「解けない問題」を知ろう

保坂 和宏 (東京大学 B2)

第 11 回 JOI 春合宿
2012/03/19

概要

- 計算量に関して
- **P** と **NP**
- **NP** 完全
- 決定不能

- いろいろな問題
- コンテストにおいて

Turing 機械

- コンピュータの計算のモデル
  - 「計算」を数学的に厳密に扱うためのもの
- メモリのテープ (0/1 の列)，ポインタ，機械の内部状態を持ち，規則に従って状態遷移をする
- 本講義では C 言語等で，入力を受け取って何か計算して出力を返すもの，とイメージしてよい
  - 入力は適切な形式で 0/1 の列になっているとする

計算量

- 計算に必要な資源
- 入力のサイズ (0/1 のビット列の長さ) の関数として表される

- 時間計算量
  - 動作するステップ数

- 空間計算量
  - 使用するメモリの量

計算量

- 単位などは実はあまり気にしない
- 「入力サイズが○倍になったときに□倍になる」のような関係が重要

Landau の記号

- 定数倍の差を気にせず大雑把に関数を扱うための記号
- n の関数 $f(n)$, $g(n)$ に対し，ある正定数 c が存在し，十分大きい n に対し $|f(n)| \leqq c\,|g(n)|$ であるとき，$f(n)$ は $O(g(n))$ 程度であるという
  - $f(n) = O(g(n))$ と書いてしまうことが多い
  - 例： $2n^3 + 5n^2 - n + 1000000000 = O(n^3)$, $3^{n+10} = O(3^n)$, $\log_2 n = O(\log_3 n)$

Landau の記号

- $|f(n)| \geqq c\,|g(n)|$ のとき $f(n) = \Omega(g(n))$ と書き，$f(n) = O(g(n))$ かつ $f(n) = \Omega(g(n))$ のとき $f(n) = \Theta(g(n))$ と書く
  - O 記号は「せいぜい定数倍」，Θ 記号は「ちょうど定数倍」

計算量

$O(n)$

$O(n \log n)$

$O(n^2)$

$O(n^{12})$

$O(2^n)$

$O(2^n n^2)$

$O(1.69^n)$

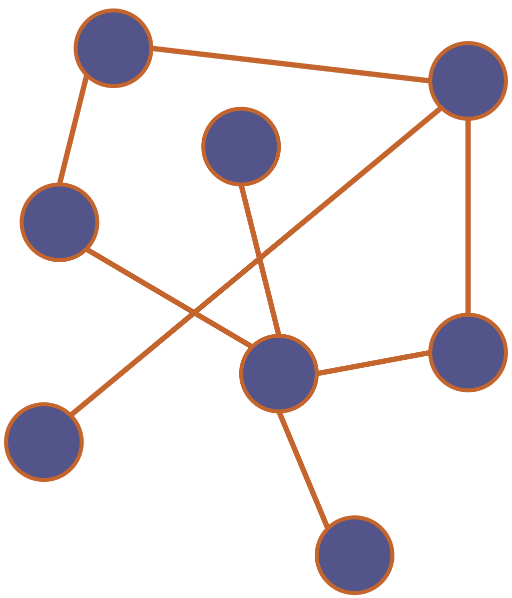
$O(3^n)$

$O(n!)$

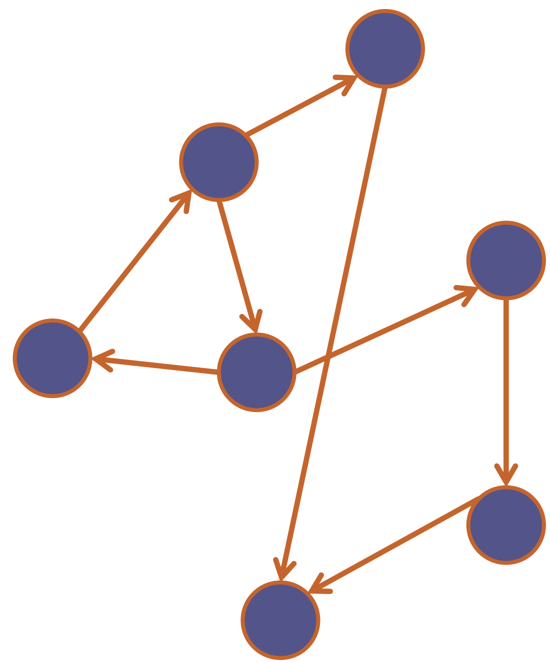
$O(n^n)$

グラフ

無向グラフ

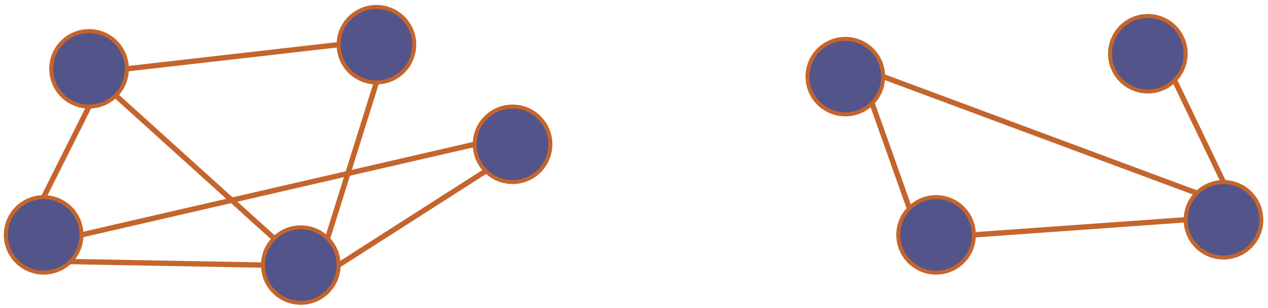


有向グラフ



決定問題

- 与えられた入力に対して，Yes か No かを答える問題
  - 例：与えられたグラフに Hamilton 閉路が存在するか？
    - Hamilton 閉路： すべての頂点をちょうど 1 回ずつ通る閉路



