COMP331/557

Chapter 8:
Complexity Theory

(Cook, Cunningham, Pulleyblank & Schrijver, Chapter 9;
Korte & Vygen, Chapter 15
Garey & Johnson)

Efficient Algorithms: Historical Remark

Edmonds (1965):

I am claiming, as a mathematical result, the existence of a *good* algorithm for finding a maximum cardinality matching in a graph.

There is an obvious finite algorithm, but that algorithm increases in difficulty exponentially with the size of the graph. It is by no means obvious whether *or not* there exists an algorithm whose difficulty increases only algebraically with the size of the graph.

The mathematical significance of this paper rests largely

Edmonds (1967):

We say an algorithm is *good* if there is a polynomial function $f(n)$ which, for every positive-integer valued n, is an upper bound on the "amount of work" the algorithm does for any input of "size" n. The concept
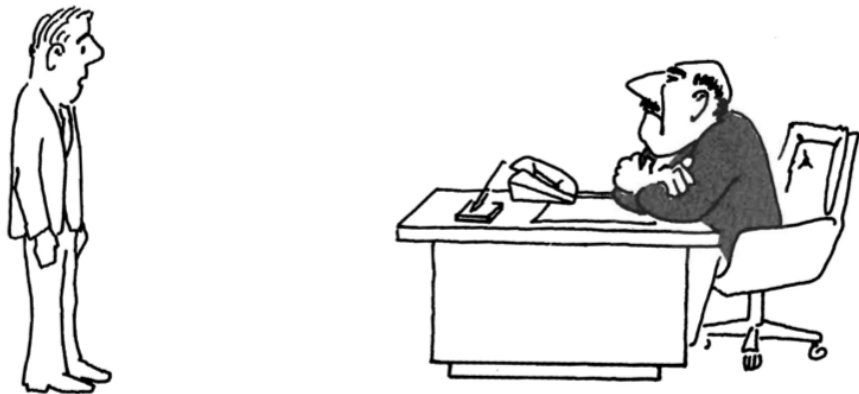
traveling saleman problem [cf. 4]. I conjecture that there is no good algorithm for the traveling saleman problem. My reasons are the same as for any mathematical conjecture: (1) It is a legitimate mathematical possibility, and (2) I do not know.

A good algorithm is known for finding in any graph

Jack Edmonds (1934–)

203

Is There a Good Algorithm for the TSP?

"I can't find an efficient algorithm, I guess I'm just too dumb."

Is There a Good Algorithm for the TSP?

"I can't find an efficient algorithm, because no such algorithm is possible!"

Source: Garey & Johnson, Computers and Intractability, 1979.

Is There a Good Algorithm for the TSP?

"I can't find an efficient algorithm, but neither can all these famous people."

Decision Problems

Most of complexity theory is based on decision problems such as, e. g.:

(Undirected) Hamiltonian Circuit Problem

Given: undirected graph $G = (V, E)$.

Task: decide whether G contains a Hamiltonian circuit.

Definition 8.1.

i A decision problem is a pair $\mathcal{P} = (X, Y)$. The elements of X are called instances of $\mathcal{P}$, the elements of $Y \subseteq X$ are the yes-instances, those of $X \setminus Y$ are no-instances.

ii An algorithm for a decision problem (X, Y) decides for a given $x \in X$ whether $x \in Y$.

Example. For Hamiltonian Circuit, X is the set of all (undirected) graphs and $Y \subset X$ is the subset of graphs containing a Hamiltonian circuit.

Further Examples of Decision Problems

(Integer) Linear Programming Problem (decision version)

Given: matrix $A \in \mathbb{Z}^{m \times n}$, vector $b \in \mathbb{Z}^m$.

Task: decide whether there is $x \in \mathbb{R}^n$ ($x \in \mathbb{Z}^n$) with $A \cdot x \geq b$.

Spanning Tree Problem

Given: graph $G = (V, E)$, edge weights $w : E \to \mathbb{Z}$, positive integer k.

Task: decide whether there is a spanning subtree of weight at most k.

Steiner Tree Problem

Given: graph $G = (V, E)$, terminals $T \subseteq V$, edge weights $w : E \to \mathbb{Z}$, positive integer k.

Task: decide whether there is a subtree of G of weight at most k that contains all terminals in T.

Complexity Classes *P* and *NP*

P

The class of all decision problems for which there is a deterministic polynomial time algorithm is denoted by *P*.

Example: The Spanning Tree Problem is in *P*.

NP

A decision problem belongs to the complexity class *NP* if solutions can be verified in polynomial time.

Examples: The Hamiltonian Circuit Problem and all problems listed on the previous slides belong to *NP*.

Nondeterministic Turing Machines

Remarks.

