近似アルゴリズム

アルゴリズム論

中野

Note 10 近似アルゴリズム

2020.5.28 作成

2020.6.23 update 2021.5.31 6.21

近似アルゴリズム

正確な値を求めるのはNP完全でも、
近似的な値を求める多項式時間アルゴリズムは
あるかもしれない。

近似比

最小化問題の近似解と最適解があるとき、近似解/最適解 を、近似解の近似比という。
最小化問題を解く近似アルゴリズムがあるとき、近似解/最適解 の上界を、その近似アルゴリズムの近似比という。

最大化問題の近似解と最適解があるとき、最適解/近似解 を、近似解の近似比という。
最大化問題を解く近似アルゴリズムがあるとき、最適解/近似解 の上界を、その近似アルゴリズムの近似比という。

例
グラフの辺を(できるだけ少ない色を使って)彩色する近似アルゴリズムが、
最小色数の高々1.2倍の色を使うとき----->近似比1.2

近似比は1に近いほうが良い。
近似比と計算時間の間にTrade offがあるアルゴリズムもある。

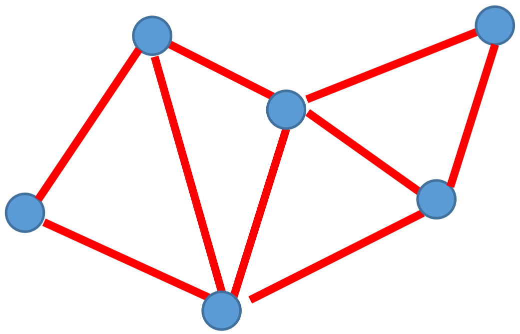
ある定数kに対して、近似比以上の近似解を求めることがNP完全である、
という問題も多数ある。

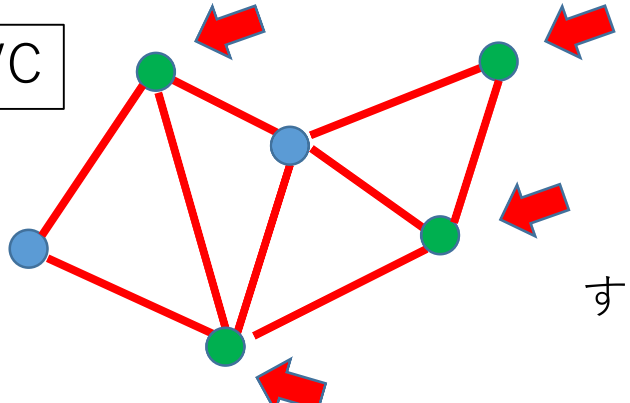
Vertex-cover問題

入力　グラフ

出力　グラフの各辺の少なくとも一方の端点を含む点の集合で
　　(Vertex-coverといいます)点の個数が最小のもの

判定問題はNP完全です。シカシ、近似比2の簡単な近似アルゴリズムがある。





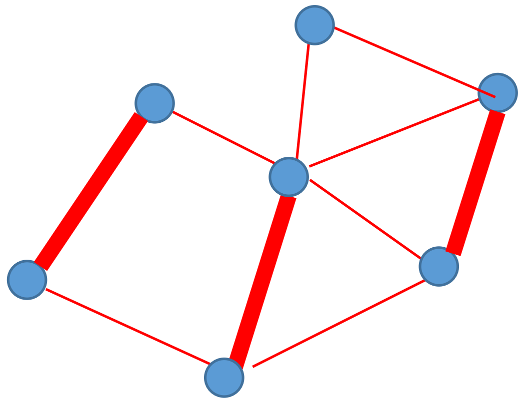
Vertex-cover問題

互いに隣りあわない辺の集合をマッチングという。
極大マッチングの辺集合Mの各辺の両端の点の集合V_Mは、
グラフの各辺の少なくとも一方の端点を含んでいる。
(もしそうでなければ、その辺をマッチングに追加できるので、極大の仮定に反する。)
つまり、V_MはVertex-coverである。
Vertex-cover問題の解の点集合をV_Cとすると、$|V_C| \leqq 2|M|$ である。