決定問題

- 補問題：決定問題の Yes/No を入れ替えた問題
  - 例：与えられたグラフに Hamilton 閉路が存在しないなら Yes，存在するなら No と答える問題
- ある決定問題とその補問題は異なる問題と考える

決定問題

- 最適化問題：何らかの制約の下で，ある値を最大化あるいは最小化する問題
  - 「その値は○以上になるか？」という形式にすることで決定問題が作れる
  - 元の問題との難易度はだいたい同じ

P

- クラス **P**：ある決定性 Turing 機械で多項式時間で解ける決定問題たち
  - ある多項式 $p(n)$ が存在して，十分大きい n に対し，入力サイズが n なら $p(n)$ ステップ以下で解ける
  - **P** に属する問題の例：グラフが連結であるか判定

- 「効率的に解ける問題」として扱われる
- **P** に属する問題の補問題は **P** に属する

NP

- クラス **NP**：(定義 1) ある非決定性 Turing 機械で多項式時間で解ける決定問題たち
  - 決定性 Turing 機械：ポインタの指す値と機械の内部状態によって次の動きが定まっている
  - 非決定性 Turing 機械：ポインタの指す値と機械の内部状態に対して，次の動きの候補が複数ある

- 正解が Yes なら Yes と出力する可能性があり，正解が No なら必ず No と出力する

NP

- **NP** に属する問題の例：Hamilton 閉路問題
- ある頂点から始めて，辺で繋がった頂点を適当に選んでいき，すべての頂点を回って最初の頂点に戻ってこられたら Yes，失敗したら No
  - Hamilton 閉路が存在するならうまくいけば Yes
  - Hamilton 閉路が存在しないなら必ず No

- 「多項式深さのバックトラックで解ける問題たち」とも言い換えられる

NP

- クラス **NP** : (定義 2) 「正解が Yes である証拠」が与えられたとき，それが正しいかを (決定性 Turing 機械で) 多項式時間で判定できる決定問題たち

- **NP** に属する問題の例 : Hamilton 閉路問題
  - Hamilton 閉路をなす頂点の列を 1 つ出力すれば，それが実際に Hamilton 閉路になっていることを検証するのは多項式時間でできる

NP

- 2 つの定義は同値
- ある問題が **NP** に属しても，その問題の補問題は **NP** に属するとは限らない
- **P** に属する問題は **NP** に属する (**P** ⊂ **NP**)
- 逆の真偽は未解決 (**P ≠ NP 予想**)

帰着

- 2 つの問題 *A*, *B* に対して，「*B* を解くアルゴリズムが存在すればそれを利用して *A* を解くことができる」とき，*A* は *B* に帰着可能という
    - *B* は *A* を「含む」と考えるとわかりやすい
- *A* の入力に対して多項式時間／線形時間の操作で *B* の入力に変換できるとき多項式帰着／線形帰着という
    - *B* は *A* と同等か *A* より難しいと考えられる

NP 完全

- NP 完全問題： **NP** に属する問題であって，**NP** に属するすべての問題から多項式帰着可能な問題
  - **NP** に属するすべての問題を「含む」
  - **NP** で最も「難しい」

- もしある **NP** 完全問題が **P** に属することがわかったら，**P = NP**

充足可能性問題 (SAT)

[入力] $x_1, \ldots, x_n$ を True または False の値をとる変数，y_{ij} は x_k または $\neg x_k$ として，

$$
\begin{aligned}
&(y_{11} \vee y_{12} \vee \ldots \vee y_{1*}) \\
\wedge\ &(y_{21} \vee y_{22} \vee \ldots \vee y_{2*}) \\
\wedge\ &\ldots \\
\wedge\ &(y_{m1} \vee y_{m2} \vee \ldots \vee y_{m*})
\end{aligned}
$$

という形の論理式

[出力] $x_1, \ldots, x_n$ に True または False を適切に割り当てて式の値を True にできるなら Yes

充足可能性問題 (SAT)

[例] $(x_1 \lor \neg x_2 \lor x_3) \land (\neg x_1 \lor x_2) \land (\neg x_3)$

- x_1 = True, x_2 = True, x_3 = False とすればよいので Yes

[例] $(x_1 \lor x_2) \land (x_1 \lor \neg x_2) \land (\neg x_1)$

- この式を True にすることはできないので No

充足可能性問題 (SAT)

NP 完全

[入力] $x_1, \ldots, x_n$ を True または False の値をとる変数，y_{ij} は x_k または $\neg x_k$ として，

$$
\begin{aligned}
&(y_{11} \vee y_{12} \vee \ldots \vee y_{1*}) \\
\wedge\ &(y_{21} \vee y_{22} \vee \ldots \vee y_{2*}) \\
\wedge\ &\ldots \\
\wedge\ &(y_{m1} \vee y_{m2} \vee \ldots \vee y_{m*})
\end{aligned}
$$

という形の論理式

[出力] $x_1, \ldots, x_n$ に True または False を適切に割り当てて式の値を True にできるなら Yes

NP 完全

充足可能性問題 (SAT)

- **NP** に属する
  - Yes の証拠として，True/False の割り当てを与えればよい

充足可能性問題 (SAT)

NP 完全

- Cook が **NP** 完全性を証明 (1971)
- 非決定性 Turing 機械で多項式時間で終了するプログラムをすべてシミュレーションできる
  - 長さ n の入力に対して $p(n)$ ステップ以内に終了する場合，長さ n の入力を受け取ったら ($p(n)$ の多項式) 個の変数をおいてうまく ($p(n)$ の多項式) 個の必要十分な節を作ることができる
    - 使いうるメモリは $2p(n) + 1$ 個のどれか，状態数はプログラムによって定まっている定数

