アルゴリズムとデータ構造

第26週目

担当　情報システム部門　徳光政弘
2025年12月23日

今日の内容

- 問題のクラス
- P、NPの違い
- P、NP困難、NP完全の関係
- 計算理論（あるいは計算論）へのみちしるべ

問題のクラス

- 問題の難しさをアルゴリズムの性能で分類を試みる
- 問題の複雑さ　アルゴリズムの最悪時間計算量
- 時間計算量の大きな分類
  - 多項式
  - 指数、階乗時間

問題のクラス

表 13.1　問題とアルゴリズムの時間計算量

問　題	アルゴリズム	最悪時間計算量
未ソートのデータに対する探索	線形探索 (アルゴリズム 4.1)	$O(n)$
ソート済みのデータに対する探索	2 分探索法 (アルゴリズム 4.2)	$O(\log n)$
ソート	ヒープソート，マージソート (アルゴリズム 5.5，アルゴリズム 7.3)	$O(n \log n)$
分割ナップサック問題	グリーディ法 (アルゴリズム 8.1)	$O(n \log n)$
部分和問題	分枝限定法 (アルゴリズム 9.2)	$O(n2^n)$
最短経路問題	ダイクストラ法 (アルゴリズム 10.1)	$O(n^2)$

問題の複雑さの比較

- 配列から目的のデータを探す問題を考える
- 少なくとも未ソートのデータに対する探索の方が複雑

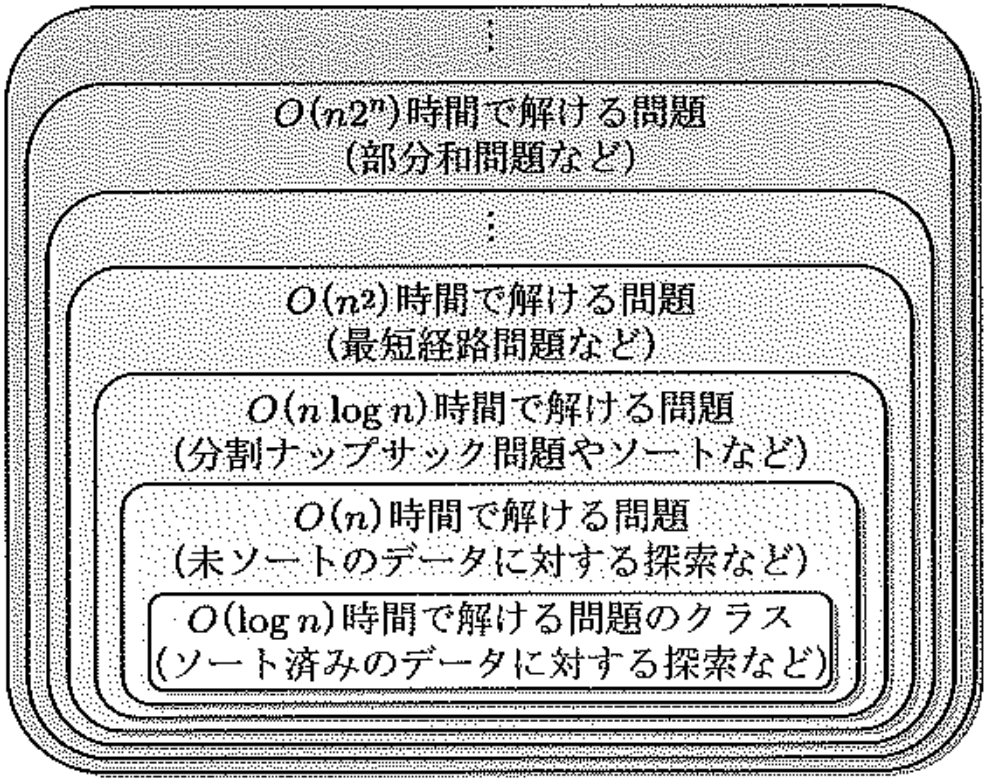
$$\left[\begin{array}{l}\text{未ソートのデータに対する}\\\text{探索の複雑さ}\end{array} = O(n)\right] \geq \left[\begin{array}{l}\text{ソート済みのデータに対する}\\\text{探索の複雑さ}\end{array} = O(\log n)\right]$$