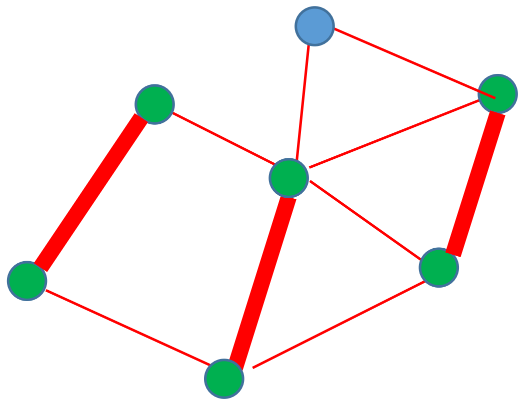
Vertex-coverは、上記の極大マッチングに含まれる各辺の少なくとも一方の端点を
含まなくてはならない。(そうでないと、カバーしていない辺があることになってしまう。)
よって、$|M| <= |V_C|$ である。

つまり、$|V_C| \leqq 2|M| \leqq 2|V_C|$ である。

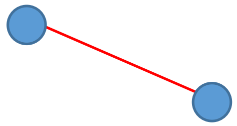
極大マッチングを求めて、マッチングの各辺の両端点からなる
点集合を求めると(2|M|個)、最適解の2倍以下の、Vertex-Coverになっている。
つまり、近似比高々2の解である！



極大マッチング



極大マッチングの両端の点は
Vertex-coverである



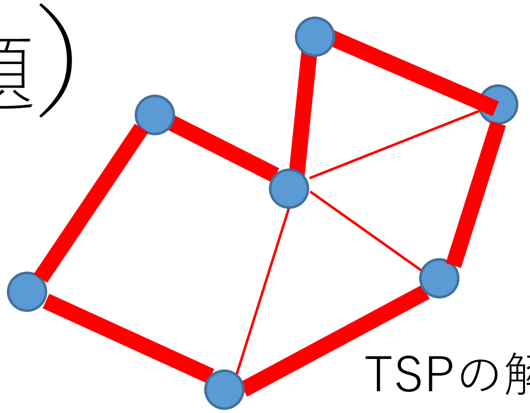
Vertex-coverでなけれ
極大マッチングでない。
矛盾！

Euclidean TSP　（セールスマン巡回問題）

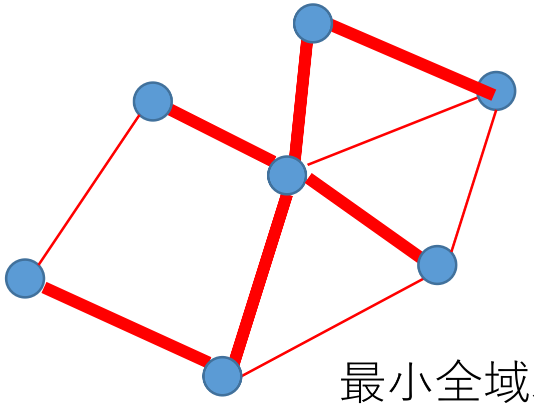
入力　　　n個の点の座標
出力　　　各点をすべて通るサイクルで距離(ユークリッド距離)が最小のもの。
対応するグラフ上の問題はNP完全です。
シカシ、 近似比高々2の簡単な多項式時間アルゴリズムが存在する。

まず、n点の全域最小木を求める。
全域最小木の辺の重みの和　 ＜　 各点をすべて通るサイクルの最小距離　である。
(サイクルから1辺削除すると、全域木なので。)
全域最小木の各辺を2重化し、この上のオイラーサイクル(ひとふで書きサイクル)を作成する。
オイラーサイクルの長さ ＝ 2 x 全域最小木の辺の重みの和
　　　　　　　　＜ 2 x 各点をすべて通るサイクルの最小距離
である。つまり、オイラーサイクルの長さは,各点をすべて通るサイクルで距離が最小のものの、
2倍以下である。つまり、近似比高々2の解である。

