グラフとネットワーク　第 5 回

マッチング：モデル化

岡本 吉央
okamotoy@uec.ac.jp

電気通信大学

2021 年 5 月 14 日

最終更新：2021 年 5 月 5 日　23:12

概要

主題

離散最適化の入門として，次を概説する

- グラフとネットワークを用いた**数理モデル化**
- **アルゴリズム**的解法の背後にある数理

キャッチフレーズ：「本当の離散数学がここから始まる」

達成目標

以下の 4 項目をすべて達成すること

1. グラフやネットワークに関する**用語**を正しく使うことができる
2. 現実世界の諸問題をグラフやネットワークで表現し，**数理モデル**を構築できる
3. アルゴリズム的解法の背後にある数理，特に，**最小最大定理**の重要性を説明でき，それを用いて最適性の**証明**ができる
4. グラフとネットワークに関する簡単な数学的事実を**証明**できる

スケジュール 前半 (予定)

0	ガイダンス	(4/9)
1	グラフの定義と次数：数理	(4/16)
2	道と閉路：数理	(4/23)
3	木：数理	(4/30)
4	マッチング：数理	(5/7)
5	マッチング：モデル化	(5/14)
6	最大流：数理	(5/21)
7	最大流：モデル化 (1) 割当	(5/28)

注意：予定の変更もありうる

概要

今日の目標

マッチングを使って問題を解決する例を見る

- XML のパターン・マッチング
- ドナー交換腎移植
- 臨床研修病院選考

数理モデルの構築 と 応用

格言

- 「応用」ということばの意味は非常にあいまい
- 「応用」というときは，学問の階層を意識する

- 研究には段階がある
- 基礎研究の応用先は別の基礎研究かもしれない
- 基礎研究の応用先は応用研究かもしれない
- 応用研究の応用先は別の応用研究かもしれない
- 応用研究の応用先は基礎研究かもしれない
- 「応用」は実世界応用を意味しないかもしれない

目次

XML (Extensible Markup Language)

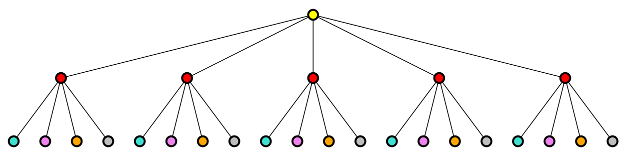
JIS による訳語：拡張可能なマーク付け言語

XML 文書の例

```
<breakfast_menu>
 <food>
  <name>Belgian Waffles</name>
  <price>$5.95</price>
  <description>
   Two of our famous Belgian Waffles with
   plenty of real maple syrup
  </description>
  <calories>650</calories>
 </food>
 <food>
  <name>Strawberry Belgian Waffles</name>
  <price>$7.95</price>
  <description>
   Light Belgian waffles covered with
   strawberries and whipped cream
  </description>
  <calories>900</calories>
 </food>
 <food>
  <name>Berry-Berry Belgian Waffles</name>
  <price>$8.95</price>
  <description>
   Light Belgian waffles covered with an
   assortment of fresh berries and whipped
   cream
  </description>
  <calories>900</calories>
 </food>
 <food>
  <name>French Toast</name>
  <price>$4.50</price>
  <description>
   Thick slices made from our homemade
   sourdough bread
  </description>
  <calories>600</calories>
 </food>
 <food>
  <name>Homestyle Breakfast</name>
  <price>$6.95</price>
  <description>
   Two eggs, bacon or sausage, toast, and
   our ever-popular hash browns
  </description>
  <calories>950</calories>
 </food>
</breakfast_menu>
```

https://www.w3schools.com/xml/simple.xml

XML 文書から色付き木へ (1)



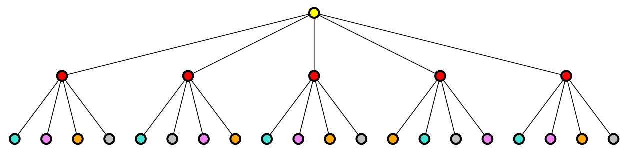
- breakfast_menu
- name
- description
- food
- price
- calories