- 少なくとも部分和問題の方が複雑

$$\left[\begin{array}{c}\text{分割ナップサック問題の複雑さ}\\ = O(n \log n)\end{array}\right] \leq \left[\begin{array}{c}\text{部分和問題の複雑さ}\\ = O(n2^n)\end{array}\right]$$

問題のクラスの階層構造

- 同じ問題でも計算量が速いアルゴリズムと遅いアルゴリズムがある
- 例えばソートアルゴリズム



クラスPとクラスNP

- クラスP　多項式時間のアルゴリズムが存在する問題
- クラスNP　問題に対する解の検証が多項式時間で確かめることができる問題

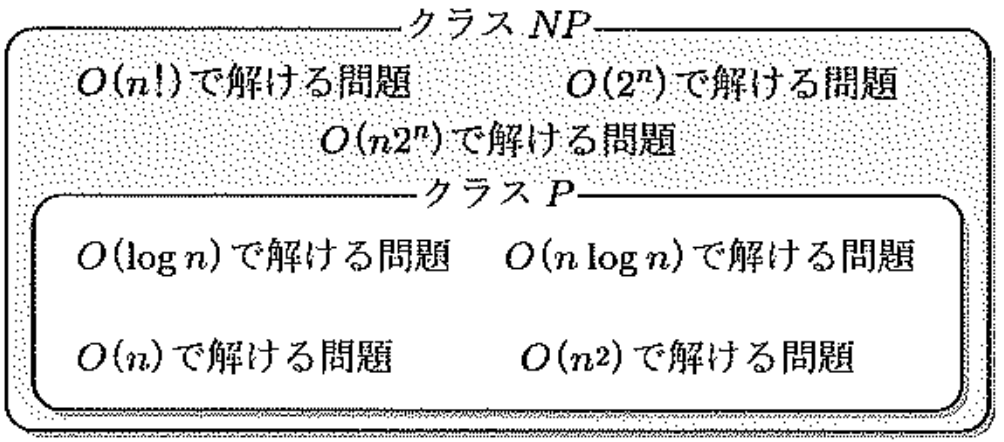


図 13.2　クラス P とクラス NP の関係

クラスNP

- 一般に問題を解くアルゴリズムは指数時間、階乗で表される問題が占める。
- 指数時間 $O(n2^n)$ 、階乗 $O(n!)$
- nが大きくなると（n=100程度）、現代の実用的な技術では解けなくなる
- 量子コンピュータが発展すれば可能性が広がる

部分和問題

【問題 9.1】 部分和問題

$\{x_1, x_2, \ldots, x_n\}$ という n 個の正の実数の集合と，s という正の実数が与えられたとする．このとき，$\{x_1, x_2, \ldots, x_n\}$ の中からその和がちょうど s になる実数の選び方を求めよ．

計算量 $O(n2^n)$
例:1, 2, 3, 4, 5, 6, 7, 8, 9, 10
和:15
10 + 5 = 15

P ≠ NP予想

- クラスPとクラスNPの包含関係はまだわかっていない。
- ミレニアム懸賞問題として、クレイ数学研究所から100万ドルの賞金がかかっている。
- 数多の計算理論の専門家が挑んだがまだ解決していない。
- たまーに「解決した！」と証明の論文が公開されるが、証明の欠点が指摘されて、振り出しに戻っている。

