

13. 近似アルゴリズム

13.1 近似アルゴリズムの種類

NP困難な問題に対しては多項式時間で最適解を求めることは困難であるので、最適解に近い近似解を求めるアルゴリズムが用いられることがある。

このように、必ずしも厳密解を求めないアルゴリズムは、大きく分けて2つの範疇に分けられる。

ヒューリスティックと近似アルゴリズム

ヒューリスティクス

(発見的解法、経験的解法)

精度保証無し、
実験的評価が主

- 遺伝的アルゴリズム (GA)
- アニーリング (焼きなまし法)
- タブサーチ
- ローカルサーチ
- ニューラルネット等

近似アルゴリズム

精度保証付き
理論的解析が主

- 定数近似アルゴリズム (APX)
- 多項式近似スキーム (PTAS)
- 完全多項式近似スキーム (FPTAS)

ここでは、近似アルゴリズムについてみていく。

α 近似アルゴリズム

定義: α 近似アルゴリズム

最小化問題に対して、いつも最適解の α 倍以下の解を入力サイズの多項式時間で求めるようなアルゴリズムを α 近似アルゴリズムという。ここで、 $\alpha \geq 1$ であり α を近似率という。すなわち、最適解を $f_{OPT}(x)$ と表し、アルゴリズムの解の評価値を $f_{AL}(x)$ と表すと、常に次の式を満足する。

$$\frac{f_{AL}(x)}{f_{OPT}(x)} \leq \alpha$$

$$\therefore f_{AL}(x) \leq \alpha f_{OPT}(x)$$

β近似アルゴリズム

定義: β近似アルゴリズム

最大化問題に対して、いつも最適解のβ倍以上の解を入力サイズの多項式時間で求めるようなアルゴリズムをβ近似アルゴリズムという。ここで、 $\beta \leq 1$ でありβを近似率という。また、βを近似保証ということもある。すなわち、最適解を $f_{OPT}(\mathbf{x})$ と表し、アルゴリズムの解の評価値を $f_{AL}(\mathbf{x})$ と表すと、常に次の式を満足する。