木 は 入れ子構造 を表現するときに 有効 ⇝ 根付き木

注

同じような考え方は他のデータ記述，プログラミング言語でも可能

XML 文書から色付き木へ (2)



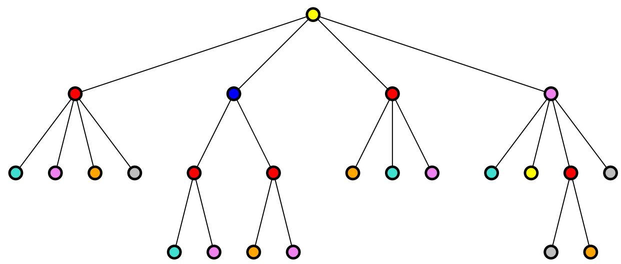
- 🟡 breakfast_menu
- 🔴 food
- name
- price
- description
- calories

このデータでは 「子の順序」 は重要ではない

注

他のデータ記述，プログラミング言語では順序が重要であることもある

色付き根付き木：もう少し複雑な例



木のパターン・マッチング

木のパターン・マッチング

今から行いたいこと

- 「色付き根付き木」を定義する
- 「木のパターン・マッチング」を定義する

目標

木のパターン・マッチングを解く

そのために，グラフのマッチングを利用する

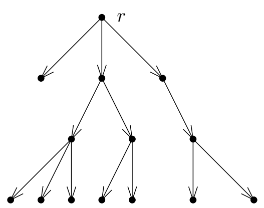
有向木

根付き木は有向木としてモデル化 (定義) する

定義：有向木とは？

有向グラフ $G = (V, A)$ が **有向木** であるとは 次を満たすこと

- ある頂点 $r \in V$ が存在して，任意の頂点 $v \in V$ に対して，
  r から v へ至る有向道がただ 1 つ存在する



用語

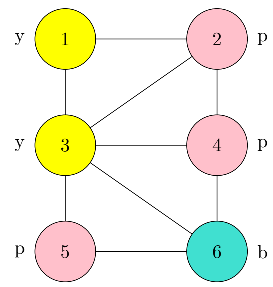
- 有向木を **外向木** と呼ぶこともある
- r は有向木 G の **根** と呼ばれる

頂点への色の割当

グラフ $G = (V, E)$，色の集合 L

色の割当

G の頂点への色の割当は写像 $c \colon V \to L$ として表せる



$L = \{$p ●, y ●, b ●$\}$

- $c(1) =$ y
- $c(2) =$ p
- $c(3) =$ y
- $c(4) =$ p
- $c(5) =$ p
- $c(6) =$ b

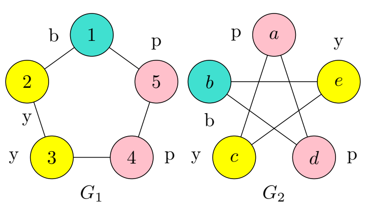
色付き同型性

グラフ $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$，色の集合 L，
色の割当 $c_1 \colon V_1 \to L, c_2 \colon V_2 \to L$

定義：色付き同型写像とは？

(G_1, c_1) から (G_2, c_2) への **色付き同型写像** とは，
次を満たす全単射 $\varphi \colon V_1 \to V_2$ のこと

- φ は G_1 から G_2 への同型写像 (辺の保存)
- 任意の頂点 $v \in V_1$ に対して，$c_1(v) = c_2(\varphi(v))$ (色の保存)



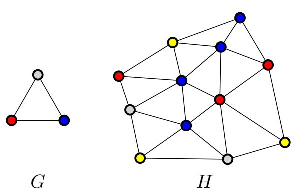
- $\varphi(1) = b$
- $\varphi(2) = e$
- $\varphi(3) = c$
- $\varphi(4) = a$
- $\varphi(5) = d$

グラフのパターン・マッチング

グラフ $G = (V_G, E_G), H = (V_H, E_H)$，色の集合 L，
色の割当 $c_G \colon V_G \to L, c_H \colon V_H \to L$

定義：グラフのパターン・マッチングとは？

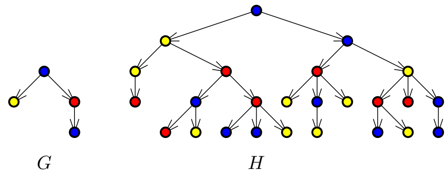
H の部分グラフ H' で，G から H' への色付き同型写像を持つものを
見つける問題