NP 完全

- Karp が SAT の **NP** 完全性を用いて 20 個の問題の **NP** 完全性を証明 (1972)
  - いくつかを講義後半で見ていきます

NP 困難

- **NP 困難問題**： ~~NP に属する問題であって，~~ **NP** に属するすべての問題から多項式帰着可能な問題
  - **NP** に属さない問題を含む (決定問題以外も含む)
  - 任意の **NP** 問題に対し同等かそれより難しい

決定不能

- 決定不能問題：「任意の入力に対して有限時間で停止し正しい出力をする」ような Turing 機械が存在しない問題
  - **NP** 困難の極端な例

停止問題

決定不能

[入力] プログラム *f*, *f* への入力 *x*
[出力] *f* が入力 *x* に対して有限時間で停止するなら Yes

- 入力 *x* は任意の 0/1 列とする
  - 「入力がグラフである」のような問題であれば，「*x* がグラフを表さなかったらすぐに停止する」というようにすればこの問題の適用ができる

決定不能

停止問題

- 停止問題を解くプログラム G が存在したとする
  - G に入力 f, x を与えた結果を $G(f, x)$ と書く
- 次のプログラム H を考える：
  - プログラム f を入力として受け取り，
    - $G(f, f)$ = Yes ならば無限ループを起こす
    - $G(f, f)$ = No ならば停止する
- $H(H)$ を考える (H に入力 H を与える) と：
  - $G(H, H)$ = Yes なら $H(H)$ が停止しないので矛盾
  - $G(H, H)$ = No なら $H(H)$ が停止するので矛盾

各種の問題

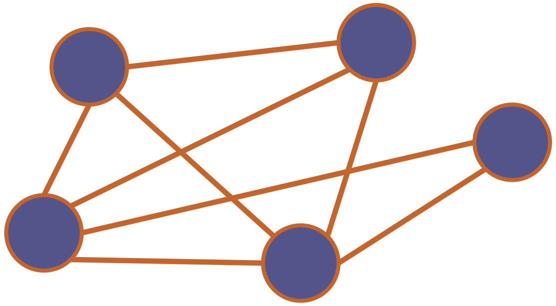
- ここからいろいろな問題を見ていきます
- **P**？ **NP** 完全？ 決定不能？ その他？ と考えてみましょう
  - 直感も大切
  - 証明は概略だけ述べたり省略したりするので後でなぜかも考えてみましょう