P ≠ NP予想

- クラスPとクラスNPの包含関係はまだわかっていない。
- クラスPはクラスNPの部分集合なのか

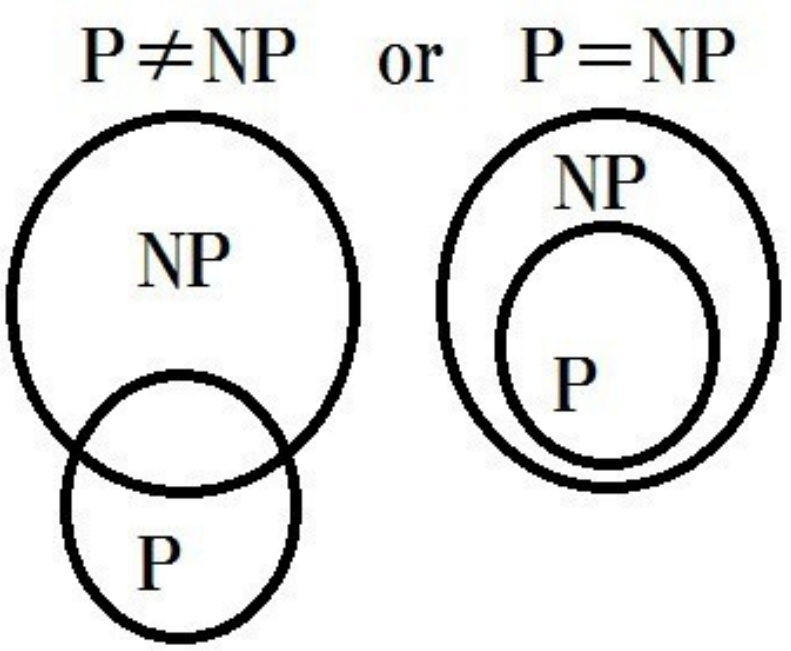
$$P \subseteq NP$$

それともクラスPとクラスNPは等しいのか

$$P = NP$$

- もし、包含関係が等しいとすると、複雑な問題を解く多項式時間アルゴリズムが存在することになる（だけど、実際にはこれまで見つかっていない）
- 大方の予想は「$P \neq NP$」となっている。

P ≠ NP予想



笑わない数学：P≠NP問題を見る – bqsfgameの日記
https://bqsfgame.hatenablog.com/entry/2022/09/10/000000