- 「見つける」は「1 つ見つける」を意味することもあるし，
  「全部見つける」を意味することもある

色付き有向木のパターン・マッチング

以下，グラフのパターン・マッチングにおける 次の場合を考える
- G と H は有向木である
- H の部分グラフ H' で，G から H' への色付き同型写像を持つものを 1 つ見つける

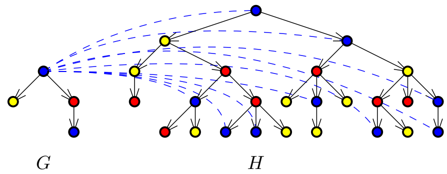


今から行うこと

グラフのマッチングを使って，この問題を解く

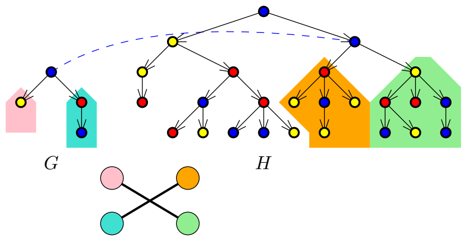
色付き有向木のパターン・マッチング：アイディア (1)

根に対応する頂点を考える



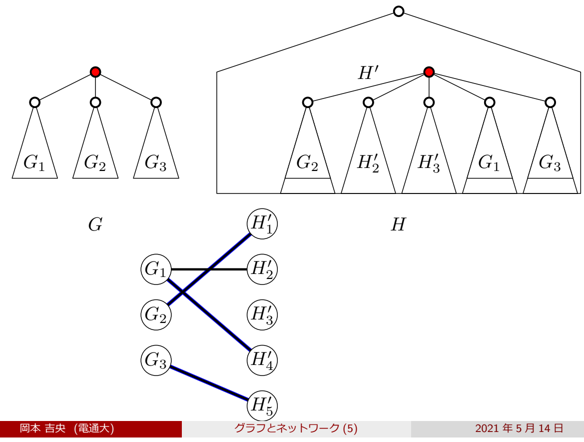
色付き有向木のパターン・マッチング：アイディア (2)

部分木の対応を考える ⇝ 二部グラフのマッチング

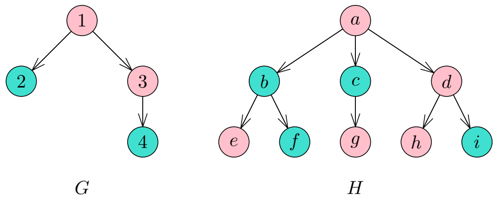


●と●の間に辺 ⇔
●の部分グラフ H'' で次を満たすものが存在
- ●の根 = H'' の根
- ●から H'' への色付き同型写像が存在

色付き有向木のパターン・マッチング：一般論



色付き有向木のパターン・マッチング：例



	e	f	b	g	c	h	i	d	a
2	×	○	○	×	○	×	○	×	×
4	×	○	○	×	○	×	○	×	×
3	×	×	×	×	×	×	×	○	○
1	×	×	×	×	×	×	×	×	○

色付き有向木のパターン・マッチング：アルゴリズム (1)

入力

有向木 $G = (V_G, E_G), H = (V_H, E_H)$，色の集合 L，
色の割当 $c_G \colon V_G \to L, c_H \colon V_H \to L$

アルゴリズム

1. V_G の要素を後行順で $v_1, v_2, \ldots, v_n$ と並べる
2. V_H の要素を後行順 $u_1, u_2, \ldots, u_m$ と並べる
3. 各 $i \in \{1, 2, \ldots, n\}$ と $j \in \{1, 2, \ldots, m\}$ に対して小さい順に次を行う
   1. $c_G(v_i) \neq c_H(u_j)$ ならば，$t[i, j] = $ ×

次のページに続く

色付き有向木のパターン・マッチング：アルゴリズム (2)

アルゴリズム (続き)

3. 各 $i \in \{1, 2, \ldots, n\}$ と $j \in \{1, 2, \ldots, m\}$ に対して小さい順に次を行う
   2. $c_G(v_i) = c_H(u_j)$ ならば，次を行う
      - a. 次のような (無向) 二部グラフ $B = (V_B, E_B)$ を作る

$$X = v_i \text{ を始点とする } G \text{ の弧の終点全体}$$
$$Y = u_j \text{ を始点とする } G \text{ の弧の終点全体}$$
$$V_B = X \cup Y$$
$$E_B = \{\{v_i, u_j\} \mid t[i, j] = \bigcirc\}$$

      - b. B が X を飽和するマッチングを持つならば，$t[i, j] = $ ○
        そうでなければ，$t[i, j] = $ ×
4. ある $j \in \{1, \ldots, m\}$ に対して，$t[n, j] = $ ○ であるとき，「Yes」を出力
   そうでないとき，「No」を出力

色付き有向木のパターン・マッチング：文献案内と補足

- 紹介したアルゴリズムは (本質的に) 次の論文による．
  - D. W. Matula. Subtree isomorphism in $O(n^{5/2})$. Ann. Discrete Math. 2 (1978) 91–106.

  この論文に Edmonds も同様なアルゴリズムを発見したと書いてあり，後年，Chung も同様なアルゴリズムを再発見し，論文として出版した．
- 紹介したアルゴリズムの計算量は「二部グラフの最大マッチング」を計算するアルゴリズムの計算量に依存する
  - 例えば，Hopcroft–Karp のアルゴリズム (というよく知られたもの) を使うと紹介したアルゴリズムの計算量は $O(|V_G|^{3/2}|V_H|)$ になる
- 今では，それよりも速いアルゴリズムが知られている
  - R. Shamir, D. Tsur. Faster subtree isomorphism. J. Algor. 33 (1999) 267–280.

目次

腎移植

腎移植とは？
ドナーからレシピエントに腎臓を移し植える医療行為のこと

ドナー： 提供者
レシピエント： 受給者

ドナー (D) $\longrightarrow$ レシピエント (R)

腎移植の主な方法
- 生体腎移植
- 死体腎移植 (献腎移植)

重要な前提
- 腎臓はヒトにおいて 2 つ存在する臓器であり，片方の腎臓があれば恒常性は維持できる (医学的な前提)
- 腎移植に金銭の授受は伴わない (倫理的・法的な前提)

生体腎移植 と 解決すべき問題

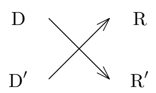
生体腎移植における根本的な問題
ドナーとレシピエントの「型」が合わないと移植が難しい

$$\mathrm{D} \longrightarrow \mathrm{R}$$

患者とその親族の型が合うとは限らない

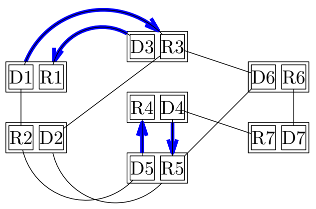
ドナー交換腎移植

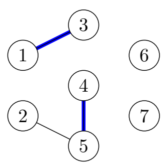
ドナー交換腎移植とは？
ドナーとレシピエントを交換して行う腎移植



ドナー交換腎移植を実現するには…

「ドナーとレシピエントの組」が多く登録されたプールが必要

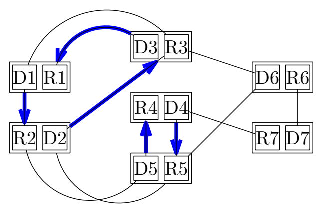


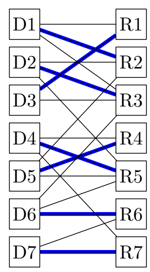


$\rightsquigarrow$ 最大マッチング

ドミノ型腎移植を実現するには…

このような腎移植も可能かもしれない (ドミノ型腎移植)



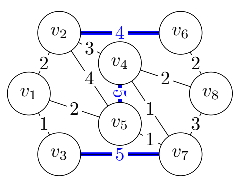


$\rightsquigarrow$ 二部グラフの最大重みマッチング

最大重みマッチング

無向グラフ $G = (V, E)$
各辺 $e \in E$ に対する非負重み $w(e) \geq 0$ (辺重み関数 $w \colon E \to \mathbb{R}$)

定義：最大重みマッチングとは？
w に関する G の**最大重みマッチング**とは
G のマッチング $M \subseteq E$ で，
G の任意のマッチング M' に対して $\sum_{e \in M} w(e) \geq \sum_{e \in M'} w(e)$ を満たすもの



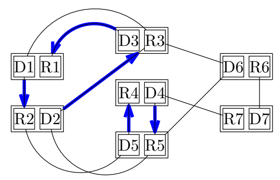
注
最大重みマッチングが
最大マッチングであるとは限らない

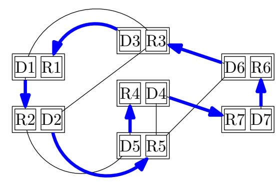
長い有向閉路の得失

長い有向閉路を許すことのメリット
- 多くのレシピエントに腎臓を提供できるようになる

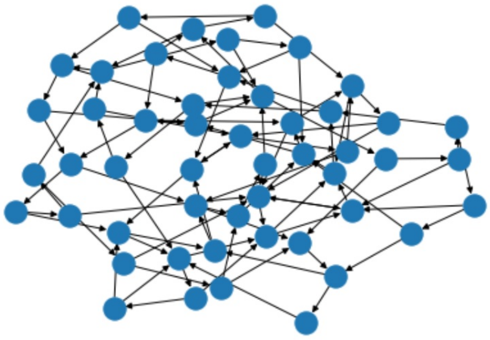
長い有向閉路を許すことのデメリット
- 同時に執刀しなくてはならない手術が増加する

- 長さ 3 の有向閉路 $\rightsquigarrow$ 6 個の手術を同時に行う
- 長さ 7 の有向閉路 $\rightsquigarrow$ 14 個の手術を同時に行う



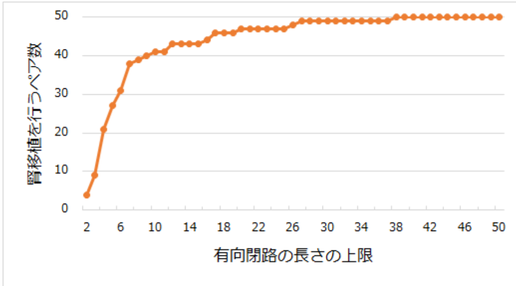


少し大きな例で計算してみた (1)



頂点数 50，弧数 98，最大閉路の長さ 50

少し大きな例で計算してみた (2)



- 長さ 2 の閉路まで ⇝ 4 ペア
- 長さ 3 の閉路まで ⇝ 9 ペア
- 長さ 4 の閉路まで ⇝ 21 ペア
- 長さ 10 の閉路まで ⇝ 41 ペア
- 長さ 16 の閉路まで ⇝ 44 ペア
- 長さ 38 の閉路まで ⇝ 50 ペア

ドナー交換腎移植：解き方のまとめ

長さ 2 の閉路まで許す

⇝ 最大マッチングを求める問題　(効率よく解ける問題)

長さ k の閉路まで許す (ただし，$k \geq 3$)

⇝ NP 困難な問題　(効率よい解き方が知られていない問題)

閉路の長さに制限がない

⇝ 二部グラフの最大重みマッチングを求める問題　(効率よく解ける問題)

他にも考えるべき条件 (制約) があるので，実際はこれほど単純ではない

長さ k の閉路まで許す問題 — ちゃんと定義する

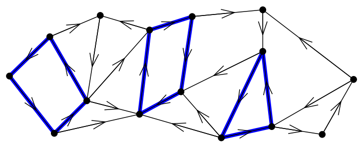
有向グラフ $G = (V, A)$，自然数 $k \geq 2$

「長さ k の閉路まで許す問題」の定義

有向閉路の集合 $\{C_1, C_2, \ldots, C_t\}$ で次を満たすものを見つける

1. $C_1, C_2, \ldots, C_t$ の長さはどれも k 以下である
2. どの頂点も $C_1, C_2, \ldots, C_t$ の中の 2 つ以上に含まれない
3. $C_1, C_2, \ldots, C_t$ に含まれない頂点の数が最小になる

