Advanced Topics in Theoretical Computer Science

Part 5: Complexity (Part 4)

7.02.2018

Viorica Sofronie-Stokkermans

Universität Koblenz-Landau

e-mail: sofronie@uni-koblenz.de

Contents

- Recall: Turing machines and Turing computability
- Register machines (LOOP, WHILE, GOTO)
- Recursive functions
- The Church-Turing Thesis
- Computability and (Un-)decidability
- Complexity

Until now

- Time and space complexity: PTIME, NTIME, PSPACE, NSPACE
- Complexity classes: P, NP, PSPACE
- Complexity classes for functions
- Polynomial time reducibility
- Complete and hard problems
- Examples of NP-complete problems

Examples of NP-complete problems

Examples of NP-complete problems:

1. Is a logical formula satisfiable? (SAT, CNF-SAT, 3-CNF-SAT))
2. Does a graph contain a clique of size k?
3. Rucksack problem
4. Can a graph be colored with three colors?
5. Is a (un)directed graph hamiltonian?
6. Multiprocessor scheduling

Summary

Examples of NP-complete problems:

1. Is a logical formula satisfiable? (SAT)
2. Does a graph contain a clique of size k?
3. Rucksack problem
4. Can a graph be colored with three colors?
5. Is a (un)directed graph hamiltonian?
6. Multiprocessor scheduling

Examples of NP-complete problems

Definition (k-colorability) A undirected graph is k-colorable if every node can be colored with one of k colors such that nodes connected by an edge have different colors.

$L_{\mathsf{Color_k}}$: the language consisting of all undirected graphs which are colorable with at most k colors.

Examples of NP-complete problems

COLOR $= \{(G, k) \mid G$ undirected graph that can be colored with k colors$\}$

COLOR is NP complete

Proof: Exercise. *Hint:*

(1) Prove that the problen is in NP.

(2) Let $F = C_1 \wedge \cdots \wedge C_k$ in 3-CNF containing propositional variables $\{x_1, \dots, x_m\}$.
Let $G = (V, E)$ be an undirected graph, that is defined as follows:

$$V = \{C_1, \dots, C_k\} \cup \{x_1, \dots, x_m\} \cup \{\overline{x_1}, \dots, \overline{x_m}\} \cup \{y_1, \dots, y_m\}$$

$$\begin{aligned} E = &\{(x_i, \overline{x_i}), (\overline{x_i}, x_i) \mid i \in \{1, ..., m\}\} \cup \{(y_i, y_j) \mid i \neq j\} \cup \\ &\{(y_i, x_j), (x_j, y_i) \mid i \neq j\} \cup \{(y_i, \overline{x_j}), (\overline{x_j}, y_i) \mid i \neq j\} \cup \\ &\{(C_i, x_j), (x_j, C_i) \mid x_j \text{ not in } C_i\} \cup \{(C_i, \overline{x_j}), (\overline{x_j}, C_i) \mid \overline{x_j} \text{ not in } C_i\} \end{aligned}$$

Use G to prove 3-CNF-SAT $\preceq_{\text{pol}}$ k-colorability.

Examples of NP-complete problems

COLOR $= \{(G, k) \mid G$ undirected graph that can be colored with k colors$\}$

> COLOR is NP-complete

Detailed proof: Available online from the website

(file: k-coloring-np-complete-proof.pdf)

3-colorability $= \{G \mid G$ undirected graph that can be colored with 3 colors$\}$

> 3-colorability is NP-complete

(for a proof see e.g. https://cgi.csc.liv.ac.uk/ igor/COMP309/3CP.pdf)

Examples of NP-complete problems

Examples of NP-complete problems:

1. Is a logical formula satisfiable? (SAT)
2. Does a graph contain a clique of size k?
3. Rucksack problem
4. Can a graph be colored with three colors?
5. Is a (un)directed graph hamiltonian?
6. Multiprocessor scheduling

Examples of NP-complete problems

Definition (Hamiltonian-path)

Path along the edges of a graph which visits every node exactly once.

Examples of NP-complete problems

Definition (Hamiltonian-cycle)
Path along the edges of a graph which visits every node exactly once and is a cycle.

$L_{\text{Ham,undir}}$: the language consisting of all undirected graphs which contain a Hamiltonian cycle

Examples of NP-complete problems

Definition (Hamiltonian-cycle)
Path along the edges of a graph which visits every node exactly once and is a cycle.

$L_{\text{Ham,undir}}$: the language consisting of all undirected graphs which contain a Hamiltonian cycle

$L_{\text{Ham,dir}}$: the language consisting of all directed graphs which contain a Hamiltonian cycle

NP-completeness: again reduction from 3-CNF-SAT.