問題の帰着、問題間の関係

- 問題の帰着　ある問題を別の問題に変換することで解くこと。帰着して解ければ解けたことになる。

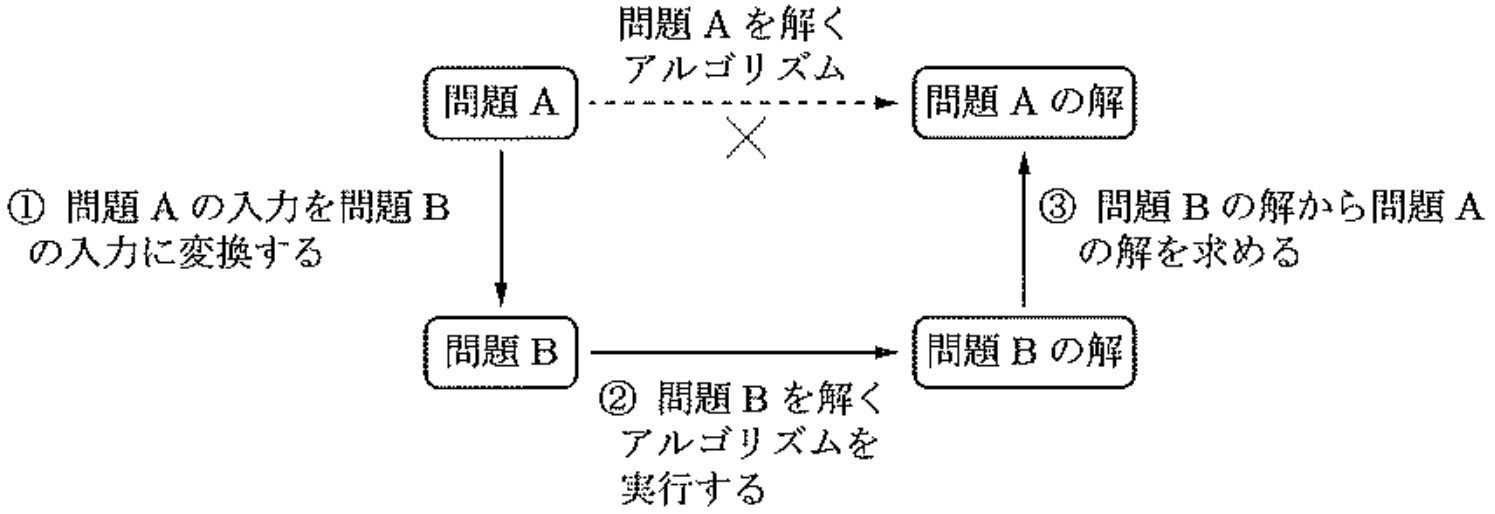


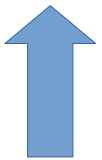
図 13.3　問題の帰着

問題の帰着の例

- ランダムに並んでいるデータから中央値($\lceil \frac{n}{2} \rceil$)を見つける。

入力：$\{17, 39, 1, 9, 5, 24, 2, 11, 23, 6\}$

出力：$\{1, 2, 5, 6, 9, 11, 17, 23, 24, 39\}$



1.クイックソートやマージソートでデータを並べ替える。
2.ソートされたデータに対して中央値を求める。

問題の帰着の手順

① 問題 A の入力を問題 B の入力に変換する.

② 変換した入力に対して，問題 B のアルゴリズムを実行する.

③ 問題 B の解から問題 A の解を求める.

問題の帰着の例

- 分割問題、部分和問題で問題の帰着を考える。

分割問題

問題 13.1

分割問題 $A = \{a_1, a_2, \ldots, a_n\}$ という n 個の正の実数の集合 A が与えられたとする．このとき，

集合 A_1 に含まれる実数の和 = 集合 A_2 に含まれる実数の和

となるように，集合 A を 2 つの集合 A_1, A_2 に分割することができるかどうかを示せ．

正数の集合：1, 2, 4, 9, 14
1 + 14 = 15
2 + 4 + 9 = 15

正数の集合：1, 4, 6, 9, 14
分割が存在しない

部分和問題

【問題 9.1】 部分和問題

$\{x_1, x_2, \ldots, x_n\}$ という n 個の正の実数の集合と，s という正の実数が与えられたとする．このとき，$\{x_1, x_2, \ldots, x_n\}$ の中からその和がちょうど s になる実数の選び方を求めよ．

正数の集合:1, 2, 3, 4, 5, 6, 7, 8, 9, 10
和:15
10 + 5 = 15

問題の帰着の例

- 分割問題、部分和問題で問題の帰着を考える。
- 考え方：部分和問題の総和を、分割問題の集合の和と結びつけて考える。

- 各 i $(1 \leq i \leq n)$ について，$x_i = a_i$ とする．
- $s = \frac{1}{2} \sum_{i=1}^{n} a_i$ とする．

問題の帰着の例

- 各 i $(1 \leq i \leq n)$ について，$x_i = a_i$ とする．
- $s = \frac{1}{2} \sum_{i=1}^{n} a_i$ とする．

正数の集合：1, 2, 4, 9, 14

S = (1 + 4 + 2 + 9 + 14) / 2 = 15

和を15にするように分割する問題に帰着できる。

問題の帰着の例

部分和問題の出力を見ることで、分割問題の結果がわかる。

- （部分和問題の出力）“和が s に等しい選び方が存在しない”
  $\Rightarrow$（分割問題の出力）“等しくなるような分割方法が存在しない”
- （部分和問題の出力）“和が s に等しくなる選び方が存在し，その組合せは集合 S である”
  $\Rightarrow$（分割問題の出力）“集合 S と集合 $A-S$ に分割すれば，和が等しくなる”