- ▶ The complexity class P consists of all decision problems that can be solved by a deterministic Turing machine in polynomial time.
- ▶ The complexity class NP consists of all decision problems that can be solved by a non-deterministic Turing machine in polynomial time.
    - ▶ Guess a solution and check in polynomial time.
- ▶ NP stands for Non-deterministic Polynomial.

Lemma 8.2.

$P \subseteq NP$.
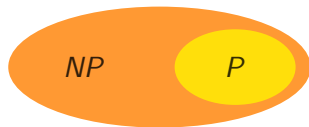
Proof: Deterministic Turing machines are a special case of non-deterministic Turing machines. □

Polynomial Transformations/Reductions

Definition 8.3.

Let $\mathcal{P}_1 = (X_1, Y_1)$ and $\mathcal{P}_2 = (X_2, Y_2)$ be decision problems. We say that $\mathcal{P}_1$ polynomially transforms to $\mathcal{P}_2$ if there exists a function $f : X_1 \to X_2$ computable in polynomial time such that for all $x \in X_1$

$$x \in Y_1 \iff f(x) \in Y_2 \ .$$

Remarks.

- ▶ A polynomial transformation is also called Karp reduction.
- ▶ Polynomial transformations are transitive.

Lemma 8.4.

Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be decision problems. If $\mathcal{P}_2 \in P$ and $\mathcal{P}_1$ polynomially transforms to $\mathcal{P}_2$, then $\mathcal{P}_1 \in P$.

NP-Hardness and *NP*-Completeness

Definition 8.5.

Let $\mathcal{P}$ be an optimization or decision problem.

i $\mathcal{P}$ is *NP*-hard if all problems in *NP* polynomially transform to $\mathcal{P}$.

ii $\mathcal{P}$ is *NP*-complete if in addition $\mathcal{P} \in NP$.

Satisfiability Problem (SAT)

Given: Boolean variables $x_1, \ldots, x_n$ and a family of clauses where each clause is a disjunction of Boolean variables or their negations.

Task: decide whether there is a truth assignment to $x_1, \ldots, x_n$ such that all clauses are satisfied.

Example: $(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3) \wedge (\neg x_1 \vee x_2)$

Cook's Theorem (1971)

Stephen Cook (1939–)

Theorem 8.6.

The Satisfiability problem is *NP*-complete.

Proof idea: SAT is obviously in *NP*. One can show that any non-deterministic Turing machine can be encoded as an instance of SAT. □

Proving NP-Completeness

Lemma 8.7.

Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be decision problems. If $\mathcal{P}_1$ is *NP*-complete, $\mathcal{P}_2 \in NP$, and $\mathcal{P}_1$ polynomially transforms to $\mathcal{P}_2$, then $\mathcal{P}_2$ is *NP*-complete.

Proof: As mentioned above, polynomial transformations are transitive. $\qquad\square$

Integer Linear Programming Problem (ILP)

Given: matrix $A \in \mathbb{Z}^{m \times n}$, vector $b \in \mathbb{Z}^m$.

Task: decide whether there is $x \in \mathbb{Z}^n$ with $A \cdot x \geq b$.

Theorem 8.8.

ILP is *NP*-complete.
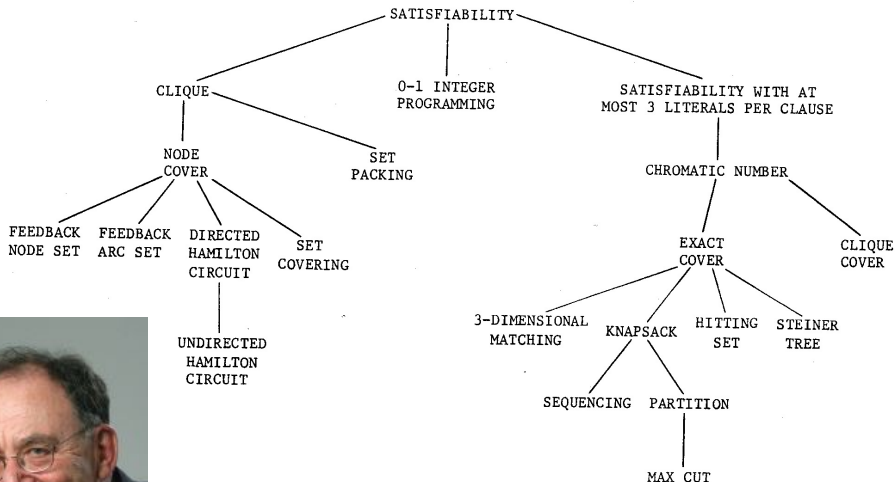
Transformations for Karp's 21 *NP*-Complete Problems

FIGURE 1 - Complete Problems

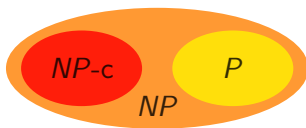Richard M. Karp (1972). *Reducibility Among Combinatorial Problems*
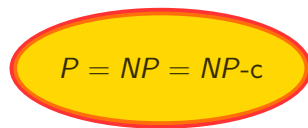
P vs. NP

> **Theorem 8.9.**
> If a decision problem $\mathcal{P}$ is *NP*-complete and $\mathcal{P} \in P$, then $P = NP$.