$$\frac{f_{AL}(\mathbf{x})}{f_{OPT}(\mathbf{x})} \geq \beta$$

$$\therefore f_{AL}(\mathbf{x}) \geq \beta f_{OPT}(\mathbf{x})$$

APX

定義: クラスAPX

α (あるいは β) 定数であるような多項式時間アルゴリズムが存在するクラスを **APX (APproXimation)** という。

PTASとFPTAS

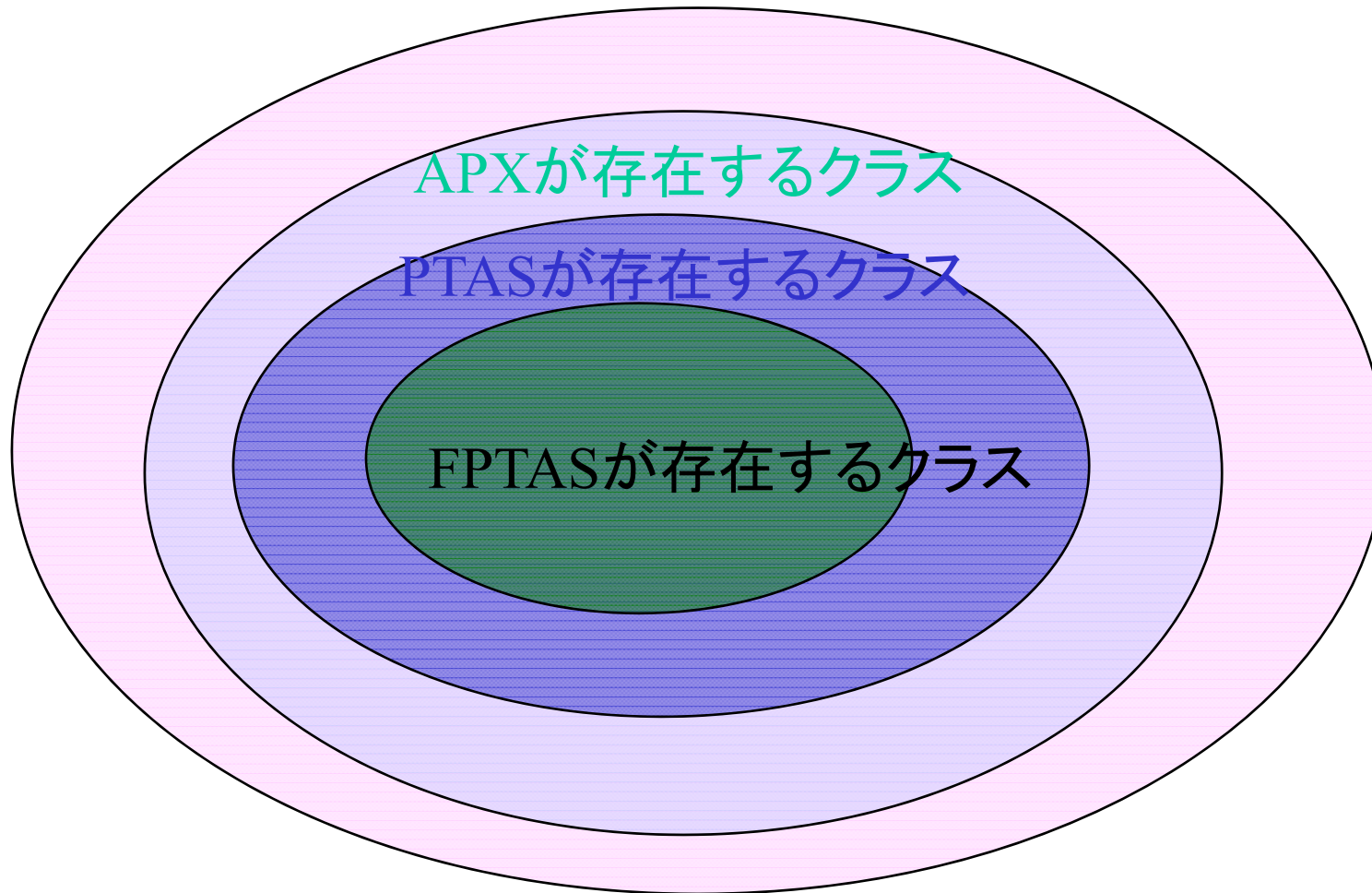
定義: PTAS、FPTAS

近似率 $\alpha(\beta)$ を限りなく1に近づけることができるとき、そのような α 近似アルゴリズムを**PTAS(Polynomial Time Approximation Scheme, 多項式時間近似スキーム)**という。すなわち、任意の正数 $\varepsilon > 0$ に対して、 $\alpha = 1 + \varepsilon$ とできる入力サイズの多項式時間アルゴリズムをPTASという。さらに、入力サイズ n と精度 $1/\varepsilon$ の多項式時間アルゴリズム**FPTAS(Fully Polynomial Time Approximation Scheme, 完全多項式時間近似スキーム)**という。

問題によっては、PTASが存在しない(知られていない)ものがある。また、定数近似すら存在しない問題もある。このようなアルゴリズムの出力は入力サイズに依存した近似値になってしまう。

近似アルゴリズムの存在

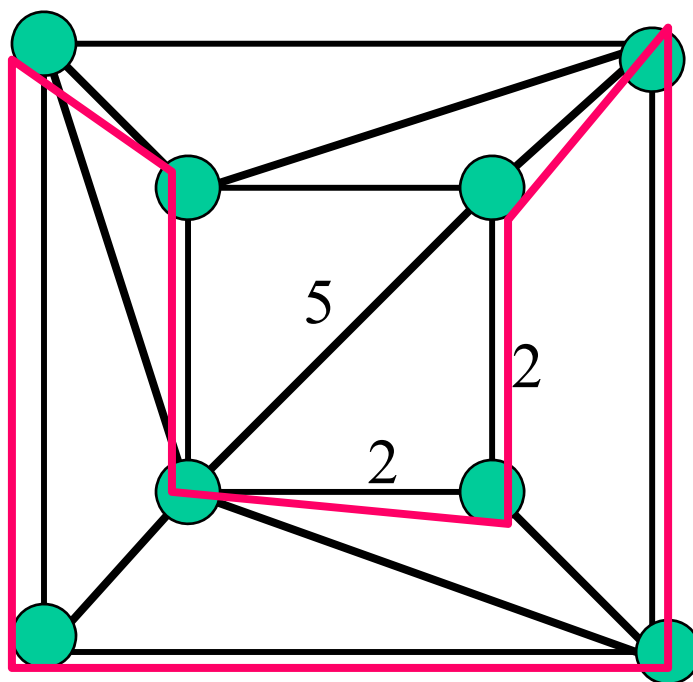
近似アルゴリズムが存在するクラス



13-2. 巡回セールスマン問題

巡回セールスマン問題には、ネットワーク型と、幾何型とがある。

ネットワーク型の巡回セールスマン問題では、入力は辺重み付きのグラフであり、出力は頂点を全て巡る閉路で重みの総和が最小のものである。



三角不等式を満足しないようなネットワーク型の巡回セールスマン問題は、近似アルゴリズムを得ることすらNP-困難である。

具体的には、定数近似アルゴリズムが存在すると、NP完全問題であるハミルトン閉路問題が多項式時間で解けてしまう。つまり、ハミルトン閉路問題をネットワーク型の定数近似巡回セールスマン問題に帰着できてしまう。このことより、ネットワーク型の巡回セールスマン問題には定数近似アルゴリズムは存在しないと考えられている。

ハミルトン閉路問題と巡回セールスマン問題

問題

名称: ハミルトン閉路問題

インスタンス: グラフ G

問い: G 中に、全ての頂点を丁度1度だけ通るような閉路は存在するか？

問題

名称: 巡回セールスマン(NTSP)