問題の帰着の計算量

① 分割問題の入力を部分和問題に変換するのに必要な時間計算量

② 部分和問題を解くのに必要な時間計算量

③ 部分和問題の出力から，分割問題の出力を決定するのに必要な時間計算量

分割問題の時間計算量 = 部分和問題の時間計算量 + $O(n)$

$$O(n2^n) + O(n) = O(n2^n)$$

部分和問題の時間計算量　$O(n2^n)$

問題 A の時間計算量 = 問題 B の時間計算量+入出力の変換に必要な時間計算量

問題の帰着の計算量

① 分割問題の入力を部分和問題に変換するのに必要な時間計算量

② 部分和問題を解くのに必要な時間計算量

③ 部分和問題の出力から，分割問題の出力を決定するのに必要な時間計算量

問題 A の時間計算量 = 問題 B の時間計算量 + 入出力の変換に必要な時間計算量

「問題Aから問題Bへ多項式時間で帰着可能」

クラスNP

- クラスNPに関する問題の難しさの定義
- 「クラスNPに含まれる問題より難しい問題の集合」と「クラスNPに含まれる問題の中でもっとも難しい問題の集合」

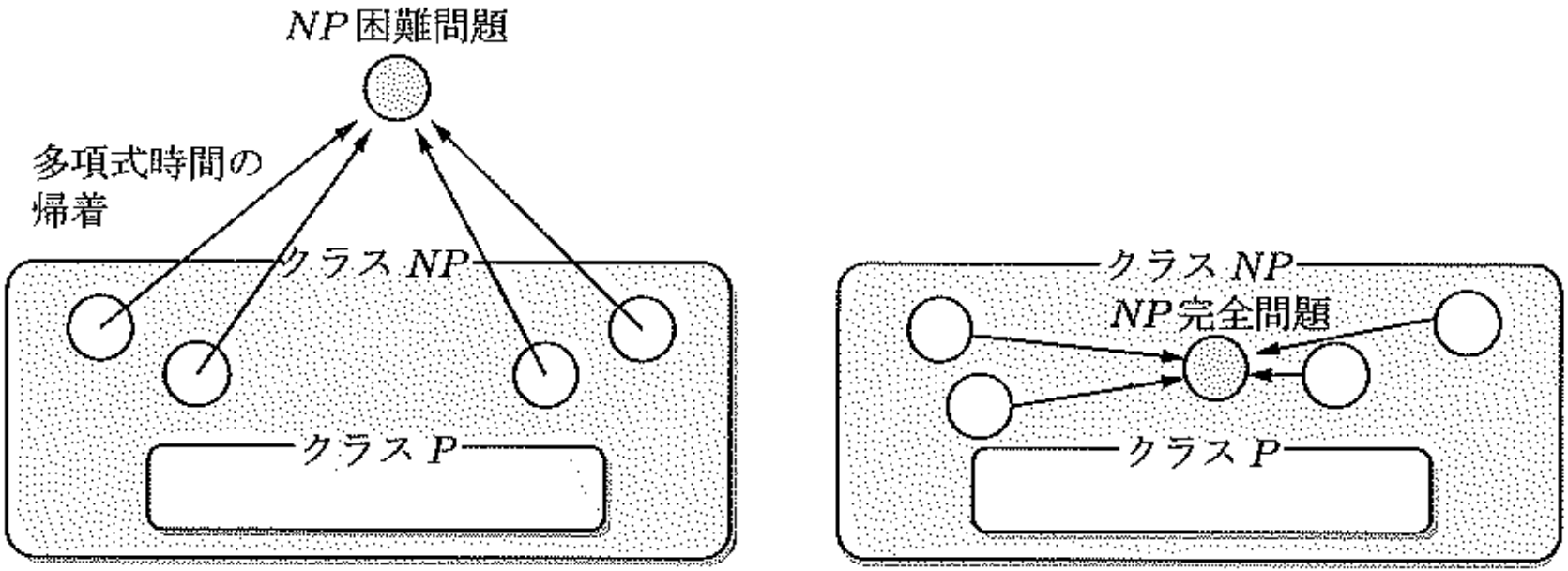
◆定義 13.1　*NP* 困難

クラス NP に含まれるすべての問題が問題 A に多項式時間で帰着可能であるとき，問題 A は *NP* 困難である．

◆定義 13.2　*NP* 完全

問題 A が NP 困難かつクラス NP に含まれるとき，問題 A は *NP* 完全である．