NP 完全

クリーク問題

[入力] 無向グラフ G, 整数 k
[出力] G にサイズ k のクリークが存在するなら Yes

- サイズ k のクリーク：互いに隣接している k 個の頂点

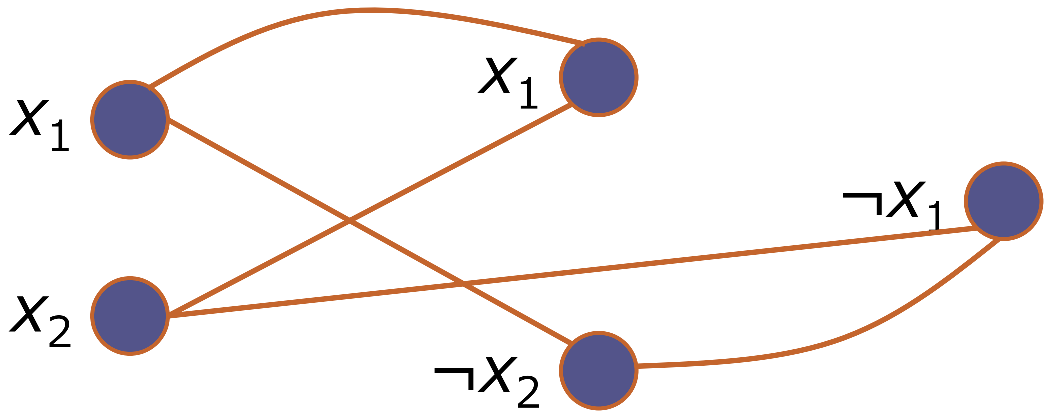


クリーク問題

NP 完全

- クリーク問題は SAT を含む

[例] $(x_1 \lor x_2) \land (x_1 \lor \neg x_2) \land (\neg x_1)$



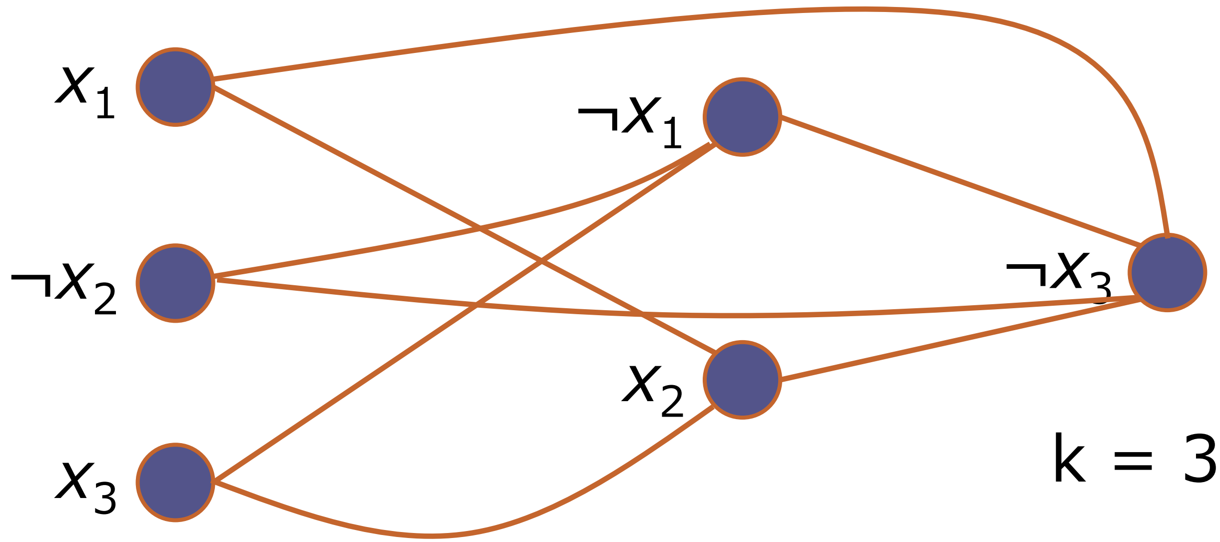
k = 3

NP 完全

クリーク問題

- クリーク問題は SAT を含む

[例] $(x_1 \lor \neg x_2 \lor x_3) \land (\neg x_1 \lor x_2) \land (\neg x_3)$



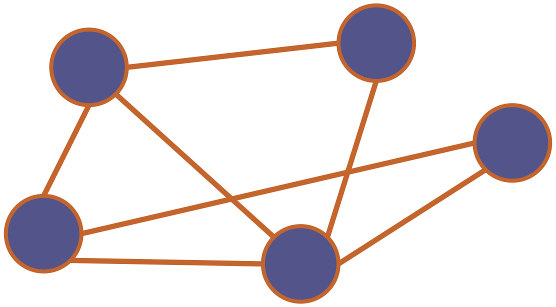
k = 3

頂点被覆問題

NP 完全

[入力] 無向グラフ G，整数 k
[出力] G にサイズ k の頂点被覆が存在するか？

- サイズ k の頂点被覆：k 個の頂点の集合 S であって，G のすべての辺について少なくとも一方の端点が S に含まれるようなもの



NP 完全

頂点被覆問題

- 頂点被覆問題はクリーク問題を含む
  - 実は線形等価

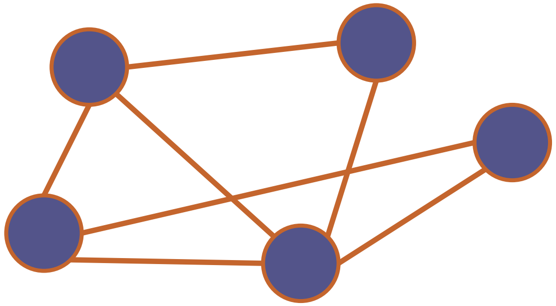
- クリーク問題の入力 (G, k) に対し，G' を G の補グラフとし，k' = (G の頂点数) - k とおくと，G のサイズ k のクリークと G' のサイズ k' の頂点被覆が対応 (互いに補集合)
  - よって (G', k') について頂点被覆問題を解けばクリーク問題も解ける

P

"無向閉路除去問題"

[入力] 無向グラフ G, 整数 k

[出力] G から k 本の辺を除去して閉路をすべてなくせるか？



"無向閉路除去問題"

P

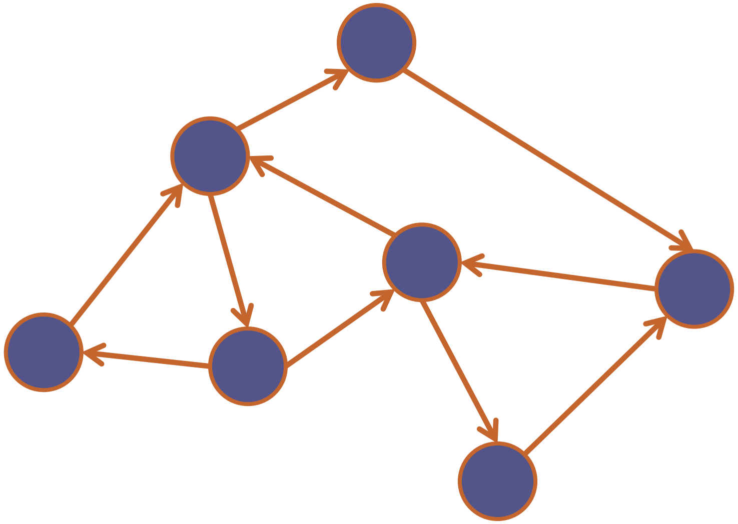
- 各連結成分に対して，(辺数) ≦ (頂点数) - 1 にはしなければならない
- それは可能
  - 連結なグラフには全域木が存在するため
- 頂点数 n，辺数 m として $O(n + m)$

NP 完全

"有向閉路除去問題"

