The Concept of NP-Completeness Introduces a Statistical Argument Against P=NP

Author: Dobri Bozhilov

June 25th, 2025

Abstract

The concept of NP-completeness summarizes a multitude of problems that can be transformed into each other through polynomial reduction. Its main idea is: "if we solve one problem in polynomial time, we will solve them all." At first glance, this is promising. But upon closer analysis, it turns out that the mechanism of reduction itself requires that each NP-complete problem has an independent polynomial-time solution. This is because the reduction represents a transition between initially necessary solutions for each separate problem, and not a means to simplify the solution by finding it in only one place. Therefore, with each new proven NP-complete problem, the necessity for a solution to exist for it is added, if such a solution is even possible. If even one does not have a solution, then none can. Thus, a statistical argument against P=NP is formed.

Thesis

The reduction between NP-complete problems is a transformation with polynomial complexity from one problem to another. But both at the "input" (first problem) and at the "output" (second problem) there are elements with non-polynomial complexity (e.g., exponential). That is, the reduction does not change the complexity of the problems. The reduction itself cannot be a solution.

It can only be a solution if it itself is non-polynomial. In this case, however, the principle of NP-completeness is violated.

In short, exponential complexity cannot be reduced to polynomial through polynomial transformation.

Formally, let:

$$f : \Sigma^* \rightarrow \Sigma^*, \quad A \leq_p B \Leftrightarrow \quad x \in A \Leftrightarrow \quad f(x) \in B$$

where $f \in P$ is a polynomial function. This means that solving problem A can be reduced to solving B. But if $B \notin P$, then $A \notin P$. The reduction does not provide an independent solution.

Therefore, the reduction is useful only if $B \in P$. Otherwise, it only transforms the problem without reducing its complexity. For example, if the solving time of a problem is $T(n) = 2^n$, and $f(n) \in poly(n)$, then the result is $T(f(n)) = 2^{poly(n)} \notin P$.

Hence, the existence of a class of NP-complete problems requires that each of them has its own hypothetical polynomial-time solution, so that any of them can exist at all. And the reduction would only represent a transition between these solutions.

At present, the reduction works successfully for transforming between non-polynomial solutions. One problem solvable in exponential time can be solved through another — again in exponential time. The transformation is polynomial.

However, this principle does not introduce facilitation (solve just one and we're done), but statistically escalating evidence that $P \neq NP$.

If we assume that for each NP-complete problem there is an optimistically defined probability of 99% to find a polynomial-time solution, then the requirement for such to exist for all problems would lead to a negligible overall probability for any solution to exist. Currently, thousands of NP-complete problems exist. Between 500 and 700 are officially documented and recognized as such. Raising the probability 0.99 to the power of 500 gives 0.0066. Or 0.66%. This is the overall probability for the existence of a polynomial solution, even under extremely optimistic assumptions (0.99 per problem) and a minimal number of proven NP-complete problems (500). The real probability is much lower.

But most importantly, with the discovery of each new NP-complete problem, it adds "statistical evidence" that $P \neq NP$.

If $n = 1000$, then:

$$0.99^{1000} \approx 4.3 \times 10^{-5}$$

And as $n \to \infty$:

$$\lim_{n\to\infty} 0.99^n = 0$$

Thus, with each new NP-complete problem we approach a complete impossibility for the existence of a universal polynomial solution. This leads to:

$P \neq NP$ (by statistical induction)

Here we must clarify that the problems in the NP-complete class are extremely different from each other and require entirely different types of solutions.

They encompass different types of structures and requirements. Some require numerical solutions, others – combinatorial, others – graphs, etc. For example:

- some are based on combinatorics (SAT),

- others on numerical relations (Subset Sum),

- others on graph constructions (Hamiltonian Path).

In order for all of them to be solved polynomially simultaneously, mutually contradictory requirements are imposed.

These problems require different types of algorithmic properties. If each problem P_i requires a specific condition s_i for polynomial solving, and if there exist pairs s_i, s_j with s_i ∩ s_j = ∅, then there cannot be a universal property valid for all problems:

$$S = \{s_1, s_2, \ldots, s_k\}, \quad \exists\, i \neq j: s_i \cap s_j = \emptyset$$

Therefore, individual polynomial solutions for each problem cannot simultaneously exist, and the reduction requires exactly that — each problem to be equivalent to the others in difficulty.

An example of such a contradiction was analyzed in depth in the article *Invariance of Initial Conditions in P and NP and Structural Incompatibility Between Problems with Hypothetical "Additional Hidden Properties" Allowing Partial Solutions to Each (Bozhilov.D)*.

That is, to simplify an NP-complete problem to a polynomial solution requires hypothetical qualities that are not the same for another NP-complete problem. And the qualities can contradict each other.

Which makes it impossible to have simultaneous individual polynomial solutions for each. And these are mandatory and follow from the proven polynomial reduction between the problems.

Conclusion

Of course, mathematics does not accept statistical arguments as definitive ones. One might ask whether the choice is either there is proof (100%) or not (0%). And intermediate levels such as 99% or multiplication of probabilities, do not exist.

But the meaning of this article is to point out that the increase in the number of documented NP-complete problems makes proving P = NP increasingly difficult. For a single polynomial solution to exist, such must exist everywhere – and in diverse problems. And the probability for such to exist is inversely proportional to the number of problems.

In theory, it has been proven that an infinite number of NP-complete problems can exist.

When using a statistical approach, this may be considered complete proof, insofar as raising an arbitrary probability to a power tending to infinity yields an overall probability tending to zero.

That is, if we proceed from such logic, then $P \neq NP$.

Acknowledgements

As this work evolved - mostly over my ideas, I strongly benefited from the great wisdom of the inspired global scientific society and its effort to boldly dream and seek the hidden truth - even where no man has looked before. And obviously, when your faith is strong enough, the inevitable brightness comes. So I acknowledge all the researchers, colleagues, advisors, and mentors that contributed over time. Beyond academia, I am deeply thankful to my devoted family and friends. High peaks are never climbed alone. To editors, reviewers, and all who engaged: heartfelt respect. I also acknowledge the open science hive - tools, online networks, software, repositories, web events, AI, that made all this become real. No single path in life is easy, but walking together success is guaranteed.

Data Availability

This article does not contain any datasets beyond the numerical values already presented within the manuscript. All relevant calculations and results are fully documented in the paper.

Conflict of Interest

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References:

1. **S. A. Cook**, *The Complexity of Theorem-Proving Procedures*, Proceedings of the 3rd Annual ACM Symposium on Theory of Computing (STOC), 1971.
2. **E. Horowitz and S. Sahni**, *Fundamentals of Computer Algorithms*, Computer Science Press, 1978.
3. **M. Sipser**, *Introduction to the Theory of Computation*, Cengage Learning, 3rd ed., 2012.
4. **C. Papadimitriou**, *Computational Complexity*, Addison-Wesley, 1994.
5. **D.Bozhilov,** *Invariance of Initial Conditions in P and NP and Structural Incompatibility Between Problems with Hypothetical "Additional Hidden Properties" Allowing Partial Solutions to Each, 2025, Zenodo, doi: 10.5281/zenodo.15286615.*

Contact:

Dobri Bozhilov
AI Laboratory, Sofia, Bulgaria

zadrugata2019@gmail.com
++359 887 467 818