Examples of NP-complete problems

Theorem. The problem whether a directed graph contains a Hamiltonian cycle is NP-complete.

Proof. (1) The problem is in NP: Guess a permutation of the nodes; check that they form a Hamiltonian cycle (in polynomial time).

(2) The problem is NP-hard. Reduction from 3-CNF-SAT.

$F = (L_1^1 \vee L_2^1 \vee L_3^1) \wedge \cdots \wedge (L_1^k \vee L_2^k \vee L_3^k)$

Construct $f(F) = G$ such that G contains a Hamiltonian cycle iff F satisfiable.

The details can be found in Erk & Priese, "Theoretische Informatik", p.466-471.

Examples of NP-complete problems

Examples of NP-complete problems:

1. Is a logical formula satisfiable? (SAT)
2. Does a graph contain a clique of size k?
3. Rucksack problem
4. Can a graph be colored with three colors?
5. Is a (un)directed graph hamiltonian?
6. Multiprocessor scheduling

Examples of NP-complete problems

Definition (Multiprocessor scheduling problem)

A scheduling problem consists of:

- n processes with durations $t_1, \ldots, t_n$
- m processors
- a maximal duration (deadline) D

The scheduling problem has a solution if there exists a distribution of processes on the processors such that all processes end before the deadline D.

L_{schedule} : the language consisting of all solvable scheduling problems

Other complexity classes

Co-NP

co-NP is the class of all laguages for which the complement is in NP

Example:
$L_{\text{tautologies}} = \{w \mid w \text{ is a tautology in propositional logic}\}$

Theorem. $L_{\text{tautologies}}$ is in co-NP.

Proof. The complement of $L_{\text{tautologies}}$ is the set of formulae whose negation is satisfiable, thus in NP.

It is not known whether NP = co-NP

PSPACE

Definition (PSPACE-complete, PSPACE-hard)
A language L is PSPACE-hard (PSPACE-difficult) if every language L' in PSPACE is reducible in polynomial time to L.

A language L is PSPACE-complete if:
- $L \in PSPACE$
- L is PSPACE-hard

Quantified Boolean Formulae

Syntax: Extend the syntax of propositional logic by allowing quantification over propositional variables.

Semantics:
$(\forall P)F \mapsto F[P \mapsto 1] \wedge F[P \mapsto 0]$
$(\exists P)F \mapsto F[P \mapsto 1] \vee F[P \mapsto 0]$

PSPACE

A fundamental PSPACE problem was identified by Stockmeyer and Meyer in 1973.

Quantified Boolean Formulas (QBF)

Given: A well-formed quantified Boolean formula

$$F = (Q_1 P_1) \ldots (Q_n P_n) G(P_1, \ldots, P_n)$$

where G is a Boolean expression containing the propositional variables $P_1, \ldots, P_n$ and Q_i is $\exists$ or $\forall$.

Question: Is F true?

(Does it evaluate to 1 if we use the evaluation rules above?)

PSPACE

Example

F propositional formula with propositional variables $P_1, \dots, P_n$

F is satisfiable iff $\exists P_1 \dots \exists P_n F$ is true.

PSPACE

Example

F propositional formula with propositional variables $P_1, \ldots, P_n$

F is satisfiable iff $\exists P_1 \ldots \exists P_n F$ is true.

If we have alternations of quantifiers it is more difficult to check whether a QBF is true.

PSPACE

Theorem QBF is PSPACE complete

Proof (Idea only)

(1) QBF is in PSPACE: we can try all possible assignments of truth values one at a time and reusing the space (2^n time but polynomial space).

(2) QBF is PSPACE complete. We can show that every language L' in PSPACE can be polymomially reduced to QBF using an idea similar to that used in Cook's theorem (we simulate a polynomial space bounded computation and not a polynomial time bounded computation).

The structure of PSPACE

NP vs. Co-NP

co-NP is the class of all laguages for which the complement is in NP

Example:

$L_{\text{tautologies}} = \{w \mid w \text{ is a tautology in propositional logic}\}$ is in co-NP.

NP vs. Co-NP

co-NP is the class of all laguages for which the complement is in NP

Example:

$L_{\text{tautologies}} = \{w \mid w \text{ is a tautology in propositional logic}\}$ is in co-NP.