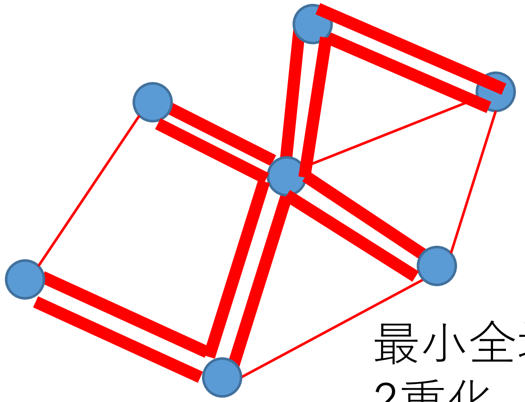
この後で、各点を丁度一回づつ通るように修正してもいい。



TSPの解



最小全域木



最小全域木
2重化

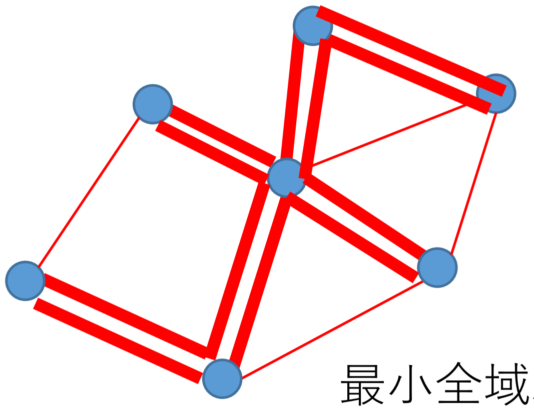
Euclidean TSP （セールスマン巡回問題）

入力　　　　n個の点の座標
出力　　　　各点をすべて通るサイクルで距離(ユークリッド距離)が最小のもの。
対応するグラフ上の問題はNP完全です。
シカシ、 近似比高々2の簡単な多項式時間アルゴリズムが存在する。

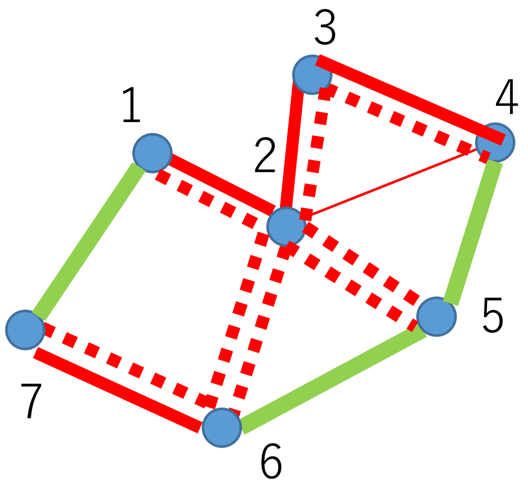
まず、n点の全域最小木を求める。
最小全域木の辺の重みの和　 <　 各点をすべて通るサイクルの最小距離である。
(サイクルから1辺削除すると、全域木なので。)

最小全域木の各辺を2重化し、この上のオイラーサイクル(ひとふで書きサイクル)を作成する。
オイラーサイクルの長さ ＝ 2 x 全域最小木の辺の重みの和
　　　　　　　　　　　　< 2 x 各点をすべて通るサイクルの最小距離
である。つまり、オイラーサイクルの長さは,各点をすべて通るサイクルで距離が最小のものの、2倍以下である。つまり、近似比高々2の解である。