Proof: See definition of *NP*-completeness and Lemma 8.4. ☐

There are two possible szenarios for the shape of the complexity world:

scenario A   scenario B

- ▶ It is widely believed that $P \neq NP$, i. e., scenario A holds.

- ▶ Deciding whether $P = NP$ or $P \neq NP$ is one of the seven millenium prize problems established by the Clay Mathematics Institute in 2000.

P vs NP Problem

Suppose that you are organizing housing accommodations for a group of four hundred university students. Space is limited and only one hundred of the students will receive places in the dormitory. To complicate matters, the Dean has provided you with a list of pairs of incompatible students, and requested that no pair from this list appear in your final choice. This is an example of what computer scientists call an NP-problem, since it is easy to check if a given choice of one hundred students proposed by a coworker is satisfactory (i.e., no pair taken from your coworker's list also appears on the list from the Dean's office), however the task of generating such a list from scratch seems to be so hard as to be completely impractical. Indeed, the total number of ways of choosing one hundred students from the four hundred applicants is greater than the number of atoms in the known universe! Thus no future civilization could ever hope to build a supercomputer capable of solving the problem by brute force; that is, by checking every possible combination of 100 students. However, this apparent difficulty may only reflect the lack of ingenuity of your programmer. In fact, one of the outstanding problems in computer science is determining whether questions exist whose answer can be quickly checked, but which require an impossibly long time to solve by any direct procedure. Problems like the one listed above certainly seem to be of this kind, but so far no one has managed to prove that any of them really are so hard as they appear, i.e., that there really is no feasible way to generate an answer with the help of a computer. Stephen Cook and Leonid Levin formulated the P (i.e., easy to find) versus NP (i.e., easy to check) problem independently in 1971.

Image credit: on the left, Stephen Cook by Jiří Janíček (cropped).   CC BY-SA 3.0

Rules:

Rules for the Millennium Prizes

Related Documents:

📄 Official Problem Description

📄 Minesweeper

Related Links:

Lecture by Vijaya Ramachandran

215

Complexity of Linear Programming

▶ As discussed in Chapter 4, so far no variant of the simplex method has been shown to have a polynomial running time.

▶ Therefore, the complexity of Linear Programming remained unresolved for a long time.

▶ Only in 1979, the Soviet mathematician Leonid Khachiyan proved that the so-called ellipsoid method earlier developed for nonlinear optimization can be modified in order to solve LPs in polynomial time.

▶ In November 1979, the New York Times featured Khachiyan and his algorithm in a front-page story.

▶ Details can, e. g., be found in the book of Bertsimas & Tsitsiklis (Chapter 8) or in the book *Geometric Algorithms and Combinatorial Optimization* by Grötschel, Lovász & Schrijver (Springer, 1988).

An Approach to Difficult Problems

Mathematicians disagree as to the ultimate practical value of Leonid Khachiyan's new technique, but concur that in any case it is an important theoretical accomplishment.

Mr. Khachiyan's method is believed to offer an approach for the linear programming of computers to solve so-called "traveling salesman" problems. Such problems are among the most intractable in mathematics. They involve, for instance, finding the shortest route by which a salesman could visit a number of cities without his path touching the same city twice.

Each time a new city is added to the route, the problem becomes very much more complex. Very large numbers of variables must be calculated from large numbers of equations using a system of linear programming. At a certain point, the compexity becomes so great that a computer would require billions of years to find a solution.

In the past, "traveling salesmen" problems, including the efficient scheduling of airline crews or hospital nursing staffs, have been solved on computers using the "simplex method" invented by George B. Dantzig of Stanford University.

As a rule, the simplex method works well, but it offers no guarantee that after a certain number of computer steps it will always find an answer. Mr. Khachiyan's approach offers a way of telling right from the start whether or not a problem will be soluble in a given number of steps.

Two mathematicians conducting research at Stanford already have applied the Khachiyan method to develop a program for a pocket calculator, which has solved problems that would not have been possible with a pocket calculator using the simplex method.

Mathematically, the Khachiyan approach uses equations to create imaginary ellipsoids that encapsulate the answer, unlike the simplex method, in which the answer is represented by the intersections of the sides of polyhedrons. As the ellipsoids are made smaller and smaller, the answer is known with greater precision. — MALCOLM W. BROWNE