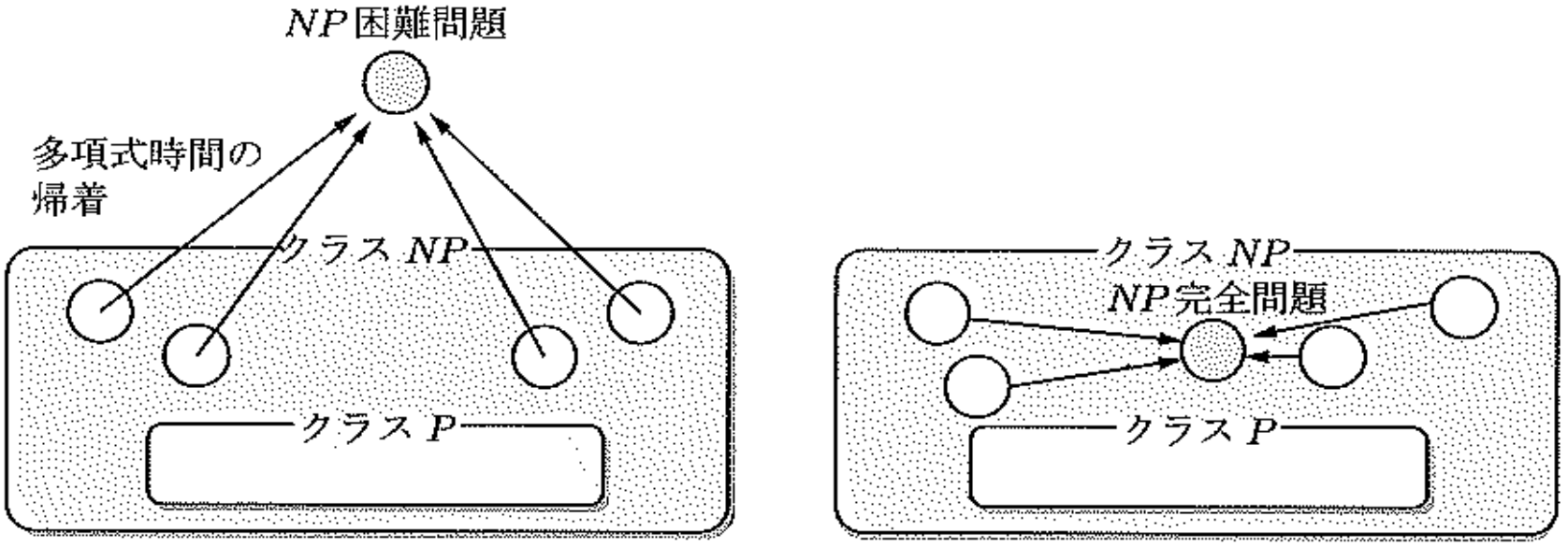
クラスNP



（a） NP困難問題　　（b） NP完全問題

図 13.4　NP 困難問題と NP 完全問題の概念図

クラスNPと計算複雑性



（a） NP困難問題　　（b） NP完全問題

クラス NP に含まれる問題の複雑さ ＝ NP 完全問題の複雑さ ＋ 多項式時間

クラスNPに関する問題の例

- 部分和問題　NP完全
- 充足可能性問題(3年生)　NP完全
- ハミルトン閉路問題(3年生)　NP完全

- ナップザック問題(0-1ナップザック問題)　NP困難
- 巡回セールスマン問題　NP困難

充足可能性問題（NP完全）

NP完全である最初の問題として証明された。（Devis・Putnam法で簡易的な問題は解ける）

充足可能性問題　n 個の 0 または 1 の変数 $x_1, x_2, \ldots, x_n$ による和積形の論理式が入力として与えられた場合に，その論理式を 1 にするような変数割当が存在するかどうかを答えよ．

論理式　$(\overline{x_1} \vee x_2 \vee \overline{x_3})(x_1 \vee x_3)(\overline{x_1} \vee \overline{x_2})$

x1=1, x2=0, x3=0の場合、出力は1となる

$(x_1 \vee \overline{x_2})(x_1 \vee x_2 \vee x_3)(\overline{x_3})(\overline{x_1} \vee x_3)$