この後で、各点を丁度一回づつ通るように修正してもいい。



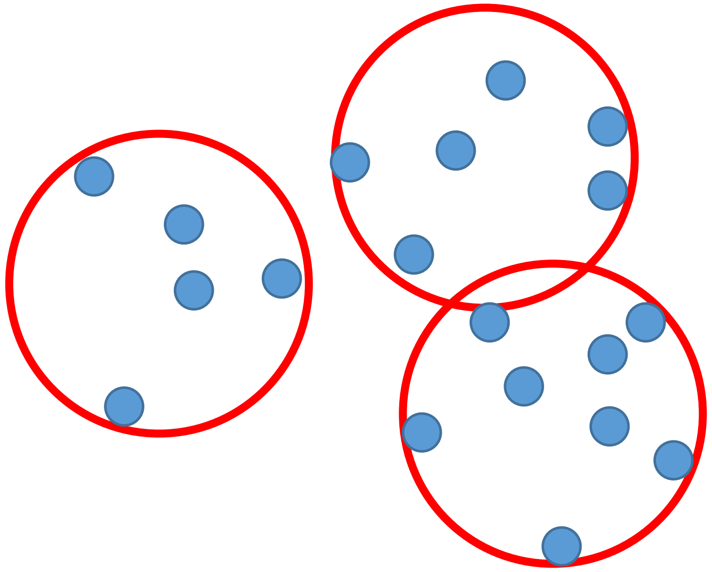
最小全域木



各点を1回ずつ通るように修正

kクラスタリング

入力 平面上のn個の点と整数k
出力 全ての点を含むk個の同じ大きさの円で最小の直径のもの
(最適解の直径をdとする)



Gonzalez(1985) 近似比 2

教師なし学習の基本！
データをグループ分けしたい！

kクラスタリング

各点について一番遠い点を計算する。この値が一番小さな点を1点目として選択する。(ひとつの円で全点をカバーするとき、この点を中心にすると半径が最小です。)

i個の点を選んだ後に、i+1個目の点を次のように選ぶ。

　まだ選ばれていない各点vで、最も近い選択済の点を求める。

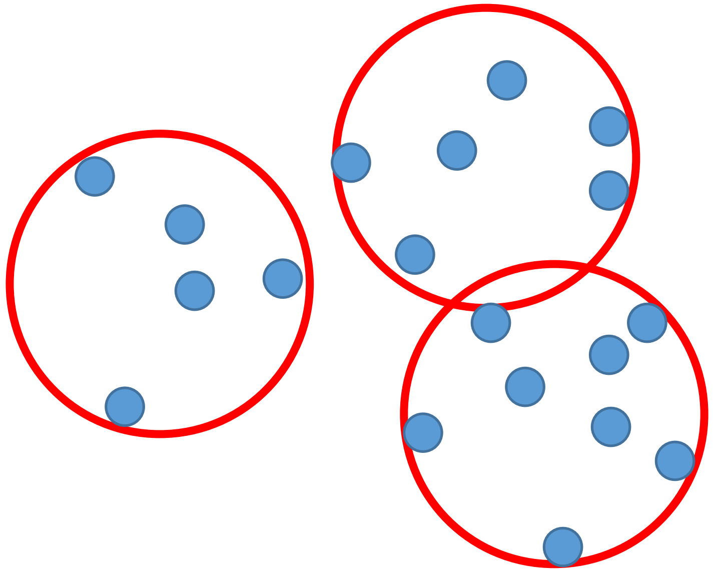
　この距離をL(v)とする。L(v)が最大の点をi+1個目の点として選ぶ。これをk+1点を選ぶまで繰り返す。

k+1個目の点をvとする。このときのvのL(v)を計算する。dを最適な場合の直径とする。

1番目からk番目までの各点を中心とした、半径L(v)の円を近似解とする。

d ≦ 2 L(v) である　(アルゴリズムは 全点を含むような直径2L(v)のk個の円＝近似解求めた。)

(最適な場合の直径dはもっと小さい可能性がある。)



一方、d ≧ L(v) である。　(選んだk+1点のうち、どの2点間も L(v)以上離れているので

最適解のk個の円のうち1つは, アルゴリズムが選んだk+1点のうちの2点をカバーする必要がある.)