[入力] 有向グラフ G，整数 k

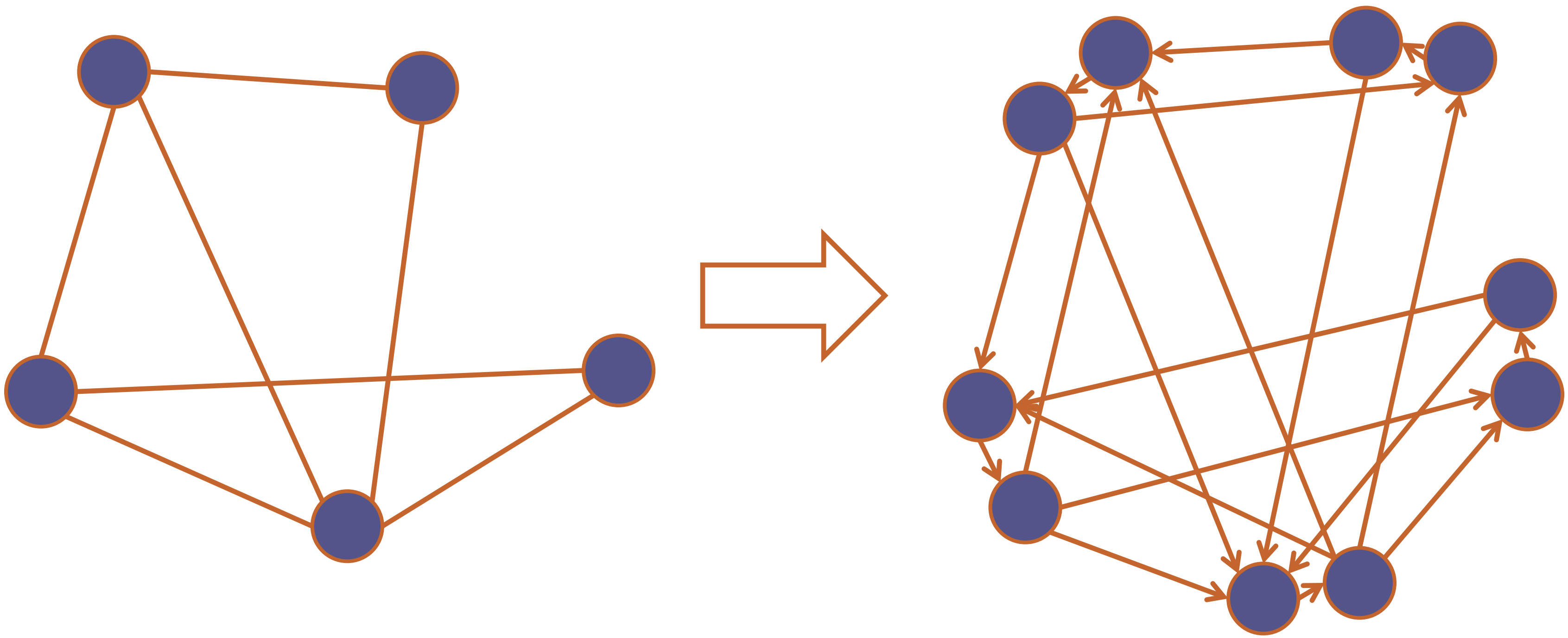
[出力] G から k 本の辺を除去して閉路をすべてなくせるか？



"有向閉路除去問題"

NP 完全

- "有向閉路除去問題" は頂点被覆問題を含む



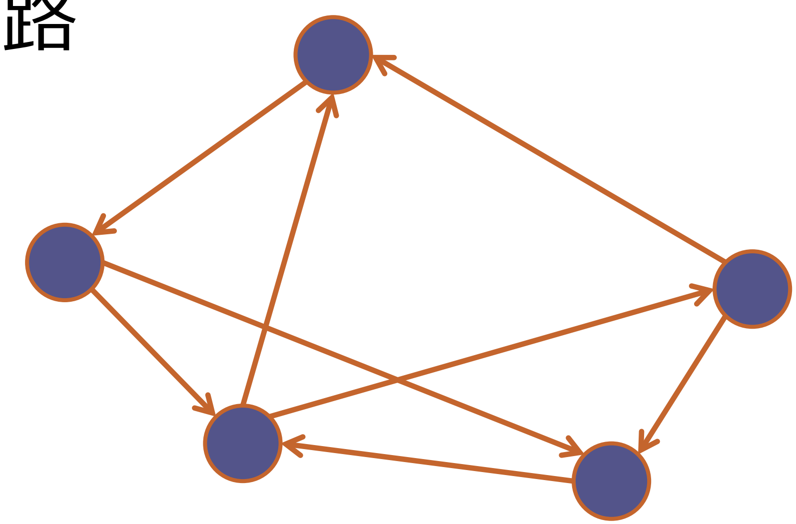
k = 3　　　　k = 3

有向 Hamilton 閉路問題

NP 完全

[入力] 有向グラフ G
[出力] G に Hamilton 閉路が存在するか？

- Hamilton 閉路：すべての頂点をちょうど 1 回ずつ通る閉路



有向 Hamilton 閉路問題

NP 完全

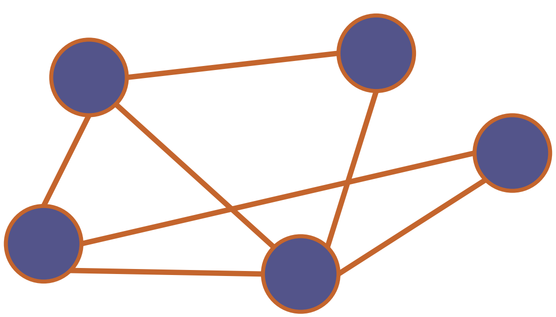
- 有向 Hamilton 閉路問題は頂点被覆問題を含む
  - 帰着が結構大変なので省略します

