2 計算効率の評価

ネットワーク流に限らず数理計画法では，アルゴリズムを開発するうえで最も注意の払われるのが計算効率の善し悪しである．通常，この計算効率の評価には次の 3 つの方法が用いられている:

経験的解析 (empirical analysis). 問題例の分布を特定してサンプリングを行い，それらに対するアルゴリズムの実際の振舞いを計測する．

最悪計算量解析 (worst-case analysis). 任意の問題例に対し，アルゴリズムが要する計算量の上界を解析する．

平均計算量解析 (average-case analysis). 問題例の分布に特定の仮定を設け，アルゴリズムが要する計算量の平均を解析する．

このうち平均計算量は一般に解析が難しいため，アルゴリズムの性能評価に広く用いられているのは経験的解析と最悪計算量解析である．この節では主に後者について解説しよう．

2.1 最悪計算量解析

最悪計算量解析では，四則演算や代入，比較など基本演算 1 回が単位時間で実行されると仮定し，問題サイズを表すいくつかのパラメータ: 頂点数 n，枝数 m，枝の費用と容量の上限 $C = \max\{c_{ij} \mid (i,j) \in E\}$, $U = \max\{u_{ij} \mid (i,j) \in\}$ を用いてアルゴリズムの実行時間の上界を算定する．例えば，最短路問題のアルゴリズムとして後述するダイクストラ法は，ある定数 p, q に対して高々 $pn(n-1) + qm$ 回の基本演算を必要とするが，これを記号 $O(\cdot)$ を用いて「ダイクストラ法の実行時間は $O(n^2)$ である」という．この $O(\cdot)$ 表記では，定数を無視して「n や m などのパラメータが十分に大きな場合に演算回数を支配する項のみ」を表す．したがって，実際の演算回数が $2^n + 500!n^2$ ならば実行時間は $O(2^n)$ となる．

多項式時間アルゴリズム. アルゴリズムの実行時間が問題の入力ビット長の多項式関数でおさえられるとき，そのアルゴリズムを**多項式時間アルゴリズム** (polynomial-time algorithm) という．最小費用流問題 (1.2) では，入力ビット長は n や m, $logC$, $\log U$ の低次多項式となるので，これらの多項式で (1.2) を解くアルゴリズムの実行時間がおさえられるなら，それは多項式時間アルゴリズムである．また，4 つのパラメータのうち n と m だけの多項式で実行時間のおさえられるアルゴリズムを特に**強多項式時間アルゴリズム** (strongly polynomial-time algorithm) という．

指数時間アルゴリズム. 実行時間が入力ビット長の多項式ではおさえられないアルゴリズムは**指数時間アルゴリズム** (exponential-time algorithm) とよばれる．例えば，$O(nC)$, $O(2^n)$, $O(n!)$

や $O(n^{\log n})$ などは指数時間である．しかし，$O(nC) = O(n2^{\log C})$ のように実行時間が，入力ビット長に関しては指数であるが，m, n, C, U に関しては多項式となるアルゴリズムも存在する．これらを**擬多項式時間アルゴリズム** (pseudopolynomial-time algorithm) という．

さて，最悪計算量解析による 2 種類の実行時間のうち，アルゴリズムとしては一体どれが好ましいか？ それは 100MPS（1 秒間に 100 万回の基本演算が可能）の計算機を想定した次の結果から一目瞭然であろう．

入力	実行時間					
n	$n\log n$	n^2	n^5	2^n	$n^2 2^n$	$n!$
10	3.32×10^{-7} 秒	1×10^{-6} 秒	0.0015 秒	2.1×10^{-5} 秒	0.001 秒	0.036 秒
20	8.64×10^{-7} 秒	4×10^{-6} 秒	0.032 秒	1.05×10^{-2} 秒	4.19 秒	771 年
30	1.47×10^{-6} 秒	9×10^{-6} 秒	0.243 秒	10 秒	2.68 時間	5.61×10^{6} 百五十億年
40	2.13×10^{-6} 秒	1.6×10^{-5} 秒	1.02 秒	3.05 時間	204 日	1.72×10^{22} 百五十億年
50	6.64×10^{-6} 秒	2.5×10^{-5} 秒	3.13 秒	103 日	8.93 世紀	6.42×10^{38} 百五十億年
100	6.62×10^{-6} 秒	1×10^{-4} 秒	1.67 分	26798 百五十億年	2.68×10^{8} 百五十億年	1.77×10^{132} 百五十億年

こので,「百五十億年」とあるのはビッグバンが起きてから現在までの時間．したがって，大規模な問題に対してはハードウェアの性能をいくらあげても，指数時間アルゴリズムを用いいる限り，焼け石に水，いや太陽に滴でしかないことがわかる．本稿では直列計算機を想定しているが，並列計算機を使っても現実には同様の問題の生じることが知られている．