インスタンス: ネットワーク N 、正数 K

問い: ネットワーク中の全ての頂点を通るような閉路で重みの総和が K 以下となるようなもの存在するか？

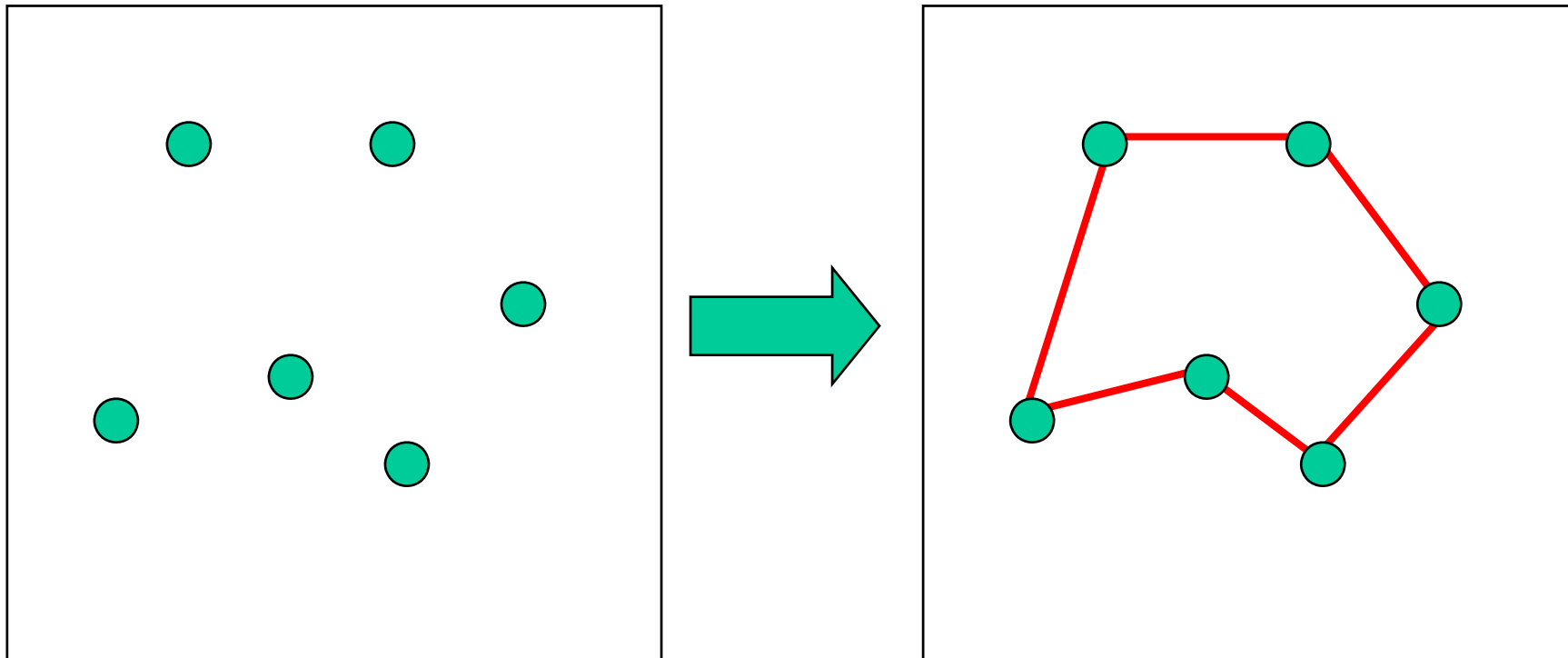
巡回セールスマン問題への帰着

ネットワーク型の巡回セールスマン問題を解く α 近似アルゴリズムが存在すると、次のようにハミルトン閉路問題を巡回セールスマン問題に帰着できる。つまり、ハミルトン閉路問題のインスタンスであるグラフ G から巡回セールスマン問題のインスタンスである辺重み付きグラフ(ネットワーク N)と整数 K を定めればめることができる。

まず、 G の辺に全て1の重みをつけてネットワーク N を構成する。次に、 K として、 $K = \frac{n}{\alpha}$ として、巡回セールスマン問題へ帰着する。このとき、 α の定数近似であるAPXが存在すると、多項式時間でハミルトン閉路の存在判定ができてしまう。

以上のことより、 $P \neq NP$ である限り、ネットワーク型の巡回セールスマン問題を解く多項式時間アルゴリズムはない。

幾何巡回セールスマン問題 (GTSP)



実は、幾何巡回セールスマン問題には、定数近似アルゴリズムが存在する。ネットワーク型と、幾何型での最大のちがいは、三角不等式が成り立つかどうかである。ここで三角不等式とは、任意の x, y, z 対して、次が成り立つことである。