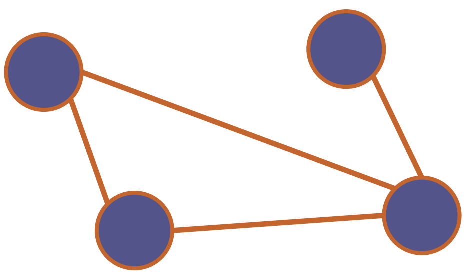
無向 Hamilton 閉路問題

NP 完全

[入力] 無向グラフ *G*
[出力] *G* に Hamilton 閉路が存在するか？

- Hamilton 閉路：すべての頂点をちょうど 1 回ずつ通る閉路

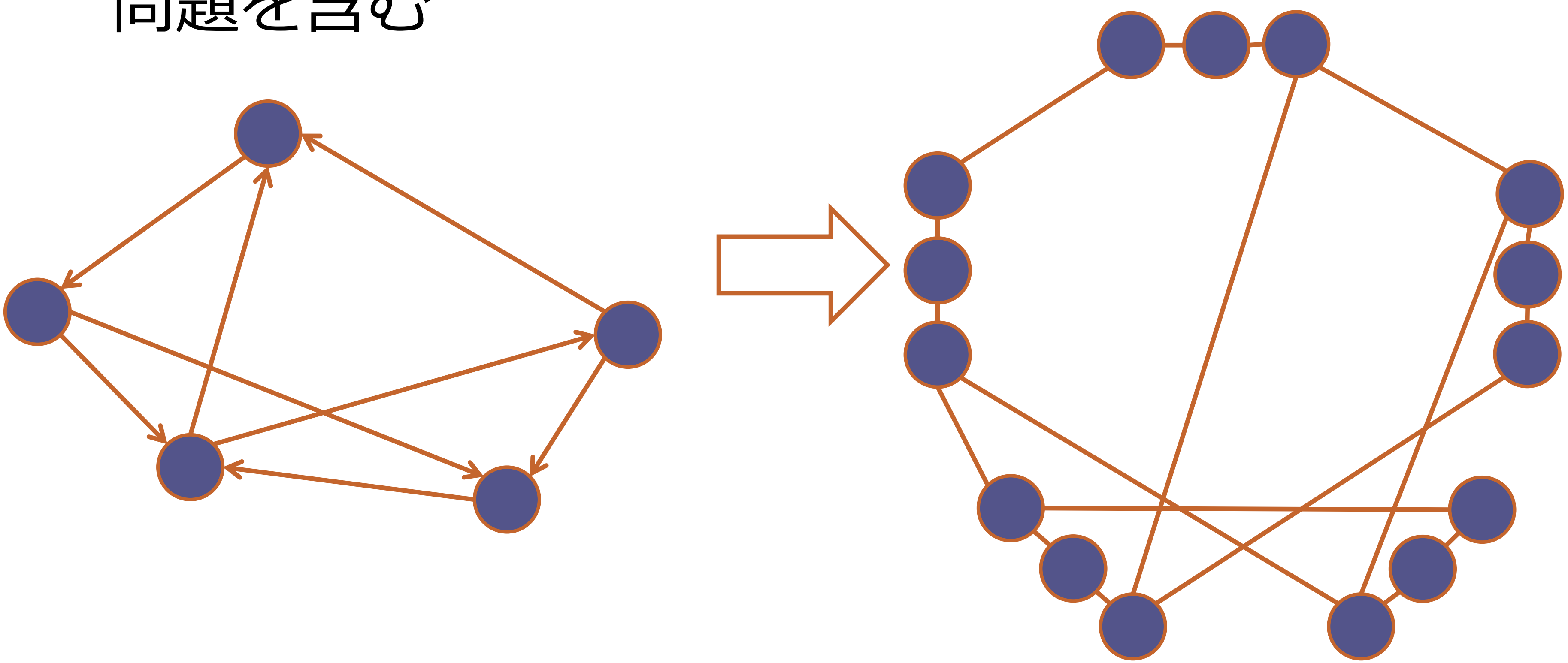




無向 Hamilton 閉路問題

NP 完全

- 無向 Hamilton 閉路問題は有向 Hamilton 閉路問題を含む

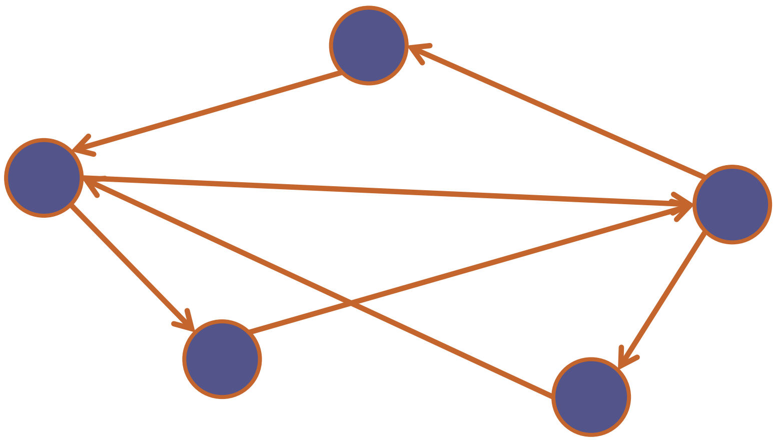


P

有向 Euler 回路問題

[入力] 有向グラフ *G*
[出力] *G* に Euler 回路が存在するか？

- Euler 回路：すべての辺をちょうど 1 回ずつ通る回路



有向 Euler 回路問題

P

- 「一筆書き問題」
- Euler 回路の存在には以下の 2 つが必要：
  - 各頂点に対して (出次数) = (入次数)
  - 孤立点を除いて連結
- これが十分条件であることも知られている

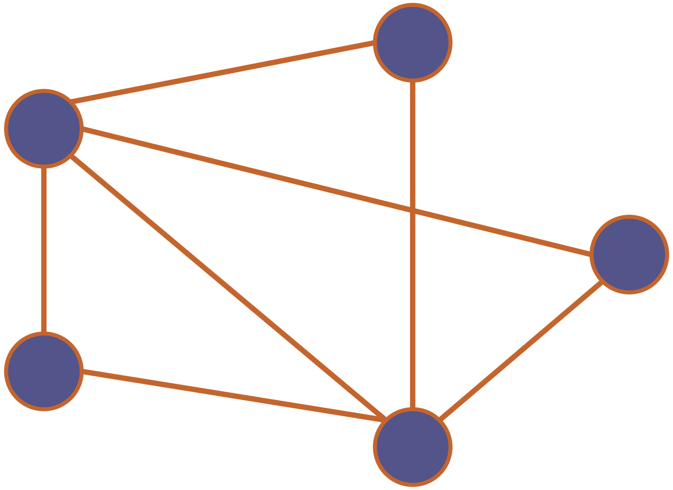
- 頂点数 n，辺数 m として $O(n + m)$

P

無向 Euler 回路問題

[入力] 無向グラフ *G*
[出力] *G* に Euler 回路が存在するか？

- Euler 回路：すべての辺をちょうど 1 回ずつ通る回路



無向 Euler 回路問題

P

- 「一筆書き問題」
- Euler 回路の存在には以下の 2 つが必要：
  - 各頂点に対して次数が偶数
  - 孤立点を除いて連結
- これが十分条件であることも知られている