Informally

$L \in NP$ iff there exists a language $L' \in P$ and a $k \geq 0$ s.t. for all $w \in \Sigma^*$:

$w \in L$ iff $\exists$ c (witness) of lenght polynomial in $|w|$ and s.t. $\langle w, c \rangle \in L'$

(can use c to check in PTIME that $w \in L$)

NP vs. Co-NP

co-NP is the class of all laguages for which the complement is in NP

Example:
$L_{\text{tautologies}} = \{w \mid w \text{ is a tautology in propositional logic}\}$ is in co-NP.

Informally

$L \in$ NP iff there exists a language $L' \in$ P and a $k \geq 0$ s.t. for all $w \in \Sigma^*$:

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ and s.t. $\langle w, c\rangle \in L'$
(can use c to check in PTIME that $w \in L$)

$L \in$ co-NP iff the complement of L is in NP (with test language L')

$w \in L$ iff $\forall c$ of lenght polynomial in $|w|$, $\langle w, c\rangle \notin L'$
(can use c to check in PTIME that $w \in L$)

NP vs. Co-NP

co-NP is the class of all laguages for which the complement is in NP

Example:

$L_{\text{tautologies}} = \{w \mid w \text{ is a tautology in propositional logic}\}$ is in co-NP.

Informally

$L \in$ NP iff there exists a PTIME deterministic verifyer M s.t. for all $w \in \Sigma^*$:

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ and s.t. $M(w, c) = 1$

$L \in$ co-NP iff the complement of L is in NP (with test language L')

$w \in L$ iff $\forall c$ of lenght polynomial in $|w|$, $M(w, c) = 1$.

The structure of PSPACE

... Beyond NP

The structure of PSPACE

Idea: (M PTIME deterministic verifyer)

NP

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ s.t. $M(w, c) = 1$.

co-NP

$w \in L$ iff $\forall c$ of lenght polynomial in $|w|$, s.t. $M(w, c) = 1$.

Σ_2^p

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ s.t.
$\forall d$ of lenght polynomial in $|w|$, $M(w, c, d) = 1$

The structure of PSPACE

Idea: (M PTIME deterministic verifyer)

NP

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ s.t. $M(w, c) = 1$.

co-NP

$w \in L$ iff $\forall c$ of lenght polynomial in $|w|$, s.t. $M(w, c) = 1$.

Σ_2^p

$w \in L$ iff $\exists\ c$ (witness) of lenght polynomial in $|w|$ s.t.
$\forall d$ of lenght polynomial in $|w|$, $M(w, c, d) = 1$

Example: QBF with one quantifier alternation

$\Sigma_2 SAT = \{F = \exists P_1 \dots P_n \forall Q_1 \dots Q_m \overline{F}(P_1, \dots, P_n, Q_1, \dots Q_n) \mid F \text{ true}\}$

The structure of PSPACE

Remarks

- in fact, $\Sigma_2 SAT$ is complete for Σ_2^p
- more alternations lead to a whole hierarchy
- all of it is contained in PSPACE

The structure of PSPACE

For $i \geq 1$,a language L is in Σ_i^p if there exists a PTIME deterministic verifyer M such that:

$$
\begin{aligned}
w \in L \quad \text{iff} \quad & \exists u_1 \text{ of lenght polynomial in } |w| \\
& \forall u_2 \text{ of lenght polynomial in } |w| \\
& \dots \\
& Q_i u_i \text{ of lenght polynomial in } |w| \\
& \text{such that } M(w, u_1, \dots, u_i) = 1
\end{aligned}
$$

where Q_i is $\exists$ if i is odd and $\forall$ otherwise.

The polynomial hierarchy is the set $PH = \bigcup_{i \geq 1} \Sigma_i^p$

$\Pi_i^p = \text{co-}\Sigma_i^p = \{CL \mid L \in \Sigma_i^p\}$

The structure of PSPACE

Formal definition (main ideas)

Extend the notion of polynomial reducibility:

Nondeterministic Turing Machine with an oracle: NTM + oracle tape

– makes initial guess

– consult an oracle

Informally: NOTM for problem P: nondeterministic algorithm with a subroutine for P.