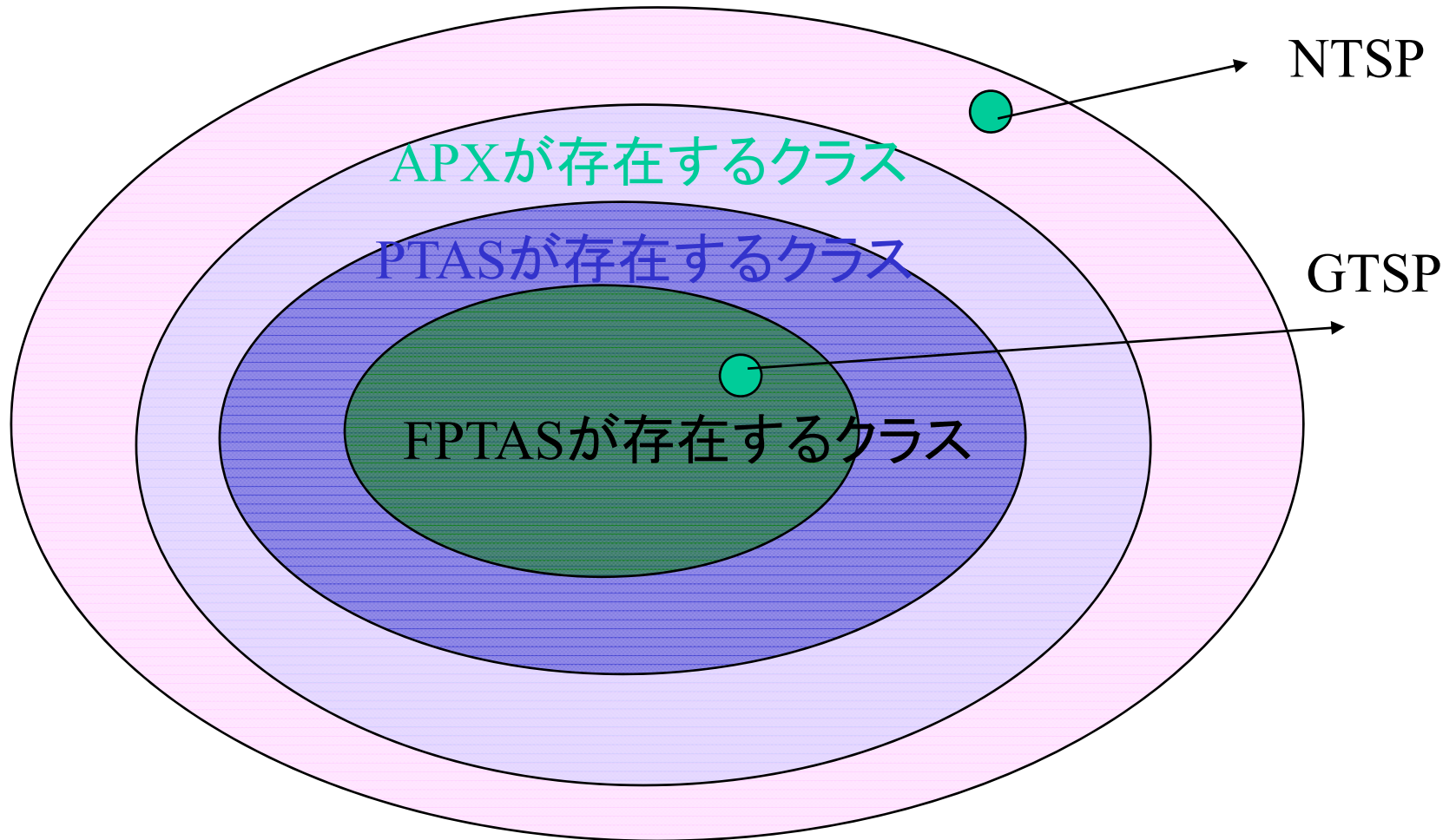
$$d(x, y) \leq d(x, z) + d(y, z)$$

2次元(ユークリッド)平面上の点集合を考えれば、2点間の距離は自動的に三角不等式を満たす。

このことは、利用できる情報(三角不等式)が多くなったということであり、ネットワーク型よりも幾何型の方が問題が簡単であることを意味する。実際、幾何型の巡回セールスマン問題は、FPTASを持つことが最近(1998年)に示された。このアルゴリズムは複雑なので、ここでは2近似アルゴリズムとそれを改善した1.5近似アルゴリズムを示す。

NTSPとGTSP

近似アルゴリズムが存在するクラス



2近似アルゴリズム

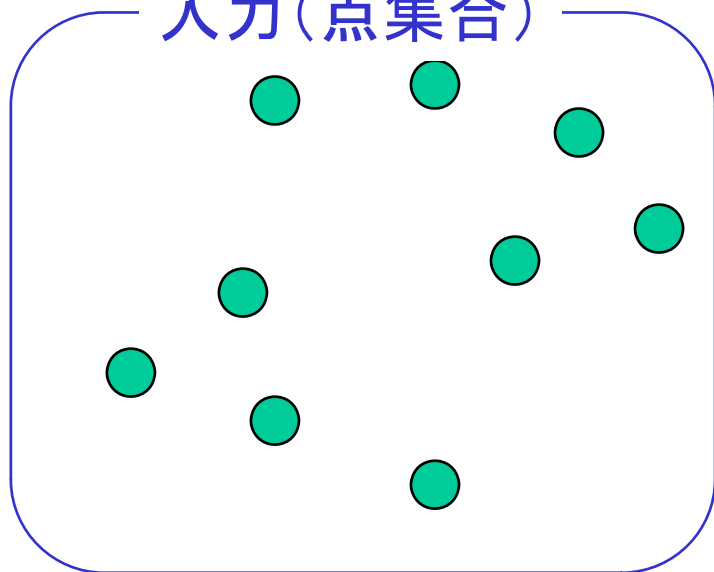
GTSPに対する2近似アルゴリズム

1. 点集合を連結する最小全域木MST T を
求める。
2. T の辺を辿りながら、全ての点を通る巡回路
 C を求める。 $(T$ の辺を全て2重化すればいい。)
3. C で一度通過した点をショートカットする順回路
 C を求める。

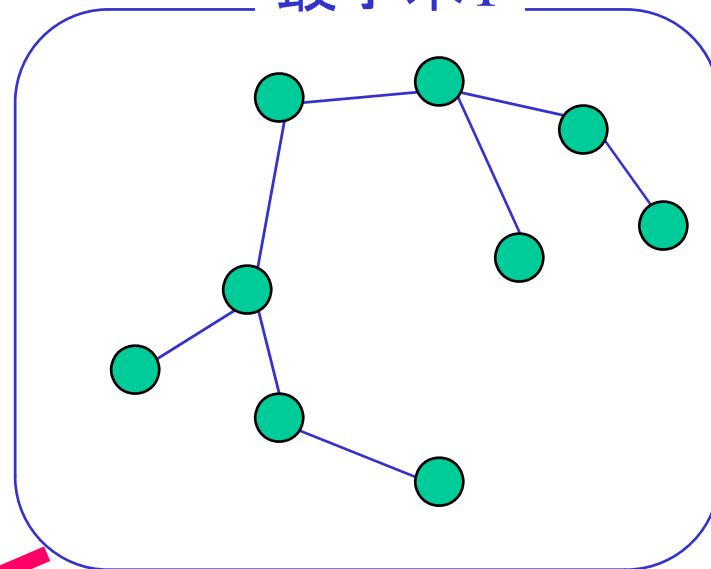
次にこのアルゴリズムの動作を示す。

2近似アルゴリズムの動作

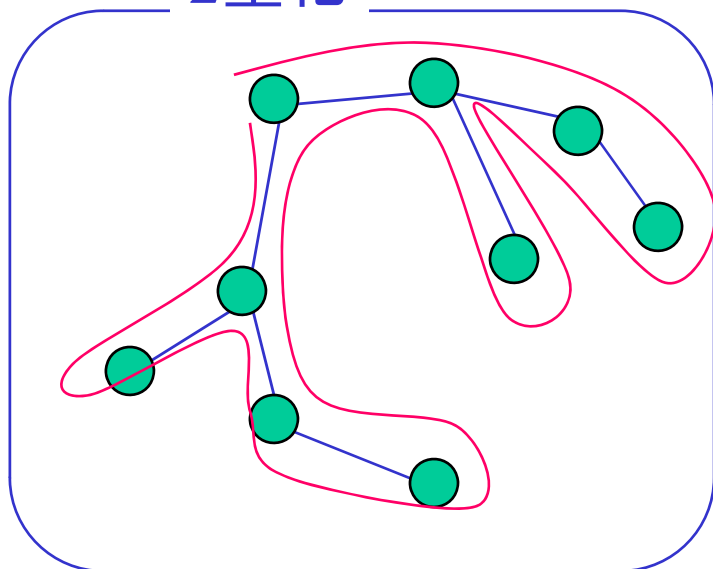
入力(点集合)



