P, NP and NP-complete

方嘉聪 2200017849

北京大学

2024 年 5 月 17 日

目录

Deterministic Turing Machine

定义 (Deterministic Turing Machine)

A k-tape TM is a 7-tuple $(Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject})$:

- Q sets of states, Σ input alphabet, Γ tape alphabet
- $\delta : Q \times \Gamma^k \to Q \times \Gamma^k \times \{L, S, R\}^k$ transition function
- q_0 initial state, q_{accept} accepting state, q_{reject} rejecting state

Deterministic Turing Machine

定义 (Deterministic Turing Machine)

A k-tape TM is a 7-tuple $(Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject})$:

- Q sets of states, Σ input alphabet, Γ tape alphabet
- $\delta : Q \times \Gamma^k \to Q \times \Gamma^k \times \{L, S, R\}^k$ transition function
- q_0 initial state, q_{accept} accepting state, q_{reject} rejecting state

定义 (Computation of a TM)

A TM M accepts input w if a sequence of configurations $C_1, C_2, \ldots, C_k$ exists such that:

- C_1 is the start configuration of M on input w;
- each C_i yeilds C_{i+1} by applying δ;
- C_k is an accepting configuration.

Universal Turing Machine

定义 (Universal Turing Machine)

A TM U is a universal TM if it can **simulate** any TM M.

定理

存在一个通用图灵机 U，对于任意图灵机 M 和输入 x, U 能够在 $O(T \log T)$ 时间内输出 $M(x)$.

形式化书写和证明这里略去.

Nondeterministic Turing Machine

定义 (Nondeterministic Turing Machine)

A k-tape NTM is a 7-tuple $(Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject})$:

- Q sets of states, Σ input alphabet, Γ tape alphabet
- $\delta : Q \times \Gamma^k \to 2^{Q \times \Gamma^k \times \{L,S,R\}^k}$ transition function
- q_0 initial state, q_{accept} accepting state, q_{reject} rejecting state

Nondeterministic Turing Machine

定义 (Nondeterministic Turing Machine)

A k-tape NTM is a 7-tuple $(Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject})$:

- Q sets of states, Σ input alphabet, Γ tape alphabet
- $\delta : Q \times \Gamma^k \to 2^{Q \times \Gamma^k \times \{L,S,R\}^k}$ transition function
- q_0 initial state, q_{accept} accepting state, q_{reject} rejecting state

- 直觉上理解, NTM 在每一步都可以选择多种可能的转移, 但只要有一种转移路径能够接受, 则 NTM 接受输入.
- 运行时间: 对于函数 $T : N \to N$, 一个 NDTM N 的运行时间是 $T(n)$ 的, 如果对于任意输入长度为 n 的输入, N 的所有分支都在 $T(n)$ 步内停机.

Class P

定义 (Class DTIME)

设函数 $T : N \to N$, 那么 DTIME($T(n)$) 为所有可以在 $O(T(n))$ 时间内被确定性图灵机判定 (decided) 的语言的集合.

定义 (Class P)

$\mathsf{P} = \bigcup_{c \in N} \mathsf{DTIME}(n^c)$

Class NP

定义 (Class NTIME)

设函数 $T : N \to N$, 那么 NTIME$(T(n))$ 为所有可以在 $O(T(n))$ 时间内被非确定性图灵机判定 (decided) 的语言的集合.

定义 (Class NP)

$\mathsf{NP} = \bigcup_{c \in N} \mathsf{NTIME}(n^c)$

Class NP(Cont.)

定义 (Another Definition of NP)

称一个语言 $L \in$NP, 如果存在一个多项式函数 $P : N \to N$, 和多项式时间的确定性图灵机 M, 使得:

$$\forall x \in \{0,1\}^*,\ x \in L \iff \exists u \in \{0,1\}^{P(|x|)},\ s.t.\ M(x,u) = 1$$

把 u 称为 x 的证书 (certificate), M 称为 L 的验证机 (verifier).

Class NP(Cont.)

定义 (Another Definition of NP)

称一个语言 $L \in$NP, 如果存在一个多项式函数 $P : N \to N$, 和多项式时间的确定性图灵机 M, 使得:

$$\forall x \in \{0,1\}^*,\ x \in L \iff \exists u \in \{0,1\}^{P(|x|)},\ s.t.\ M(x,u) = 1$$

把 u 称为 x 的证书 (certificate), M 称为 L 的验证机 (verifier).