2.2　NP 完全

さいわいにして最小費用流問題 (1.2) には強多項式時間アルゴリズムが存在している．しかし，数理計画問題の中で多項式時間に解決可能な問題はむしろ希で，多くはその可能性すら絶望視されている．そうした問題の代表が**巡回セールスマン問題** (traveling salesman problem) である:

最適化巡回セールスマン問題
入力: 有向グラフ $G = (V, E)$ と各枝 $(i, j) \in E$ の長さ $c_{ij} \in \mathbf{Z}$
出力: グラフ G の巡回路 (有向ハミルトン閉路)H で距離 $\sum_{(i,j)\in H} c_{ij}$ が最小のもの．

この問題には多項式時間アルゴリズムなど存在しないと大方の研究者に共通の見解であるが，その理論的な裏付けとなっているのが NP 完全の概念である．

大まかにいって次の特徴をもつ**認識問題** (recognition problem; ‘yes’ か ‘no’ で答えられる問題) のクラスを，Nondeterministc Polynomial-time solvable の頭文字を取って NP とよぶ:

(a) 答の証拠を教えてもらえば，それが問題の条件を満たすかどうかのチェックを多項式時間でできる．

例えば，認識問題:

有向ハミルトン閉路
入力: 有向グラフ $G = (V, E)$
質問: グラフ G に有向ハミルトン閉路は存在するか？

は，答の証拠として頂点 $i, \ldots, n \in V$ の順列 $(i_1, \ldots, i_n)$ が与えられれば，

$$(i_k, i_{k+1}) \in E, \;\; k = 1, \ldots, n-1; \;\; (i_n, i_1) \in E$$

が満たされるかどうかを確かめればよく，明らかに多項式時間で処理できる．したがって，有向ハミルトン閉路問題は NP に属する．

NP 問題を「解く」には，しかし

(b) 答の候補が入力ビット長の指数関数個もあり，それらの証拠を 1 つ 1 つチェックするという単純な方法では指数時間を必要とする

可能性もある．この可能性を示唆するときに計算の複雑さの理論で用いられるロジックは，

> 「この問題 A が多項式時間で解けるくらいなら，あの難問 B だって多項式時間で解ける」（だから，A はたぶん多項式時間では解けない）

といういささか消極的なものである．

多項式時間還元 (polynomial-time reduction)**:** 問題 A を 1 回の基本演算と同じ時間で解くアルゴリズム α があれば，α をサブルーチンとして用いることで問題 B を多項式時間で解くことができるとき，問題 B は A に多項式時間還元可能であるという．

この定義を使えば，(b) の可能性のある NP 問題を次のように分類することができる:

NP 完全 (NP-complete)**:** 問題 A が NP で，すべての NP 問題が A へ多項式時間還元可能なとき，問題 A は NP 完全であるという．

言い換えれば，NP 完全な問題はどんな NP 問題と比べても**それ以上に**難しい．ところが，NP 問題は 1 つではなく，星の数ほど存在する．したがって，考えている問題が NP 完全か否かを判定するのは不可能にも思えるが，例えば先の有向ハミルトン閉路問題から，次の定理によって芋づる式に明らかにすることができる:

定理 2.1. 問題 B を任意の NP 完全問題とするとき，B が問題 A へ多項式時間還元可能ならば，A は NP 完全である．

試しに有向ハミルトン問題が次の問題へ多項式時間で還元できることを示し，これが NP 完全であることを証明してみよう:

認識巡回セールスマン

入力: 有向グラフ $G=(V,E)$，整数 b，および各枝 $(i,j)\in E$ の長さ $c_{ij}\in\mathbf{Z}$
質問: グラフ $G=(V,E)$ の巡回路 H で $\sum_{(i,j)\in H} c_{ij}\le b$ を満たすものは存在するか？

以下の手順で有向ハミルトン問題を解くアルゴリズム β を作成する:
まず，入力のグラフ $G=(V,E)$ をもとに

$$c_{ij}=\begin{cases}1, & (i,j)\in E \text{ のとき}\\ 2, & (i,j)\notin E \text{ のとき},\end{cases}$$

と定義し，各枝 $(i,j)\in E'$ が長さ $c_{ij}\in Z$ をもつ完全グラフ $G'=(V,E')$ をつくる．この c_{ij} の定義から直ちに

$$G' \text{ の巡回路 } H \text{ が } \sum_{(i,j)\in H} c_{ij}\le n \text{ を満たす} \Longleftrightarrow H\subseteq G$$

であることがわかる．したがって，G', $b=n$ および c_{ij} を入力とする認識巡回セールスマン問題が ‘yes’ の答をもてば，そのときに限ってもとのグラフ G には有向ハミルトン閉路が存在する．そこで，認識巡回セールスマン問題を解くアルゴリズム α をサブルーチンとして呼び出せばアルゴリズム β の完成である．