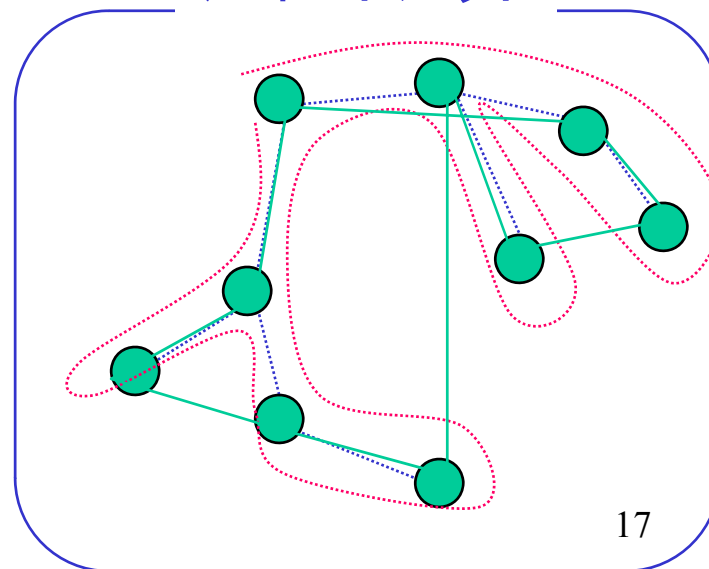
最小木T



2重化



ショートカット



近似率の解析

最適な順回路を C_{OPT} とし、アルゴリズムで得られる順回路を C_{AL} とする。また、 $|C_{OPT}|, |C_{AL}|$ でそれぞれの長さを表すものとする。

順回路 C_{OPT} から任意に一本辺を除去すると、全域木が得られる。よって、最小全域木 T の方が辺の重み総和は小さい。(そもそも、頂点を連結するもののなかで、重みが最小なものが最小全域木であった。)

よって、次が成り立つ。

$$|T| \leq |C_{OPT}|$$

また、アルゴリズムで得られる閉路では、最小木を2重化したものより小さいので、次が成り立つ。

$$|C_{AL}| \leq 2|T|$$

$$\therefore \frac{|C_{AL}|}{2} \leq |T|$$

以上より、

$$\therefore \frac{|C_{AL}|}{2} \leq |T| \leq |C_{OPT}|$$

$$\therefore |C_{AL}| \leq 2|C_{OPT}|$$

$$\therefore \frac{|C_{AL}|}{|C_{OPT}|} \leq 2$$

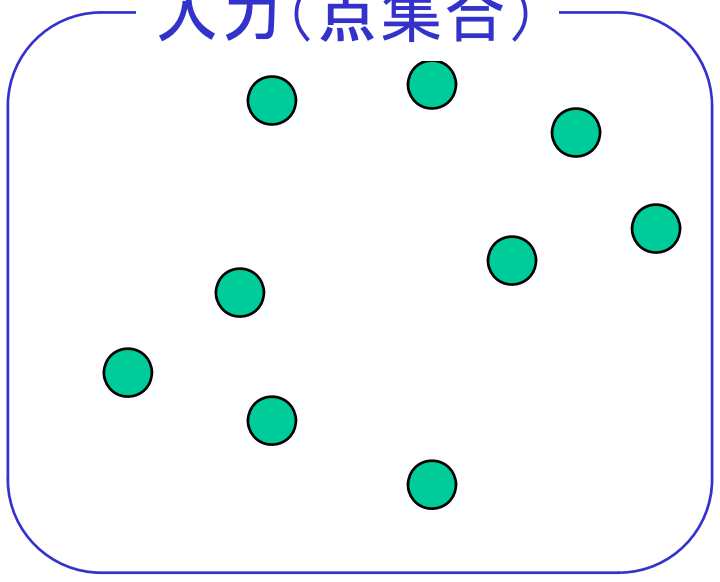
1. 5近似アルゴリズム

GTSPに対する1. 5近似アルゴリズム

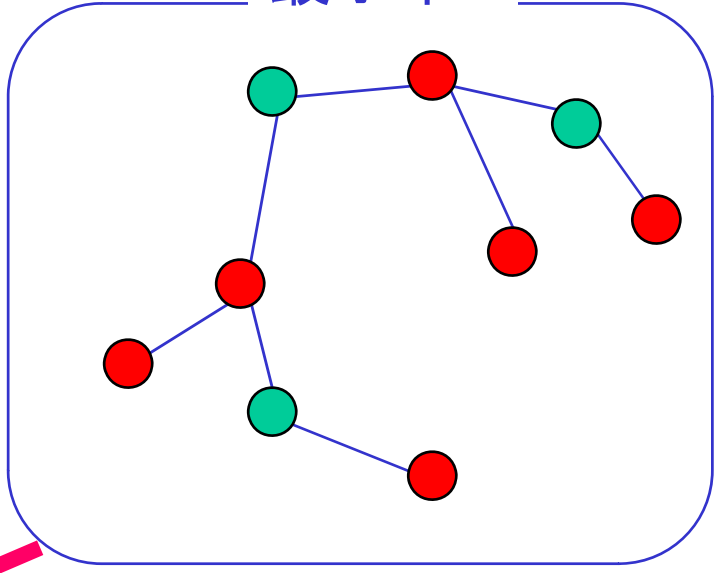
1. 点集合を連結する最小全域木MST T を
求める。
2. T において、次数が奇数の点からなる完全グラフを作り、最小重みマッチング M を求める。
3. $T+M$ はオイラーグラフであるので、一筆書きに対応する順回路 C を求める。
4. 3の順回路 C からできるだけショートカットとして、
順回路 C' を構成する。

1. 5近似アルゴリズムの動作

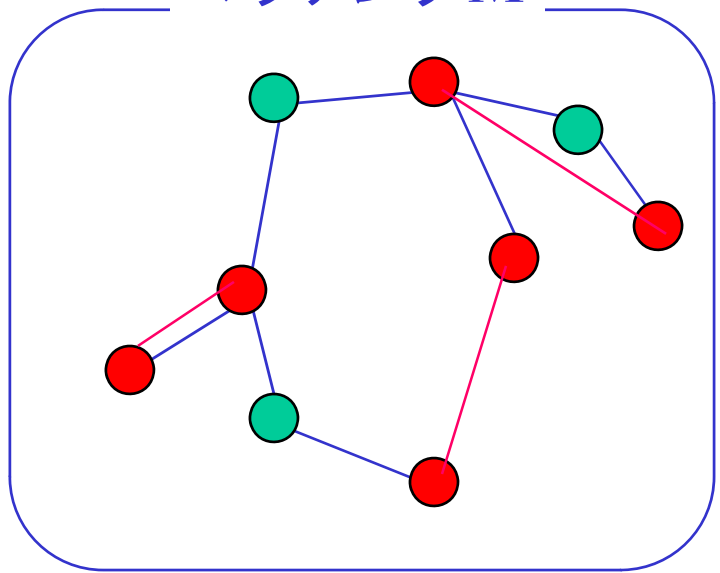
入力(点集合)