- 直觉上理解, NP 包括了所有可以在多项式时间内验证一个解是否正确的问题.
- 上述两个定义是等价的, 这里不展开证明细节.

Polynomial-time Reduction

定义 (Polynomial-time Reduction)

称一个语言 A 多项式时间归约 (is polynomial time reducible) 到另一个语言 B, 记作 $A \leq_p B$, 如果存在一个多项式时间可计算的函数 $f : \Sigma^* \to \Sigma^*$ 使得:

$$\forall w, w \in A \iff f(w) \in B$$

Polynomial-time Reduction

定义 (Polynomial-time Reduction)

称一个语言 A 多项式时间归约 (is polynomial time reducible) 到另一个语言 B, 记作 $A \leq_p B$, 如果存在一个多项式时间可计算的函数 $f : \Sigma^* \to \Sigma^*$ 使得:

$$\forall w, w \in A \iff f(w) \in B$$

若 $A \leq_p B$:

- $B \in \mathsf{P} \implies A \in \mathsf{P}$
- $A \leq_p B, B \leq_p C \implies A \leq_p C$

直观上: $A \leq_p B$ 意味着 B 比 A 更困难. 设计规约 f 可以理解成设计一个算法, 将 A 的问题转化为 B 的问题.

NP hard and NP complete

定义 (NP-hard)

称一个语言 A 是 NP-hard 的, 如果对于任意 $L \in$NP, $L \leq_p A$.

定义 (NP-complete)

称一个语言 A 是 NP-complete 的, 如果 A 是 NP-hard $\wedge A \in$NP.

NP hard and NP complete

定义 (NP-hard)

称一个语言 A 是 NP-hard 的, 如果对于任意 $L \in$NP, $L \leq_p A$.

定义 (NP-complete)

称一个语言 A 是 NP-complete 的, 如果 A 是 NP-hard $\wedge A \in$NP.

- 直观上, NP-complete 问题是 NP 中最困难的问题.
- 若存在一个 NP-complete 问题 A 可以在多项式时间内被解决, 那么 P=NP.