The structure of PSPACE

Extend the notion of polynomial reducibility:

Nondeterministic Turing Machine with an oracle: NTM + oracle tape

– makes initial guess

– consult an oracle

Informally: NOTM for problem P: nondeterministic algorithm with a subroutine for P.

defines a so-called (polynomial time) nondeterministic Turing reduction

The structure of PSPACE

The polynomial hierarchy

$P^Y = \{L \mid$ there exists a language $L' \in Y$ such that $L \preceq_{\text{pol}} L'\}$

$NP^Y = \{L \mid$ there exists a language $L' \in Y$ such that there exists a nondeterministic Turing reduction from L to $L'\}$

$\Sigma_0^p = \Pi_0^p = \Delta_0^p = P.$

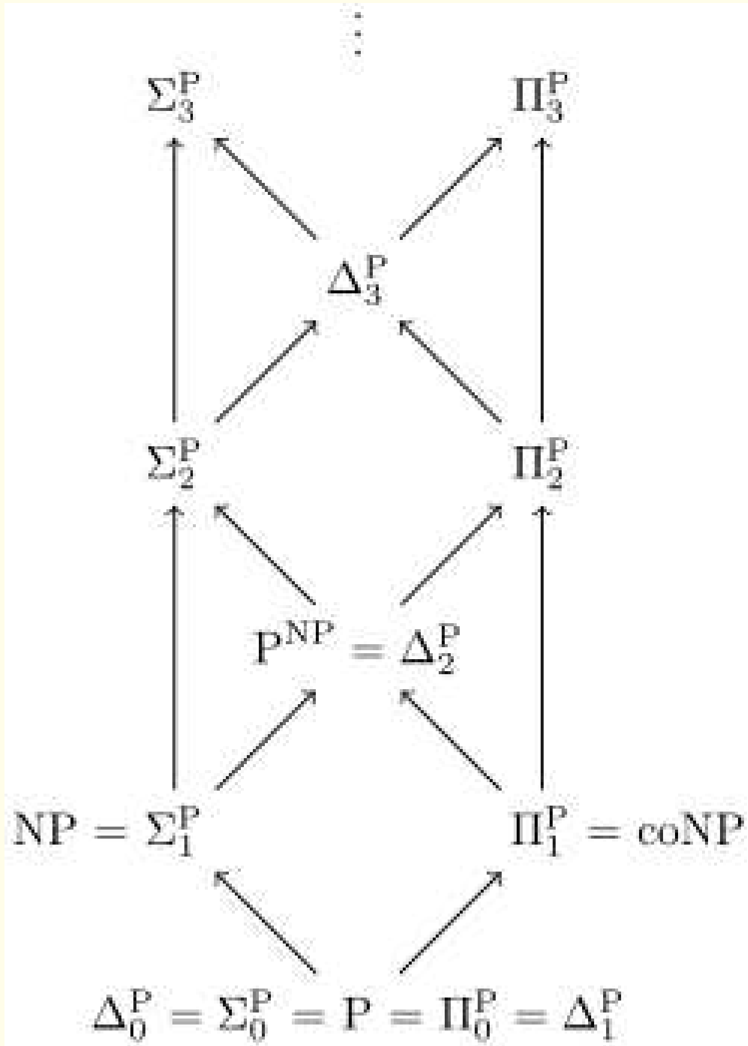
$\Delta_{k+1}^p = P^{\Sigma_k^p}$
$\Sigma_{k+1}^p = NP^{\Sigma_k^p}$
$\Pi_{k+1}^p = \text{co-}NP^{\Sigma_k^p}$

$\Pi_1^p = \text{co-NP}^P = \text{co-NP}$; $\Sigma_1^p = NP^P = NP$; $\Delta_1^p = P^P = P$.

$\Delta_2^p = P^{NP}$; $\Sigma_2^p = NP^{NP}$

The structure of PSPACE

PSPACE



The structure of PSPACE

It is an open problem whether there is an i such that $\Sigma_i^p = \Sigma_{i+1}^p$.

This would imply that $\Sigma_i^p = PH$: the hierarchy collapses to the i-th level.

Most researchers believe that the hierarchy does not collapse.

If NP = P then PH = P, i.e. the hierarchy collapses to P.

The structure of PSPACE

A complete problem for Σ_k^{P} is satisfiability for quantified Boolean formulas with k alternations of quantifiers which start with an existential quantifier sequence (abbreviated QBF_k or $QSAT_k$).

(The variant which starts with $\forall$ is complete for Π_k^{P}).

Beyond PSPACE

EXPTIME, NEXPTIME

DEXPTIME, NDEXPTIME

EXPSPACE, ...

Discussion

- In practical applications, for having efficient algorithms polynomial solvability is very important; exponential complexity inacceptable.
- Better hardware is no solution for bad complexity

Question which have not been clarified yet:

- Does parallelism/non-determinism make problems tractable?
- Any relationship between space complexity and run time behaviour?

Other directions in complexity

Parameterized complexity

Pseudopolynomial problems

Approximative and probabilistic algorithms

Motivation

Many important problems are difficult (undecidable; NP-complete; PSPACE complete)

- **Undecidable:** validity of formulae in FOL; termination, correctness of programs
- **NP-complete:** SAT, Scheduling
- **PSPACE complete:** games, market analyzers

Motivation

Possible approaches:

- Identify which part of the input is cause of high complexity
- Heuristic solutions:
  - use knowledge about the structure of problems in a specific application area;
  - renounce to general solution in favor of a good "average case" in the specific area of applications.
- Approximation: approximative solution
  - Renounce to optimal solution in favor of shorter run times.
- Probabilistic approaches:
  - Find correct solution with high probability.
  - Renounce to sure correctness in favor of shorter run times.

(I) Parameterized Complexity

Parameterized complexity is a branch of computational complexity theory that focuses on classifying computational problems according to their inherent difficulty with respect to multiple parameters of the input.

This allows the classification of NP-hard problems on a finer scale.

$\mapsto$ Fixed parameter tractability.

Example: SAT

Assume that the number of propositional variables is a parameter.

A given formula of size m with k variables can be checked by brute force in time $O(2^k m)$

For a fixed number of variables, the complexity of the problem is linear in the length of the input formula.

(I) Parameterized Complexity

Fixed parameter tractability parameter specified: Input of the form (w, k) L is fixed-parameter tractable if the question $(w, k) \in L$? can can be decided in running time $f(k) \cdot p(|w|)$, where f is an arbitrary function depending only on k, and p is a polynomial.

An example of a problem that is thought not to be fixed parameter tractable is graph coloring parameterised by the number of colors.

It is known that 3-coloring is NP-hard, and an algorithm for graph k-colouring in time $f(k)p(n)$ for $k = 3$ would run in polynomial time in the size of the input.

Thus, if graph coloring parameterised by the number of colors were fixed parameter tractable, then $\mathrm{P} = \mathrm{NP}$.

(II) Approximation

Many NP-hard problems have optimization variants

- Example: Clique: Find a possible greatest clique in a graph

... but not all NP-difficult problems can be solved approximatively in polynomial time:

- Example: Clique: Not possible to find a good polynomial approximation (unless P = NP)

(III) Probabilistic algorithms

Idea

- Undeterministic, random computation
- Goal: false decision possible but not probable
- The probability of making a mistake reduced by repeating computations
- 2^{-100} below the probability of hardware errors.

Probabilistic algorithms

Example: probabilistic algorithm for 3-Clique

NB: 3-Clique is polynomially solvable (unlike Clique)

Given: Graph $G = (V, E)$

Repeat the following k times:

- Choose randomly $v_1 \in V$ and $\{v_2, v_3\} \in E$
- Test if v_1, v_2, v_3 build a clique.

Error probability:

$k = (|E| \cdot |V|)/3$: Error probability < 0.5

$k = 100(|E| \cdot |V|)/3$: Error probability $< 2^{-100}$

Overview

- Register machines (LOOP, WHILE, GOTO)
- Recursive functions
- The Church-Turing Thesis
- Computability and (Un-)decidability
- Complexity
- Other computation models

Overview

- Register machines (LOOP, WHILE, GOTO)
- Recursive functions
- The Church-Turing Thesis
- Computability and (Un-)decidability
- Complexity
- Other computation models

Other computation models

- Variations of register machines (one register; two registers)
- Variations of TM; links with register machines
- Reversible computations: e.g. chemical reversibility or reversibility as in physics
- DNA Computing and Splicing
  Computing machines consisting from enzymes and molecules

Other computation models

- Variations of register machines (one register; two registers)
- Variations of TM; links with register machines
- Reversible computations: chemical and psysichal reversibility
- DNA Computing and Splicing
  Computing machines consisting from enzymes and molecules

Variants of automata

- Tree automata
- Automata over infinite words

Variants of automata

Tree automata

Like automata, but deal with tree structures, rather than the strings.

Tree automata are an important tool in computer science:

- compiler construction
- automatic verification of cryptographic protocols.
- processing of XML documents.

Variants of automata

Automata on infinite words (or more generally: infinite objects)

ω-Automata (Büchi automata, Rabin automata, Streett automata, parity automata and Muller automata)

- run on infinite, rather than finite, strings as input.
- Since ω-automata do not stop, they have a variety of acceptance conditions rather than simply a set of accepting states.

Applications: Verification, temporal logic

Look forward

Next semester:

- Seminar: Decision procedures and applications $\mapsto$ emphasis on decidability and complexity results for various application areas.

Various possibilities for BSc/MSc thesis and Forschungspraktika.