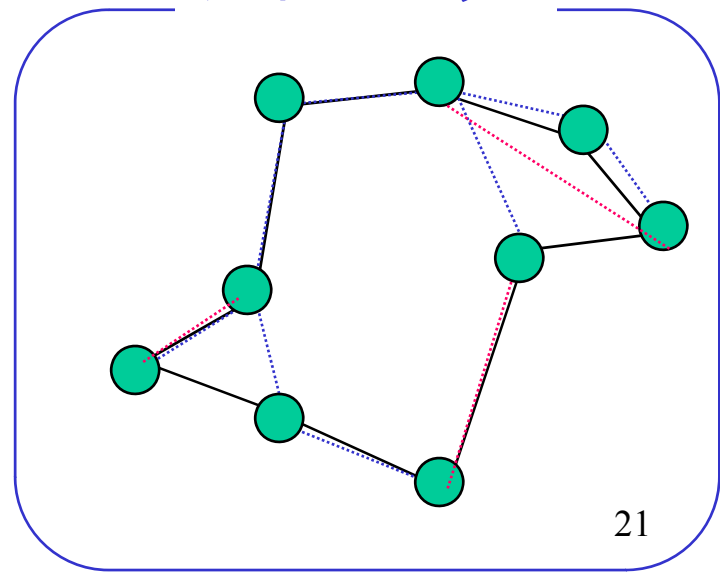
最小木T



マッチングM



ショートカット



近似率の解析

最適な順回路を C_{OPT} とし、アルゴリズムで得られる順回路を C_{AL} とする。また、 $|C_{OPT}|, |C_{AL}|$ でそれぞれの長さを表すものとする。

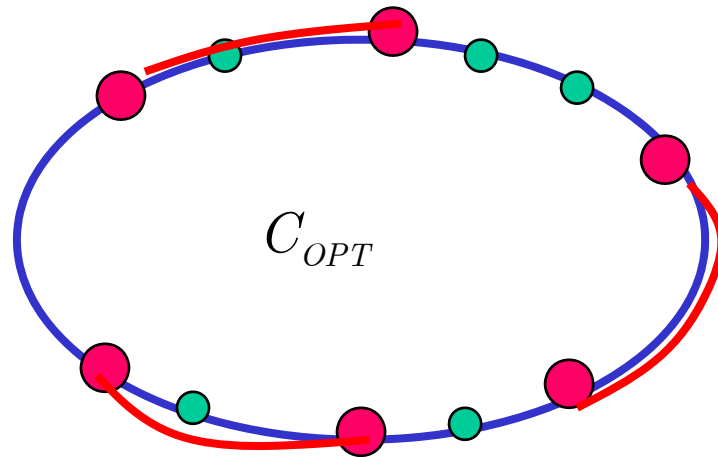
2近似アルゴリズムの解析と同様にして、

$$|T| \leq |C_{OPT}|$$

を得る。

また、 C_{OPT} 上で、次数が奇数の点(奇点)を結ぶパスを考える。 C_{OPT} 上で、奇点を結ぶパスを交互に選ぶことにより、 C_{OPT} 上の辺集合を2つに分類する。

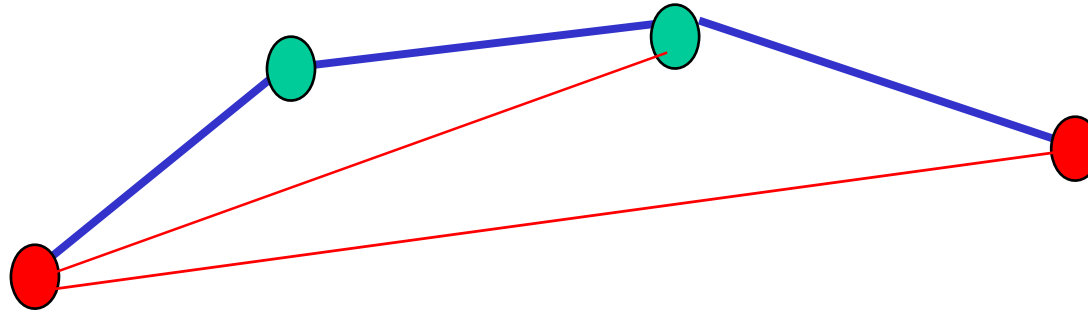
奇点が偶数個しかないことに注意すると、いつもきちんと2分割することができることがわかる。



このように、 C_{OPT} の辺集合を、
 P_1 P_2 と分割できる。もちろん、 $C_{OPT} = P_1 \cup P_2$

よって、 $|C_{OPT}| = |P_1| + |P_2|$

また、三角不等式がなりたっているので、
パスで結ぶより直接辺でたどったほうが短い。



よって、最小マッチング M の重み総和に関して、
次が成り立つ。

$$|M| \leq \min \{|P_1|, |P_2|\}$$

$$\therefore |C_{OPT}| = |P_1| + |P_2|$$

$$\geq 2 \min \{|P_1|, |P_2|\}$$

$$\geq 2|M|$$

一方、アルゴリズムより、

$$|C_{AL}| \leq |T| + |M|$$

以上より、

$$\begin{aligned} |C_{AL}| &\leq |T| + |M| \\ &\leq |C_{opt}| + \frac{1}{2}|C_{opt}| \\ &= \frac{3}{2}|C_{opt}| \\ \therefore \frac{|C_{AL}|}{|C_{opt}|} &\leq 1.5 \end{aligned}$$