目录

NPC maps

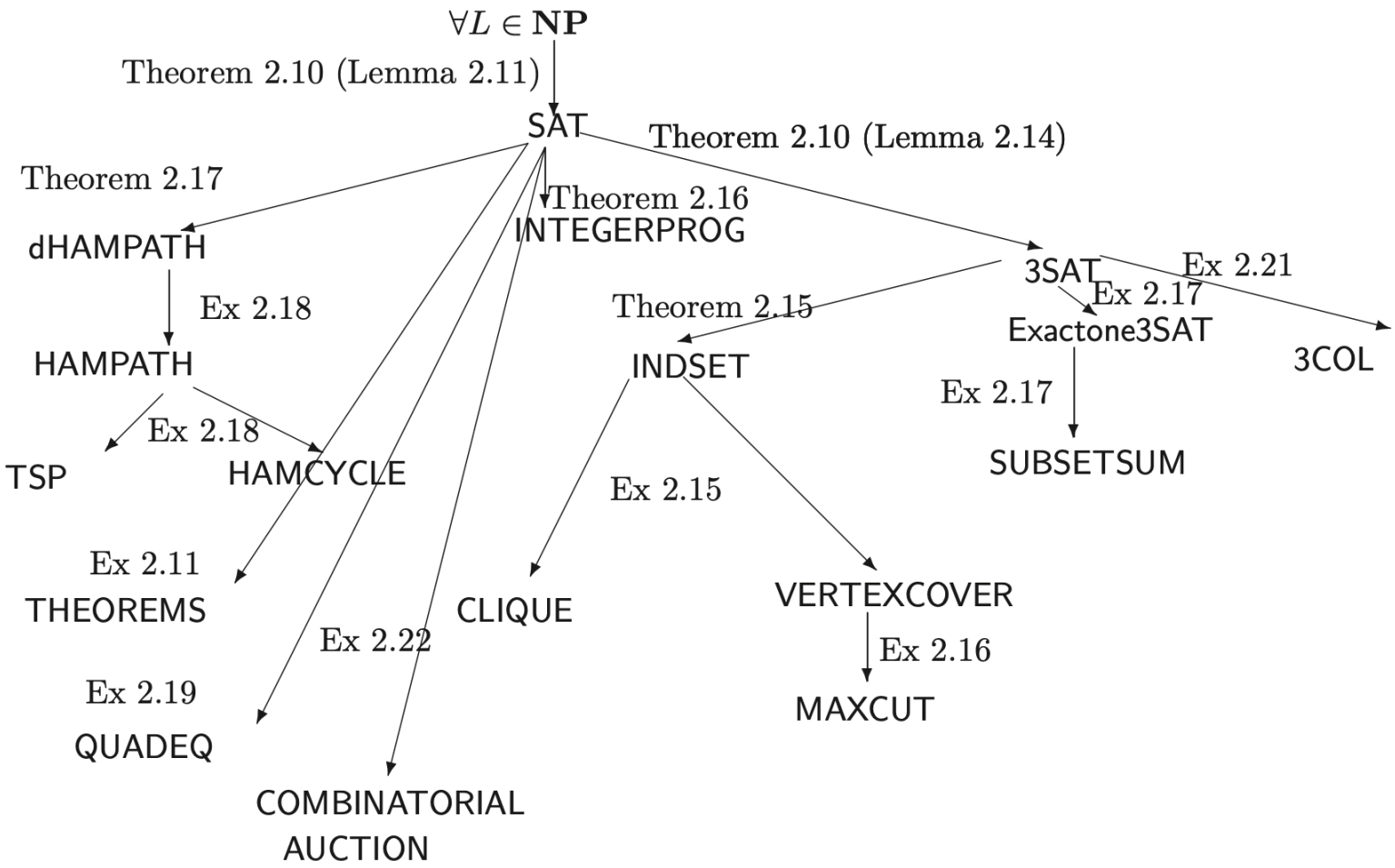


Figure 2.4 Web of reductions between the **NP**-completeness problems described in this chapter and the exercises. Thousands more are known.

Definition

定义 (析取范式 (CNF))

一个布尔表达式 φ, 称 φ 为析取范式 (CNF), 如果 φ 形如:

$$\bigwedge_i \left(\bigvee_j v_{i,j} \right)$$

称 $v_{i,j}$ 为 φ 的变量 (literal), $(\bigvee_j v_{i,j})$ 是一个子句 (clause).

Definition

定义 (析取范式 (CNF))

一个布尔表达式 φ, 称 φ 为析取范式 (CNF), 如果 φ 形如:

$$\bigwedge_i \left(\bigvee_j v_{i,j} \right)$$

称 $v_{i,j}$ 为 φ 的变量 (literal), $(\bigvee_j v_{i,j})$ 是一个子句 (clause).

定义 (SAT and 3SAT)

SAT $= \{$all satisfiable CNF formulas$\}$
3SAT $= \{$all satisfiable 3CNF formulas$\}$.

Cook-Levin Theorem

定理 (Cook-Levin Theorem)

3SAT 是 NP-complete 的.

证明: (1) SAT $\in$ NP 是显然的, 考虑证书为一组赋值即可, 验证机只需验证是否可满足.
(2) SAT 是 NP-hard 的. 这个比较困难. 不加证明的给出一个引理.

引理

对任意一个布尔函数 $f: \{0,1\}^n \to \{0,1\}$, 可以在 $\Theta(2^n)$ 的时间内构造一个 *CNF* F, 使得 $f(x) = 1 \iff F(x) = 1$.

3SAT

SAT $\leq_p$ 3SAT.

- 设 φ 是一个 CNF 公式, 考虑如下规约 f.
- 例如: $(a_1 \vee a_2 \vee a_3 \vee a_4) \to (a_1 \vee a_2 \vee z) \wedge (\bar{z} \vee a_3 \vee a_4)$
- 一般情况下, 考虑一个有 l 个 lieral 的子句 $(a_1 \vee a_2 \vee \cdots \vee a_l)$:

$$(a_1 \vee a_2 \vee z_1) \wedge (\bar{z}_1 \vee a_3 \vee z_2) \wedge \cdots \wedge (\bar{z}_{l-3} \vee a_{l-1} \vee a_l)$$

总共 $l-2$ 个子句.

MAX-SAT

定义 (MAX-SAT)

给定一个 CNF 公式 φ(n 个变量和 m 个子句) 和正整数 k, 如果存在赋值使得 φ 中至少有 k 个子句为真, 则 $\langle \varphi, k \rangle \in$MAX-SAT.

SAT $\leq_p$ MAX-SAT.

- $\forall \varphi \in$ SAT, 令 $k = m$, 则 $\langle \varphi, k \rangle \in$ MAX-SAT.

MAX-SAT $\in$ NP. 考虑证书为一组赋值即可.

INDSET

定义 (INDSET)

给定一个图 G 和正整数 k, 判定是否存在一个大小为 k 的独立集.

3SAT $\leq_p$ INDSET.

- 设 φ 是一个 3CNF 公式 (存在 n 个变量和 m 个子句), 构造一个图 $G = (V, E)$ 如下:
- V: 每个子句 C_i 对于七个点. 这七个点对应 7 种可能的赋值情况.
- E: 对应的赋值情况发生冲突, 则在这两点间连边.

Vertex Cover and CLIQUE

定义

- **Vertex Cover**: 给定一个图 G 和正整数 k, 判定是否存在一个大小不超过 k 的顶点覆盖.
- **CLIQUE**: 给定一个图 G 和正整数 k, 判定是否存在一个大小不小于 k 的团.

有如下的引理:

引理

对任意的无向图 $G = (V, E)$ 和子集 $V' \subseteq V$, 下列命题等价:

- V' 是 G 的一个顶点覆盖.
- $V - V'$ 是 G 的独立集.
- $V - V'$ 是补图 $G_c = (V, E_C)$ 的团.

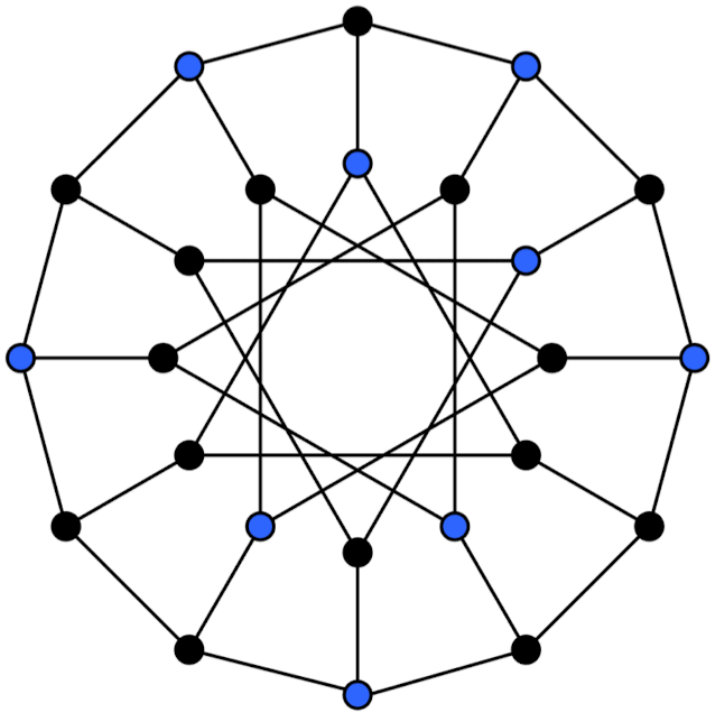
Vertex Cover and CLIQUE(Cont.)

INDSET $\leq_P$ Vertex Cover.

- $f: \langle G, k \rangle \to \langle G', k' \rangle$, 令 $G' = G$, $k' = |V| - k$.

INDSET $\leq_P$ CLIQUE.

- $f: \langle G, k \rangle \to \langle G', k' \rangle$, 令 $G' = G_c$, $k' = |V| - k$.



HAMILTONIAN 相关问题

定义

- **dHAMPATH**: 有向图版本的哈密顿路径.
- **dHAMCYCLE**: 有向图版本的哈密顿回路.
- **uHAMPATH**: 无向图版本的哈密顿路径.
- **uHAMCYCLE**: 无向图版本的哈密顿回路.

下面的证明路径是:

$$\text{3SAT} \leq_p \text{dHAMPATH} \leq_p \text{dHAMCYCLE}$$
$$\text{dHAMPATH} \leq_p \text{uHAMPATH}.$$

TSP

定义 (TSP)

给定一个完全图 G, 每条边有一个距离 d_{ij}, 给定正整数 k, 判定是否存在一个哈密顿回路, 使得总距离不超过 k.

dHAMCYCLE $\leq_p$ TSP.