- 頂点数 n，辺数 m として $O(n + m)$

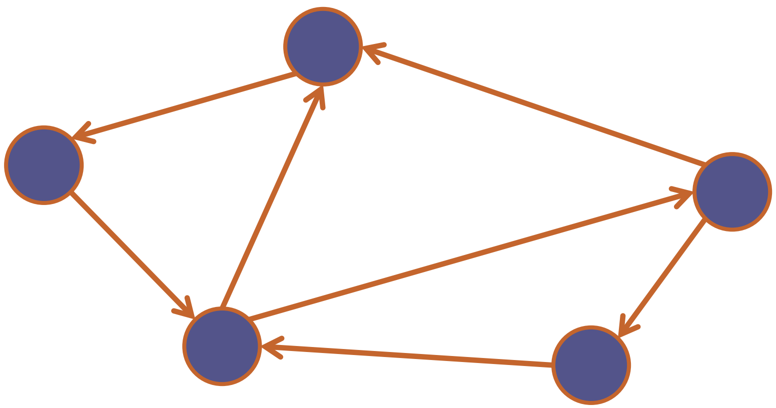
NP 完全

"強連結部分グラフ問題"

[入力] 有向グラフ G, 整数 k

[出力] G の辺を k 本だけ残した部分グラフで強連結なものは存在するか？

- 強連結：どの 2 頂点間にもパスが存在



"強連結部分グラフ問題"

NP 完全

- "強連結部分グラフ問題" は有向 Hamilton 閉路問題を含む

- G の頂点数を n とする
- 頂点数 n，辺数 n の強連結グラフは長さ n の閉路に限られる
  - 辺数 n 未満では強連結にならない
- よって有向 Hamilton 閉路問題は "強連結部分グラフ問題" の $k = n$ の場合

2-充足可能性問題 (2-SAT)

P

[入力] $x_1, \ldots, x_n$ を True または False の値をとる変数，y_{ij} は x_k または $\neg x_k$ として，

$$
\begin{aligned}
& (y_{11} \lor y_{12}) \\
\land\ & (y_{21} \lor y_{22}) \\
\land\ & \ldots \\
\land\ & (y_{m1} \lor y_{m2})
\end{aligned}
$$

という形の論理式

[出力] $x_1, \ldots, x_n$ に True または False を適切に割り当てて式の値を True にできるか？

2-充足可能性問題 (2-SAT)

P

- 節 $a \vee b$ を ¬a ⇒ b, ¬b ⇒ a と書き換えて, $x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n$ を頂点とする有向グラフを作る
- x_i と $\neg x_i$ が同一の強連結成分に属さないことが必要十分

- 変数数 n, 節数 m として $O(n + m)$

3-充足可能性問題 (3-SAT)

NP 完全

[入力] $x_1, \ldots, x_n$ を True または False の値をとる変数，y_{ij} は x_k または $\neg x_k$ として，

$$
\begin{aligned}
&(y_{11} \lor y_{12} \lor y_{13}) \\
\land\ &(y_{21} \lor y_{22} \lor y_{23}) \\
\land\ &\ldots \\
\land\ &(y_{m1} \lor y_{m2} \lor y_{m3})
\end{aligned}
$$

という形の論理式

[出力] $x_1, \ldots, x_n$ に True または False を適切に割り当てて式の値を True にできるか？

NP 完全

3-充足可能性問題 (3-SAT)

- SAT は 3-SAT に多項式変換可能
- $(a_1 \lor a_2 \lor \ldots \lor a_l)$ $(l \geqq 4)$ を次の節たちに置き換える ($u_1, \ldots, u_{l-3}$ は新しい変数)
  - $(a_1 \lor a_2 \lor u_1)$
  - $(\neg u_1 \lor a_3 \lor u_2)$

    …
  - $(\neg u_{l-4} \lor a_{l-2} \lor u_{l-3})$
  - $(\neg u_{l-3} \lor a_{l-1} \lor a_l)$

3-充足可能性問題 (3-SAT)

NP 完全

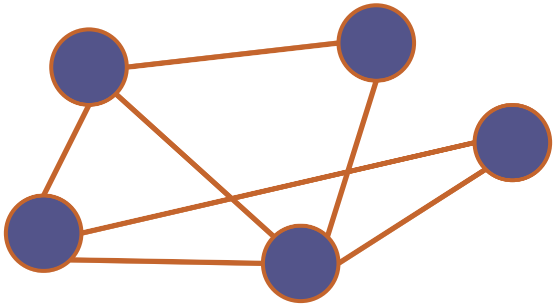
- SAT は 3-SAT に多項式変換可能
- 長さ 2 以下の節の処理は簡単
  - 3-SAT と "節の長さ高々 3 の SAT" はほぼ同じ
  - "節の長さ高々 3 の SAT" を 3-SAT として用いることが多い

頂点彩色問題

NP 完全

[入力] 無向グラフ G，整数 k
[出力] G の頂点を k 色で彩色できるか？

- 頂点彩色：各頂点に色をつけ，どの辺の両端の頂点も異なる色であるようにすること



頂点彩色問題

NP 完全

- 頂点彩色問題は 3-SAT を含む
  - 帰着が結構大変なので省略します

NP 完全

"集合分割問題"

[入力] 有限集合 U の部分集合 $S_1, \ldots, S_r$
[出力] S_i のうちいくつかで U 全体をちょうど覆うことはできるか？

$U = \{ 1, 2, 3 \}$
$S_1 = \{ 1, 2 \}, S_2 = \{ 2, 3 \}, S_3 = \{ 3 \}$

$U = \{ 1, 2, 3, 4 \}$
$S_1 = \{ 1, 3 \}, S_2 = \{ 1, 2, 4 \}, S_3 = \{ 1, 4 \}$

NP 完全

"集合分割問題"