このように、いろいろな技法を組み合わせ、近似率の改善が行われる。

NP完全問題に対しても、厳密解でなくてもよければ、近似アルゴリズムの適用を考えると良い。

13-3. ナップザック問題

ナップザック問題の一般形は次のよう書ける。

問題： ナップザック

P 特徴ベクトル $\boldsymbol{x} = {}^t(x_1, x_2, \dots, x_n)$

最大化

$$f(\boldsymbol{x}) = \sum_{i=1}^n v_i x_i$$

条件

$$\sum_{i=1}^n s_i x_i \leq b$$

$$x_i \in \{0, 1\}$$

ナップザック問題における欲張り法 (グリーディ法、Greedy法)

連続ナップザック問題のように、単位価値の高い法から順に選んでいく方法を考察する。このように、部分的に評価関数を改善するだけの方法を**欲張り法 (Greedy法)**という。(欲張り法でも近似アルゴリズムになっていることもある。これらは、問題やアルゴリズムをきちんと解析しないとわからない。)

ナップザックに対する欲張り法

1. 単位価値の高い順にならべる。すなわち、必要なら添え字を付け換えて、

$$\frac{v_1}{s_1} \geq \frac{v_2}{s_2} \geq \dots \geq \frac{v_n}{s_n}$$

とする。

2. $i = 1$ から n まで順番に

$$s_i \leq b - \sum_{j=1}^{i-1} s_j x_j \quad \text{なら} \quad x_i = 1 \quad \text{とし、}$$

そうでないなら

$$x_i = 0$$

とする。

欲張り法の性能

欲張り法で得られる解を $x^g = (x_1^g, x_2^g, \dots, x_n^g)$ とおき、最適解を $x^o = (x_1^o, x_2^o, \dots, x_n^o)$ とおく。

$x^g = x^o$ なら、欲張り法と最適解が一致している。

以下では、 $x^g \neq x^o$ のときを考えよう。

このときは、最適解には採用されたが、欲張り解には採用されなかった最初の要素を考えてその添え字をとする。すなわち、

$$0 = x_j^g < x_j^o = 1$$

このとき、任意の $i \leq j - 1$ に対して、

$$x_j^g \geq x_j^o$$

である。

よって、

$$\begin{aligned} f(\mathbf{x}^g) &= \sum_{i=1}^n v_i x_i^g \geq \sum_{i=1}^j v_i x_i^g = \sum_{i=1}^j v_i x_i^o + \sum_{i=1}^j v_i (x_i^g - x_i^o) \\ &\geq \sum_{i=1}^j v_i x_i^o + \sum_{i=1}^j \frac{v_j}{s_j} s_i (x_i^g - x_i^o) = \sum_{i=1}^j v_i x_i^o + \frac{v_j}{s_j} \left(\sum_{i=1}^j s_i x_i^g - \sum_{i=1}^j s_i x_i^o \right) \end{aligned}$$

という式が成り立つ。ここで、次のようなことに注意する。

- 任意の $i \geq j+1$ に対して、 $\frac{v_i}{s_i} \leq \frac{v_j}{s_j}$ である。
- 欲張り法で j が選ばれなかったことから、

$$\sum_{i=1}^j s_i x_i^g = \sum_{i=1}^{j-1} s_i x_i^g > b - s_j$$

- 最適解でも制約式を満たすので、

$$\begin{aligned} \sum_{i=1}^n s_i x_i^o &\leq b \\ \therefore \sum_{i=1}^j s_i x_i^o &\leq b - \sum_{i=j+1}^n s_i x_i^o \end{aligned}$$

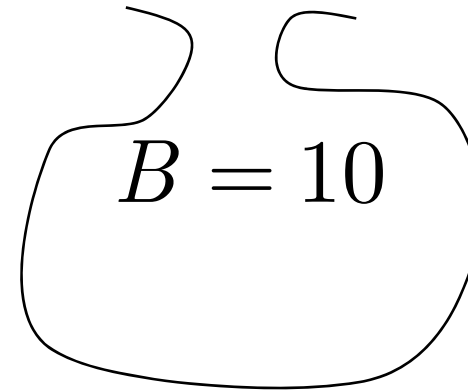
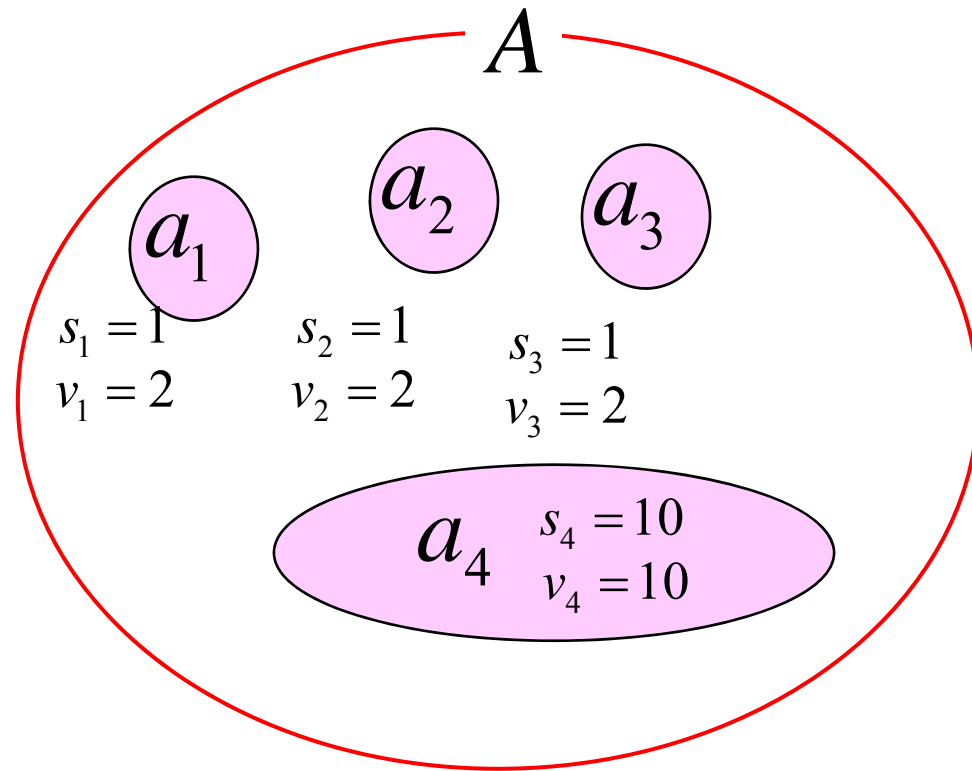
以上より、次の関係式が導ける。