グラフ G' と c_{ij} の定義に G の頂点対の数 $n(n-1)/2$ に比例する手間がかかるが，α が単位時間しか必要としなければ，アルゴリズム β は多項式時間であり，したがって「有向ハミルトン閉路問題は認識巡回セールスマン問題へ多項式時間還元可能」といえる．さて，残された問題は認識巡回セールスマン問題が NP かどうかだが，その答の証拠として巡回路 H が与えられれば，$O(n)$ で $\sum_{(i,j)\in H} c_{ij}\le b$ をチェックできることは明らかであろう．

ところで，上に作成した有向ハミルトン問題を解くアルゴリズム β のサブルーチンとして，認識巡回セールスマン問題のアルゴリズム α を呼ぶ代わりに最適化巡回セールスマン問題のアルゴリズム α' を呼べばどうなるであろう？ アルゴリズム α' が返す巡回路 H は最小の距離をもつので，$\sum_{(i,j)\in H} c_{ij}\le b$ ならば ‘yes’，そうでなければ ‘no’ をアルゴリズム β の出力とすれば良さそうである．アルゴリズム α' も単位時間しか必要としないことを仮定すれば，この修正によって β が指数時間になることもないので,「有向ハミルトン問題は最適化巡回セールスマン問題へ多項式時間還元可能」である．ならば,「最適化巡回セールスマン問題も NP 完全か？」というと，答は ‘no’ である．

NP 困難 (NP-hard)**:**　すべての NP 問題が A へ多項式時間還元可能なとき，問題 A は NP 困難であるという．

NP 完全の定義と比べると,「問題 A が NP で」という部分が抜け落ちている．最適化巡回セールスマン問題は，そもそも ‘yes’ や ‘no’ で答えられる認識問題ではなく，したがって NP には属さ

ない．しかし，任意の NP 完全問題から多項式時間で還元できるので，どんな NP 問題と比べてもそれ以上に難しい．そこで，最適化巡回セールスマン問題のような問題は NP 困難と総称されているわけであるが，その「困難」さはあくまで最悪計算量解析の結果であって平均的に速いアルゴリズムの存在まで否定するものではない．

最後になってしまったが，NP 完全に対し，与えられた証拠のチェックだけでなく答までも多項式時間で得られる認識問題の集合を，Polynoimal-time solvable の頭文字を取って**クラス** P という．巡回セールスマン問題に多項式時間アルゴリズムが絶望視されている理由は，一般に P $\neq$ NP が信じられているからにすぎない．たった 1 問の NP 完全問題でも多項式時間で解けることが証明できれば，定理 2.1 から直ちにこの不等号は否定される（演習問題: それはなぜか？）のだが，それに成功した研究者はまだ一人もいない．

参考文献

NP 完全の話は，本稿の扱うネットワーク流問題には直接関係しないが，これを読んで次のような疑問を感じ，少しでも興味を抱いてくれれば結構である: 認識巡回セールスマン問題が有向ハミルトン閉路問題から多項式時間で還元できるので NP 完全であることは認めるが，有向ハミルトン問題の NP 完全性はどうやって証明されたのか？ — それは，認識巡回セールスマン問題の NP 完全性が認められたので，そこから有向ハミルトン閉路問題へ多項式時間で還元させる — ことももちろんできるが,「卵が先か，ニワトリが先か」の議論になってしまう．実は，有向ハミルトン閉路問題は 72 年に R.M. Karp によって NP 完全問題:

頂点被覆

入力: グラフ $G = (V, E)$ と整数 $k \leq n$
質問: グラフ G に大きさが k 以下の頂点被覆 — $|V'| \leq k$ で，各枝 $(i, j) \in E$ に対して頂点 i と j の少なくとも一方を含む頂点集合 $V' \subseteq V$ — は存在するか？

から多項式時間で還元されている．それでは，頂点被覆はといえば，やはり Karp が

3 充足性

入力: 論理変数 $u_1, \ldots, u_n$ と高々 3 個の論理変数の和からなる項 $C_1, \ldots, C_m$
質問: $C_1 \wedge \cdots \wedge C_m = 1$ を満たす $u_1, \ldots, u_n$ は存在するか？

から多項式時間還元されている．それでは，3 充足性はといえば，71 年に S.A. Cook が

充足性

入力: 論理変数 $u_1, \ldots, u_n$ と項 $C_1, \ldots, C_m$
質問: $C_1 \wedge \cdots \wedge C_m = 1$ を満たす $u_1, \ldots, u_n$ は存在するか？

から多項式時間還元させている．それでは，充足性はといえば， — NP 完全問題の記念すべき第 1 号で，Cook は 71 年に定理 2.1 の助けを借りずに（何しろ充足性問題に還元すべき NP 完全問題はそれまで発見されていない）証明している: