y \in 2^n$ is a leaf, let $v(i, y)$ be the highest ancestor of y in T_i and $h(i, y)$ be the distance from $v(i, y)$ to y. Define $g_{i,y} = g_{y_n} \circ \cdots \circ g_{y_{n-h(i,y)+1}}$. Let $x_v \in_u 2^k$ be a k-bit random variable for each $v \in T$, each chosen independently. Define the hybrid random function

$$h_i(y) = g_{i,y}(x_{v(i,y)}) \mod 2.$$

Then h_0 has the same distribution as h and $h_{2^n - 1}$ has the same distribution as $f_k(x)$. Using (3), the triangle inequality, and the fact that some value, over i, must achieve the average, we have $\exists i \in 2^n - 1$ with

$$|Pr(h_i \in C^*) - Pr(h_{i+1} \in C^*)| \geq 2^{-O(n)}. \tag{4}$$

By considering conditional probabilities, we can fix all of the values of the x_v's except for $x_{v_{i+1}}$ so as to maximize the lhs. So now we have a statistical test breaking G: on input $X \in 2^{2k}$, first create the truth table for the random function $t : 2^n \to 2$ (whose randomness comes from X) defined by

```
$t(y)$
  if $y$ is a descendent of $v_{i+1}$,
    if $y_{n-h(i+1,y)+1} = 0$
      $r \leftarrow$ least significant $k$ bits of $X$
    else
      $r \leftarrow$ most significant $k$ bits of $X$
    return $g_{i,y}(r) \mod 2$
  else
    return $h_i(y)$
```

If X is uniform, then t will have the same distribution as h_i. If X is distributed as $G_k(x)$, then t will have the same distribution as h_{i+1}. Notice that we can evaluate $\langle t \rangle$ with a circuit of size $2^{O(n)}$. We then use our circuit for deciding whether t has property C^*. From (4), we have $H(G_k) \leq 2^{O(n)} \leq 2^{O(k^\epsilon)}$, which completes the proof. □

3 Appendix

We prove some properties of $\widetilde{\forall}, \widetilde{\exists}$ that were used in the proof above, the later of which are somewhat nontrivial.

Lemma 2. *$\neg\widetilde{\forall} n\ \phi(n) \Leftrightarrow \widetilde{\exists} n\ \neg\phi(n)$.*

We will see in the proof that quantification over ω is not strictly necessary: any nonempty poset S with no maximal element and with the property

$$\forall n_0 \in S\ |\{n \in S \mid n \not\geq n_0\}| < \infty \tag{5}$$

will do. For example, the poset $S = \omega$ with the partial ordering $\{(a, b) \in S^2 \mid b - a \geq 2\}$ has these properties. But these are strange properties, so we will content ourselves to assume $S = \omega$ with the usual ordering.

Proof. ($\Rightarrow$) This relies crucially on the fact that the poset S over which we quantify is nonempty and has no maximal element. We will recursively develop a sequence $n_0 < n_1 < \cdots \in S$ such that $\forall i \in \omega\ \neg\phi(n_i)$. To get started, we observe that since $S \neq \emptyset$, $\exists n'_0 \in S$. By hypothesis $\exists n_0 \geq n'_0\ \neg\phi(n_0)$. Now suppose that $n_0 < \cdots < n_{i-1}$ have been constructed such that $\neg\phi(n_0), \ldots, \neg\phi(n_{i-1})$. Since S has no maximal element, $\exists n'_i > n_{i-1}$. Again by hypothesis, $\exists n_i \geq n'_i\ \neg\phi(n_i)$, which completes the induction.

($\Leftarrow$) Suppose indirectly that $\exists n_0\ \forall n \geq n_0\ \phi(n)$. By (5), $|\{n \mid n \not\geq n_0\}| < \infty$, and yet $|\{n \mid \neg\phi(n)\}| = \infty$. So $\{n \mid \neg\phi(n)\} - \{n \mid n \not\geq n_0\} \neq \emptyset$. So $\exists n \geq n_0\ \neg\phi(n)$, a contradiction. □

Using the new suggestive notation $\widetilde{\forall}, \widetilde{\exists}$, we can see the following theorem as allowing an inversion of quantifier order, as so many theorems in mathematics are.

Theorem 3. *Let $X_0, X_1, \ldots$ be a sequence of sets, $X = \prod_{i\in\omega} X_i$, and $\forall n \in \omega$ let $U_n \subseteq X_n$. Then*

$$\forall x \in X\ \widetilde{\exists} n\ x_n \in U_n \Leftrightarrow \widetilde{\exists} n\ \forall x \in X\ x_n \in U_n \tag{6}$$

$$\text{and} \quad \forall x \in X\ \widetilde{\forall} n\ x_n \in U_n \Leftrightarrow \widetilde{\forall} n\ \forall x \in X\ x_n \in U_n. \tag{7}$$

Proof. We prove this directly by using combinatorics and the axiom of choice. (Is it possible to prove it with the Tychonoff theorem, which states that the product of compact topological spaces is compact, or the compactness theorem from sentential logic?)

If $\exists i\ X_i = \emptyset$, then the theorem is vacuously true, so assume $\forall i\ X_i \neq \emptyset$. Wlog assume that the X_i are disjoint. Consider the infinite digraph $G = (V, E)$ where $V = \cup_i X_i \cup \{r\}$, $r \notin \cup_i X_i$, and $E = \{(x, y) \mid \exists i\ x \in X_i, y \in X_{i+1}\} \cup \{(r, y) \mid y \in X_0\}$.

Since the X_i are disjoint, there are edges from X_i to X_j iff $j = i + 1$. Furthermore, there are edges from the root r to each element of X_0. Define the n^{th} *level* of G as $X_n \subseteq V$. We may consider the root r as belonging to level -1.

There is a bijective correspondence between elements of X and infinite paths in G starting from r (from now on called simply *paths*). Color each node $x_n \in X_n \subseteq V$ *black* if $x_n \in U_n$ and *white* otherwise.

The forward direction of (6) says that if each path in G has an infinite number of black nodes, then an infinite number of levels are completely black. To prove this, suppose indirectly that the conclusion is false so that $\widetilde{\forall} n\ \exists x \in X\ x_n \notin U_n$; in other words, $\exists n_0$ so that all levels below n_0 each have a white node. Use the axiom of choice to choose a path $x \in X$ that involves 1 such white node in each level below n_0. Then x has only finitely many black nodes, contradicting hypothesis.

The reverse implication says that if there are an infinite number of completely black levels, then each path has an infinite number of black nodes, which is obvious.

The forward direction of (7) says that if each path is eventually all black, then below some level, the digraph is all black. To prove this, suppose indirectly that $\widetilde{\exists} n\ \exists x \in X\ x_n \notin U_n$; in other words, an infinite number of levels each have a white node. Use the axiom of choice to choose a path $x \in X$ that hits a white node at each of these levels. Then x is not eventually all black, contradicting hypothesis.

The reverse direction says that if below some level G is all black, then every path is eventually all black, which is obvious. □

References

[1] T. Baker, J. Gill, R. Solovay, Relativizations of the $P = NP$ question, *SIAM Journal on Computing*, 4:431-442, 1975

[2] A. Razborov, S. Rudich, Natural proofs, *Journal of Computer and System Sciences*, vol. 55, issue 1, 1997

[3] S. Goldwasser, S. Micali, Probabilistic Encryption, *Journal of Computer and System Science*, vol. 28, no. 2, 270-299, 1984 (preliminary version in 14th STOC, 1982)