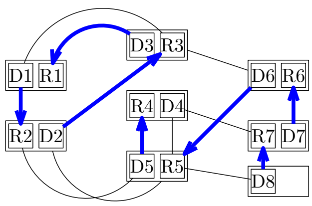
例：$k = 4$ のとき



ドナー交換腎移植：献体がある場合

献体はレシピエントのいないペアとして扱うことができる

- 献体は優先的に割り当てたい (腎臓を長期間保存できないので)



⇝ 二部グラフの最大重みマッチング　(演習問題)

ドナー交換腎移植：補足

ドナー交換腎移植の歴史

- 1986 年：概念の提唱 (Rapaport による)
- 1991 年：韓国で世界初のドナー交換腎移植
- 1999 年：ヨーロッパで初のドナー交換腎移植 (スイス)
- 2000 年：アメリカで初のドナー交換腎移植
- 2003 年：日本で初のドナー交換腎移植

現在，日本ではドナー交換腎移植が推進されていない

ドナー交換腎移植に関する数理 (経済学的) 研究

- Roth, Sönmez, Ünver (2004, 2005, 2007)
- Ashlagi, Roth (2012)
- その後，多くの研究

⇝ Roth はノーベル経済学賞 (2012) を受賞

目次

臨床研修病院選考

日本では 2004 年に始まった (諸外国でも様々な形で存在する)

医師臨床研修制度とは？

医師国家試験を受けて医師免許を受けた者に対して
2 年間の臨床研修を義務化するもの

臨床研修 ＝ 病院で実際に診療等を行う研修

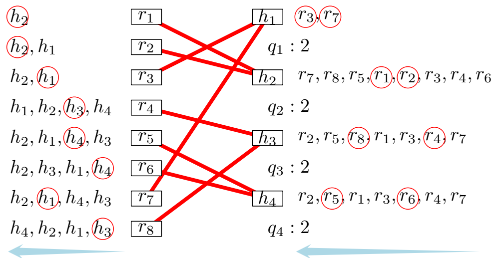
2004 年に始まった制度では

- 臨床研修を希望する者が研修を行う病院を希望できる

⇝ 問題：どのように研修医を病院に割り当てるのか？

研修医配属問題 (1)

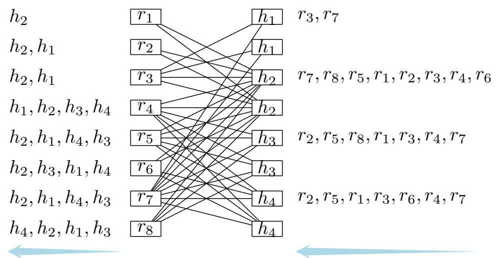
研修医 (resident) 病院 (hospital)



https://www.jrmp.jp/ の例

研修医配属問題 (2)

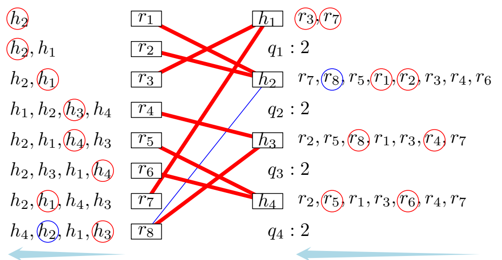
研修医 (resident) 病院 (hospital)



$\rightsquigarrow$ 最大重みマッチング

研修医配属問題 (3)

r_8 と h_2 はこの配属を受け入れられない



「ブロッキング・ペア」という考え方

ブロッキング・ペア

研修医の配属割当 μ

定義：ブロッキング・ペアとは？

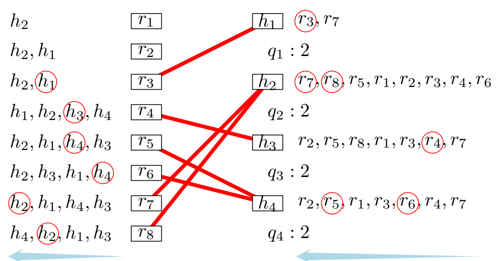
μ における **ブロッキング・ペア** とは，
研修医 r と病院 h の組 (r, h) で次の条件をすべて満たすもの