3変数でこの式の長さは真理値表で手計算できる

ハミルトン閉路問題（NP完全）

ハミルトン経路問題　n 個の頂点のグラフとそのグラフの2つの頂点が出発頂点と到着頂点として与えられた場合に，出発頂点から始まって各頂点を1回ずつ通り，到着頂点に達するような経路をみつけよ．

ハミルトン閉路問題がNP完全のはず（教科書と比較）

コストは無視する
出発頂点　v1
到着頂点　v7

経路

$(v_1, v_2, v_3, v_4, v_6, v_5, v_7)$

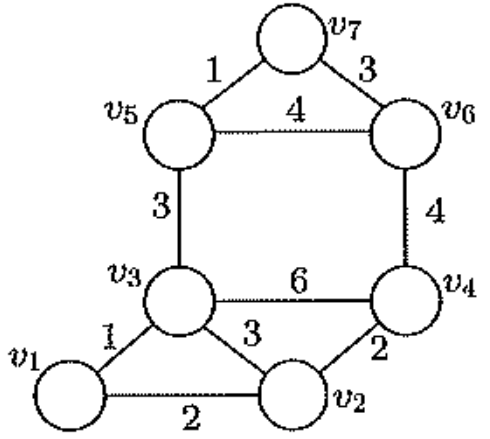


図 13.5　ハミルトン経路問題と巡回セールスマン問題の入力

巡回セールスマン問題(NP困難)

巡回セールスマン問題　ハミルトン経路問題と同様の n 個の頂点のグラフが与えられ，各辺の長さが決められているものとする．このとき，各頂点を１回ずつ通る最短のハミルトン経路をみつけよ．

原題　出発頂点v1からv1へ戻るコスト最小の経路はどれか
(出発頂点と到着頂点が同じ頂点)