- 设 $G = (V, E)$, $G' = (V', E,)$, $f : \langle G \rangle \to \langle G', d_{ij}, k \rangle$
- $V' = V$, 如果 $(i, j) \in E$, 则 $d_{ij} = 1$, 否则 $d_{ij} = +\infty$.
- 令 $k = |V|$. 那么有 $G \in$ dHAMPATH $\iff \langle G', d_{ij}, k \rangle \in$ TSP

Karp's 21 Problems

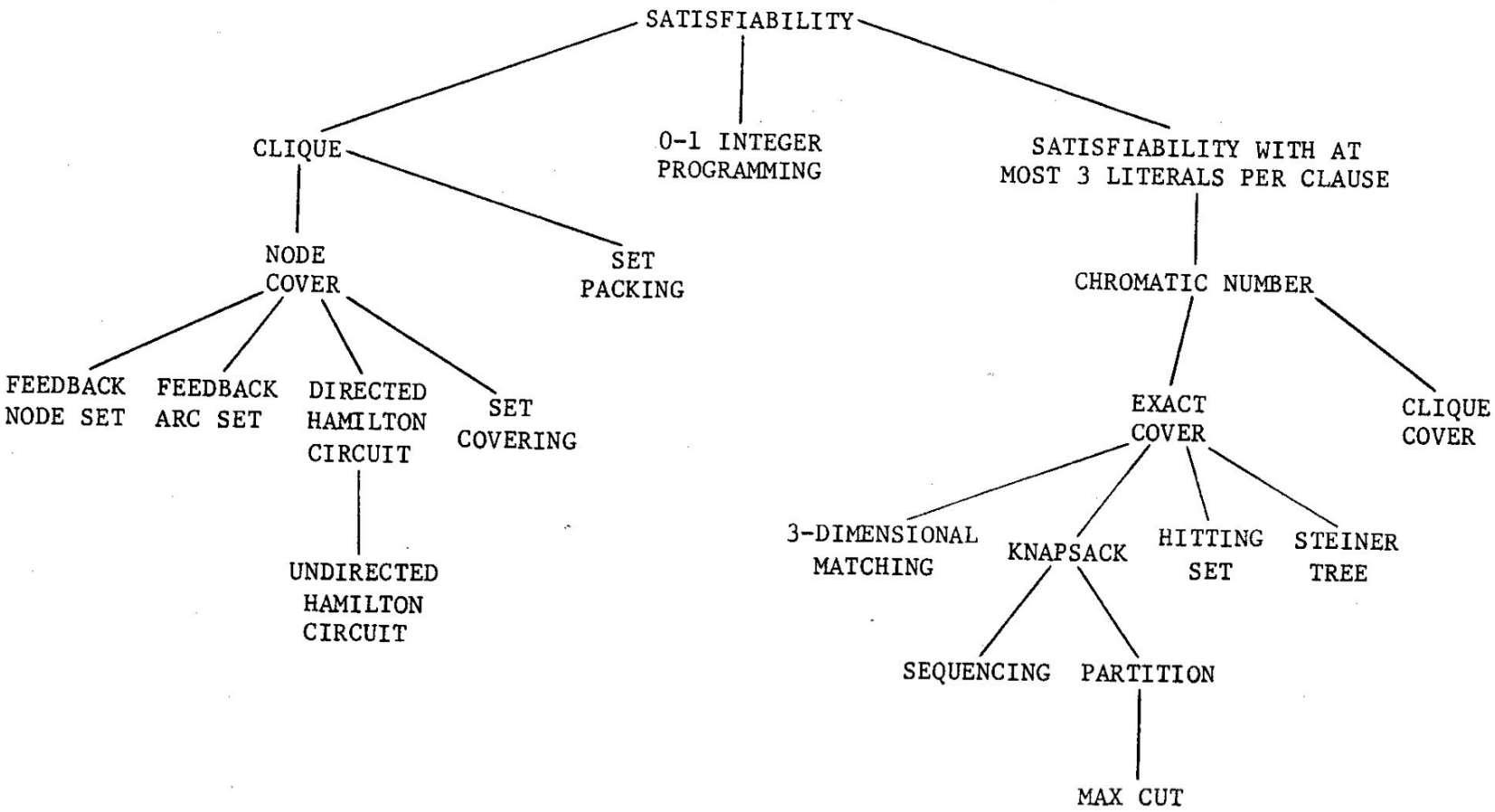


FIGURE 1 - Complete Problems