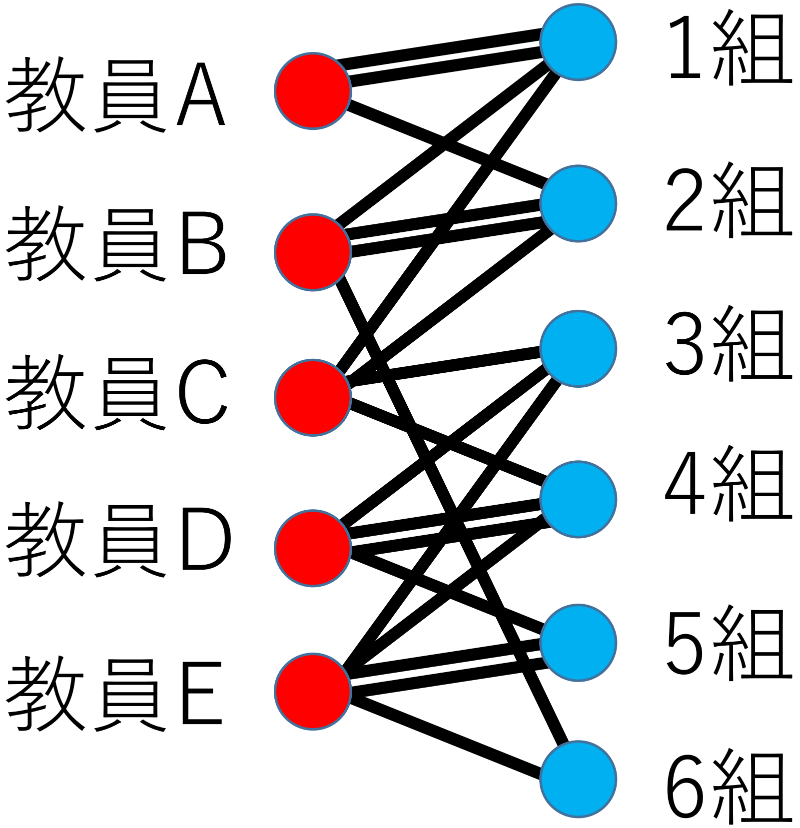
よって d ≦ 2 L(v) ≦ 2 d

つまりアルゴリズムの求めた近似解の直径2L(v)は、最適解の直径dの2倍以下である。

つまり高々2近似である。

多重グラフの辺彩色

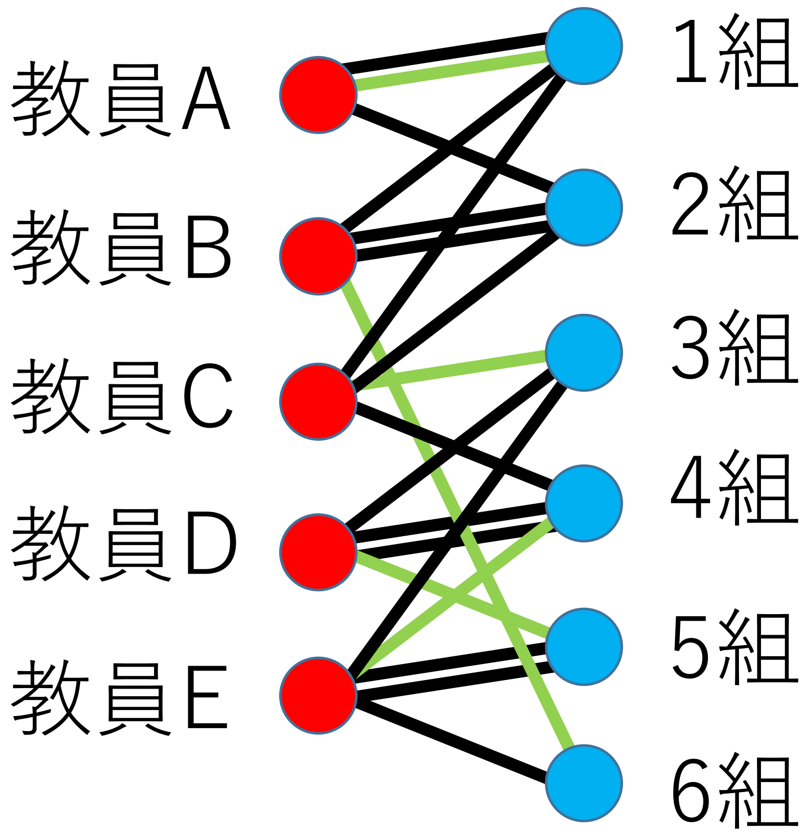
入力　　多重グラフ(2点間に辺が複数あってもいいグラフ)
出力　グラフの隣り合う各辺が異なる色になるように、
最小個数の色で、すべての辺を彩色せよ。