最小のコストは求めることが難しい。
組合せ爆発をすぐに起こす代表的な問題になっている。

応用上の問題(非常に重要)
- バス、タクシーの巡回(営業所に戻る必要がある)
- 電子回路基板の半田付けの順番

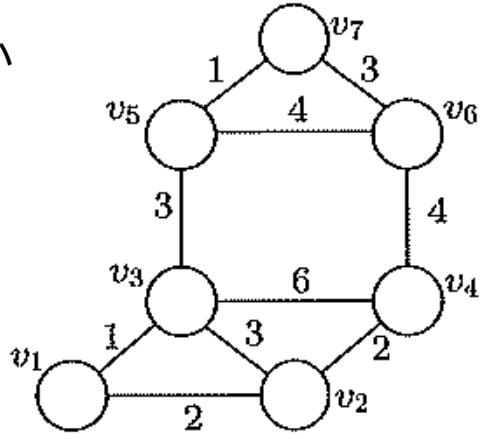
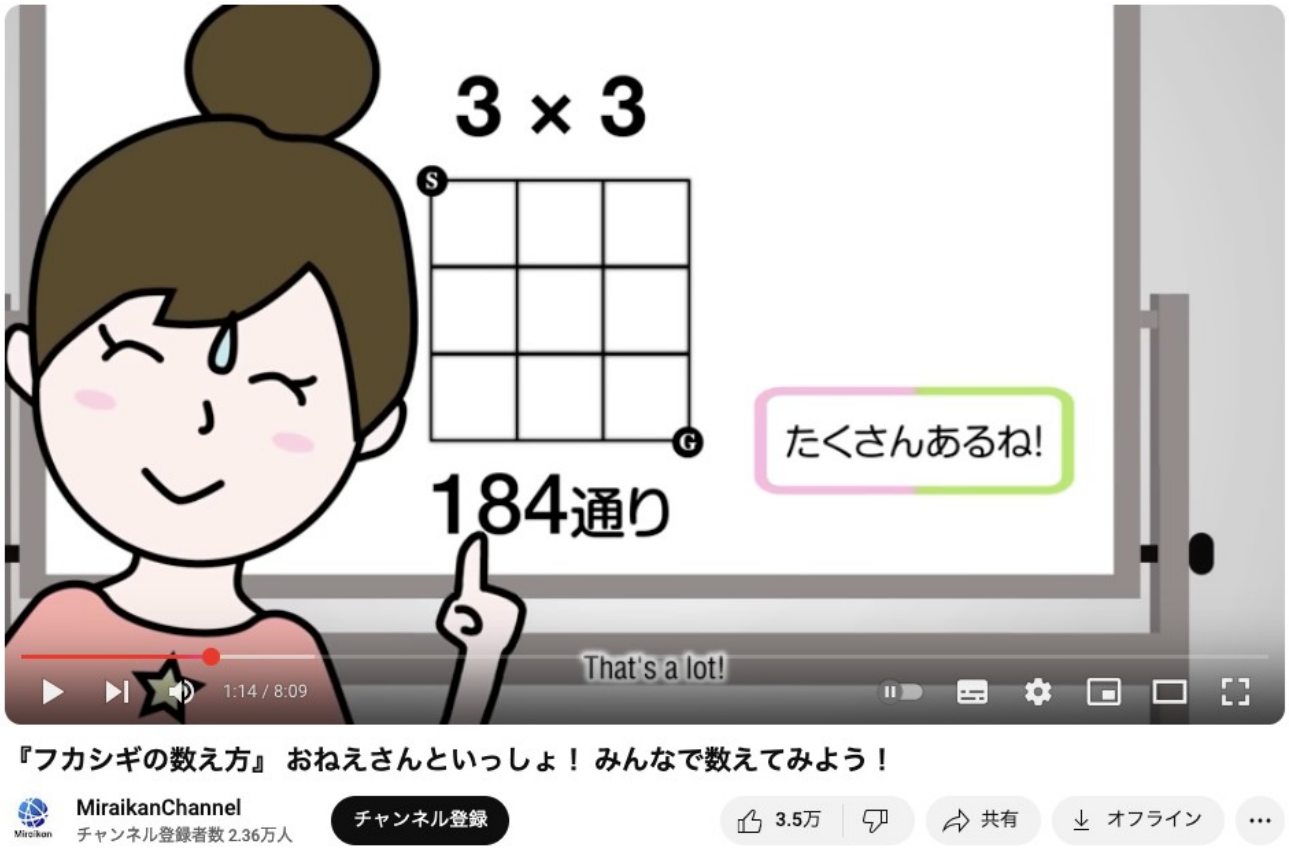


図 13.5　ハミルトン経路問題と巡回セールスマン問題の入力

手計算をするおねえさんの動画



日本未来科学館が公開している計算複雑性を説明している真面目な動画(かつて、徳光がプログラミング担当のときには紹介していました。)

もし、この問題がP=NPだとしたら、こんな苦労はなかったのかもしれない・・・。

『フカシギの数え方』おねえさんといっしょ！ みんなで数えてみよう！ - YouTube
https://www.youtube.com/watch?v=Q4gTV4r0zRs

P ≠ NP予想（再訪）

- クラスPとクラスNPの包含関係はまだわかっていない。
- クラスPはクラスNPの部分集合なのか

$$P \subseteq NP$$

それともクラスPとクラスNPは等しいのか

$$P = NP$$

もし、P=NPだとすると、世の中が大変なことになる。暗号を解くのが多項式時間でできてしまう（つまり現実的な時間）。解けない方が嬉しいのか、解けることでより便利になるのか。さてどちらをとるか。

- もし、包含関係が等しいとすると、複雑な問題を解く多項式時間アルゴリズムが存在することになる（だけど、実際にはこれまで見つかっていない）
- 大方の予想は「$P \neq NP$」となっている。

暗号と計算複雑性

- 現代の暗号の主流はRSA暗号
- 桁数が長い素因数分解の難しさを利用している
- 鍵の長さが2048ビットだと今のスーパーコンピュータでも1億年以上（事実上不可能、人類が存在というか地球がそもそもあるのか？）