- "集合分割問題" は頂点彩色問題を含む

- 頂点彩色問題の入力 (G, k) に対し,
  - U = { G の頂点 } ∪ { G の辺 } ∪ { (u, e, f) | u: 頂点, e: u に接続する辺, $1 \leqq f \leqq k$}
  - S_i は以下の形のもの：
    - 各 u, f に対し, $\{ u \} \cup \{ (u, e, f) \}$
    - 各 e, f_1, f_2 $(f_1 \neq f_2)$ に対し, $\{ e \} \cup \{ (u, e, g_1) \mid g_1 \neq f_1) \cup \{ (u, e, g_2) \mid g_2 \neq f_2 \}$

NP 完全

ナップサック問題

[入力] r 個の品物の容量 a_i と価値 b_i，ナップサックの容量 c (すべて整数)，整数 d

[出力] 品物をいくつか選んで，容量の和が c 以下で価値の和を d 以上にできるか？

$a_1 = 3$	$a_2 = 6$	$a_3 = 8$
$b_1 = 4$	$b_2 = 7$	$b_3 = 9$

$c = 10$

$d = 10$

ナップサック問題

NP 完全

- ナップサック問題は集合分割問題を含む

- U = { 0, 1, ..., n - 1 } とする
- 集合分割問題の入力 (S_i) に対し，
  - $a_i = b_i = \sum_{j \in S_i} (r + 1)^j$
  - $c = d = \sum_{0 \leqq j < n} (r + 1)^j$

- DP で $O(r\ c)$ * (整数演算のコスト) 時間で解けるがこれは入力長に対する多項式ではない

Post の対応問題

[入力] 0/1 の列たち $a_1, b_1, a_2, b_2, \ldots, a_r, b_r$
[出力] $a_{i1}a_{i2}\ldots a_{ik} = b_{i1}b_{i2}\ldots b_{ik}$ となるような $i_1, i_2, \ldots, i_k$ ($k > 0$) は存在するか？ (i_j は同じ値を複数回含んでもよい)

Post の対応問題

[例] 0/111, 1/011, 11/1

- Yes
  - 11 11 11 0 1 11
  - 1 1 1 111 011 1

[例] 0/01, 1/00, 011/0, 1/110

- No

[例] 0/001, 001/1, 1/0

- Yes (やってみてください)

決定不能

Post の対応問題

[入力] 0/1 の列たち $a_1, \ldots, a_r, b_1, \ldots, b_r$
[出力] $a_{i1}a_{i2}\ldots a_{ik} = b_{i1}b_{i2}\ldots b_{ik}$ となるような $i_1, i_2, \ldots, i_k$ $(k > 0)$ は存在するか？ (i_j は同じ値を複数回含んでもよい)

決定不能

Post の対応問題

- Yes の証拠として $i_1, i_2, \ldots, i_k$ が使えない
  - k が r の多項式程度で収まるとは限らないので，多項式時間で検証できない
- 実は Turing 機械のシミュレーションが行える
- よって停止問題を含む

"ジャッジ問題"

決定不能

[入力] 任意の入力に対して停止し 0 または 1 を出力することが保証されているプログラム f

[出力] f は必ず 0 を出力するか？

決定不能

"ジャッジ問題"

- "ジャッジ問題" を解くプログラム G が存在したとする
- 次のプログラム *H* を考える：
  - Post の対応問題の入力 $((a_i), (b_i))$ を受け取り，次のプログラム *h* を考える：
    - 整数 *n* を受け取り，Post 対応問題 $((a_i), (b_i))$ が $k \leqq n$ で解をもつなら 1，そうでないなら 0 を出力
      - 全探索させれば停止するものが作れる
  - *G*(*h*) が Yes/No なら No/Yes を返す
- H は Post の対応問題を正しく解くので矛盾

コンテストにおいて

- 例：文字列がたくさん与えられるので，それらすべてを使った巡回しりとりが存在するかどうか答えよ

- 文字列を頂点として，文字列 s の次に文字列 t が使えるとき s から t へ辺をはった有向グラフ
  - Hamilton 閉路問題になる
  - **NP** 完全！

コンテストにおいて

- 例：文字列がたくさん与えられるので，それらすべてを使った巡回しりとりが存在するかどうか答えよ

- 文字を頂点として，文字 a で始まり文字 b で終わる文字列に対応させて a から b へ辺をはった有向グラフ
  - Euler 閉路問題になる
  - 線形時間！

コンテストにおいて

- 「問題 *A* を解くには，問題 *B* を解けばよい」
- *B* が **NP** 困難問題の場合は多項式時間で解くことは諦めましょう
  - もしかすると **P = NP** かもしれないとはいえ，短時間の競技中にできるなんて考えてはダメです

コンテストにおいて

- 「問題 *A* を解くには，問題 *B* を解けばよい」

- *B* が **NP** 困難問題だったら
  - 指数時間が想定されている可能性
  - *A* から変換した問題に制約がある場合
    - 特に，特殊なグラフに関して *B* を解けばよい，ということはよくあります
  - 方針が誤りなので考え直す

コンテストにおいて

- 「問題 *A* を解くには，問題 *B* を解かなければならない」
  - 例：クエリが複数個来る問題でクエリ 1 個の場合
  - 例：グラフの問題で木の場合

- *B* が **NP** 困難の場合は *A* の想定解法はまず間違いなく指数時間以上です
- *B* の解法を考えることは *A* を解くために役立つ
- *B* が手に負えないと感じた場合は諦めましょう

今回扱った問題

- 充足可能性問題 (SAT)
- 停止問題
- クリーク問題
- 頂点被覆問題
- "無向閉路除去問題"
- "有向閉路除去問題"
- 有向 Hamilton 閉路問題
- 無向 Hamilton 閉路問題
- 有向 Euler 閉路問題
- 無向 Euler 閉路問題
- "強連結部分グラフ問題"
- 2-充足可能性問題 (2-SAT)
- 3-充足可能性問題 (3-SAT)
- 頂点彩色問題
- "集合分割問題"
- ナップサック問題
- Post の対応問題
- "ジャッジ問題"