多重グラフの辺彩色

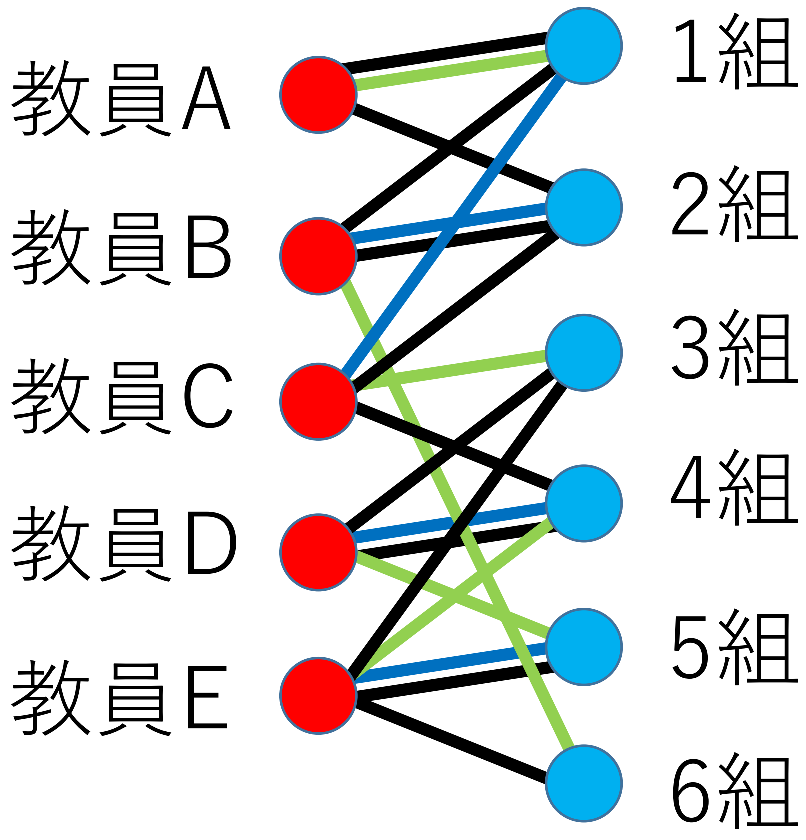
入力　　多重グラフ(2点間に辺が複数あってもいいグラフ)
出力　グラフの隣り合う各辺が異なる色になるように、
最小個数の色で、すべての辺を彩色せよ。

月　1コマ



多重グラフの辺彩色

入力　多重グラフ(2点間に辺が複数あってもいいグラフ)
出力　グラフの隣り合う各辺が異なる色になるように、
最小個数の色で、すべての辺を彩色せよ。



月　1コマ
月　2コマ
月　3コマ
……

Δ色で2部グラフを彩色できます。
1.5Δ色で一般のグラフを彩色する簡単なアルゴリズムがあります。
Δは最大次数