$$\begin{aligned} \frac{v_j}{s_j} \left(\sum_{i=1}^j s_i x_i^g - \sum_{i=1}^j s_i x_i^o \right) &\geq \frac{v_j}{s_j} \left((b - s_j) - \left(b - \sum_{i=j+1}^n s_i x_i^o \right) \right) \\ &\geq \frac{v_j}{s_j} \left(\left(\sum_{i=j+1}^n s_i x_i^o \right) - s_j \right) \geq \sum_{i=j+1}^n v_i x_i^o - v_j \end{aligned}$$

$$\begin{aligned} \therefore f(\mathbf{x}^g) &\geq \sum_{i=1}^j v_i x_i^o + \frac{v_j}{s_j} \left(\left(\sum_{i=j+1}^n s_i x_i^o \right) - s_j \right) \\ &\geq \sum_{i=1}^n v_i x_i^o - v_j \geq f(\mathbf{x}^o) - v_j \geq f(\mathbf{x}^o) - v_{\max} \end{aligned}$$

なお、ここで、 v_{\max} は価値 v_i の最大値である。

欲張り法で悪い評価値を出すインスタンス



$$S = (1, 1, 1, 10)$$

$$V = (2, 2, 2, 10)$$

$$b = 10$$

$$\mathbf{x}^g = (1, 1, 1, 0)$$

$$f(\mathbf{x}^g) = 6$$

$$\mathbf{x}^o = (0, 0, 0, 1)$$

$$f(\mathbf{x}^o) = 10$$

ナップザックの1/2近似アルゴリズム

ナップザックに対する1/2近似アルゴリズム

1. 欲張り法によって解 x^g を求める。
2. 価値が最大のものを一つだけ選ぶ。
3. 上の2つのうちで大きい方を解 x^a として出力する。

近似率

アルゴリズムの出力(解) \mathbf{x}^a とする。このとき、以下の式が成り立つ。

$$f(\mathbf{x}^a) \geq \frac{f(\mathbf{x}^g) + v_{\max}}{2} \geq \frac{1}{2} f(\mathbf{x}^o)$$

以上より、1/2近似アルゴリズムであることがわかる。

ナップザック問題に対するFPTAS

ナップザック問題に対しては、任意の正数 ε に対する近似保証 $(1 - \varepsilon)$ のアルゴリズムが知られている。

ナップザックに対するFPTAS

1. 与えられた ε に対して、 $K = \frac{\varepsilon v_{\max}}{n}$ とおく。
2. 各要素 i に対して、価値を $v_i' = \left\lfloor \frac{v_i}{K} \right\rfloor$ と修正する。
3. 修正した価値のもとで、ナップザックの最適解 x^r を動的計画法で求める。
4. 上記の解 x^r と最初の価値のもとでの最大値を比べて評価値の高いものをアルゴリズムの出力(解) x^a とする。

FPTASの評価

$$v_i' = \left\lfloor \frac{v_i}{K} \right\rfloor \quad \text{とおいていることにより、} \quad 0 \leq v_i - Kv_i' < K$$

よって、任意の解 \boldsymbol{x} に対して、その修正後の評価値 $f'(\boldsymbol{x})$ に関して次式が成立する。

$$0 \leq f(\boldsymbol{x}) - Kf'(\boldsymbol{x}) < nK$$

一方、 \boldsymbol{x}^r は修正した価値での最適解なので

$$f'(\boldsymbol{x}^r) \geq f'(\boldsymbol{x}^o)$$

また、

$$f(\boldsymbol{x}^a) \geq v_{\max}$$

よって、

$$\begin{aligned} f(\mathbf{x}^a) &\geq f(\mathbf{x}^r) \geq Kf'(\mathbf{x}^r) \geq Kf'(\mathbf{x}^o) \\ &\geq f(\mathbf{x}^o) - nK = f(\mathbf{x}^o) - \varepsilon v_{\max} \geq f(\mathbf{x}^o) - \varepsilon f(\mathbf{x}^a) \end{aligned}$$

以上より、

$$\begin{aligned} (1 + \varepsilon)f(\mathbf{x}^a) &\geq f(\mathbf{x}^o) \\ \therefore f(\mathbf{x}^a) &\geq \frac{1}{1 + \varepsilon} f(\mathbf{x}^o) > (1 - \varepsilon)f(\mathbf{x}^o) \end{aligned}$$

計算時間

ステップ3の動的計画法の部分について考察しよう。

まず、動的計画法に基づくナップザックアルゴリズムとして、大きさが決まっているときの価値の最大値を表として構成していた。この動的計画法に基づいた場合、 $O(nb)$ 時間のアルゴリズムが得られた。

ここでは、この動的計画法を以下の方針に切り替える。価値が決まっているときに、大きさの最小値として構成する。このような動的計画法も構成できることに注意する。この場合、価値の最大値は、 v_{\max} であるので、評価値の最大値は、 $O(nv_{\max})$ である。よって、アルゴリズムの計算量は、

$O(n^2 v_{\max})$ 時間である。

FPTASでは修正した値を用いるので、結局、

$$O\left(n^2 \left\lceil \frac{v_{\max}}{K} \right\rceil\right) = O\left(n^2 \left\lceil \frac{n}{\varepsilon} \right\rceil\right) = O\left(\frac{1}{\varepsilon} n^3\right)$$

ナップザック問題の近似可能性