1. μ において，r は h に配属されていない
2. r は h を希望順に入れ，h も r を希望順に入れている
3. r が次のいずれかを満たす
   - r は μ でどの病院にも配属されていない
   - r は μ で配属される病院より，h を好む
4. h が次のいずれかを満たす
   - h は μ で空席を持て余している
   - h は μ で配属される研修医の誰かより，r を好む

ブロッキング・ペア (r, h) は μ を受け入れられない

ブロッキング・ペア：例

この配属割当には，ブロッキング・ペアが存在しない



安定な配属割当

定義：安定な配属割当

研修医の配属割当が **安定** であるとは，
それがブロッキング・ペアを持たないこと

- 安定な配属割当が望ましい
- 安定な配属割当の中で，研修医ができるだけ多く配属されるものが望ましい

根本的な問題

安定な配属割当はどんな場合でも存在するのか？

安定な配属割当：存在性

性質：安定な配属割当の存在性 (Gale, Shapley '62)

研修医配属問題において，どんな場合でも，安定な配属割当が存在する

証明：次の単純な場合に限って証明を行う (一般の場合は演習問題)

- どの病院の定員も 1 である
- 研修医の総数 ＝ 病院の総数
- 各研修医の希望順には，すべての病院が入っている
- 各病院の希望順には，すべての研修医が入っている

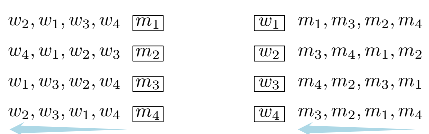
この場合

「研修医配属問題」を「安定結婚問題」と呼ぶことが多い

- 「研修医」を「男性」，「病院」を「女性」とみなす
- 「安定な配属割当」を「安定な結婚」と呼ぶ

安定結婚問題

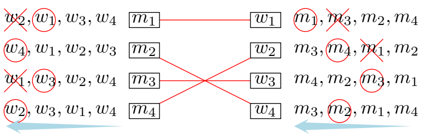
男性 女性



証明の手法

実際に，安定な結婚を求めるアルゴリズムを与える
$\rightsquigarrow$ **受入保留アルゴリズム**

受入保留アルゴリズム：動作例



受入保留アルゴリズム

まだパートナーが決まっていない男性 m がいる限り

1. m のリストの中で，まだ断られていない最上位の女性を w とする
2. w のパートナーが決まっていない，あるいは，
   w が m を w の現在のパートナーよりも好む ならば
   - (w は現在のパートナーとの関係を解消する (断る))
   - m と w はパートナーとなる
3. そうでないならば，w は m を断る

これで最終的に割当 (完全マッチング) が得られる

受入保留アルゴリズム：停止性

受入保留アルゴリズムの性質：停止性

安定結婚問題において，受入保留アルゴリズムは必ず停止し，
すべての男性と女性にパートナーを作る

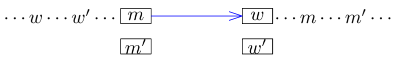
証明 (背理法)：受入保留アルゴリズムが停止しない場合を考える
- そのとき，ある男性 m とある女性 w にパートナーがいない
- アルゴリズムにおいて，いつか m は w に申し込む
- つまり，w にはパートナーがいることになり，矛盾 □

受入保留アルゴリズム：安定性 (1)

受入保留アルゴリズムの性質：安定性

安定結婚問題において，受入保留アルゴリズムの出力は
必ず 安定な割当である

証明 (背理法)：ブロッキング・ペア (m, w) が存在すると仮定する
- アルゴリズムの出力で，m のパートナーは w' であり，
  w のパートナーは m' であるとする
- (m, w) はブロッキング・ペアなので，m は w' よりも w を好む
- アルゴリズムの出力で，m のパートナーは w' なので，
  アルゴリズムにおいて，m は w に申し込みをしている



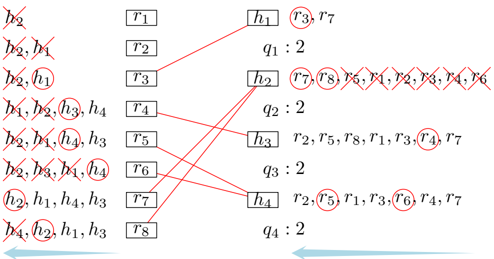
受入保留アルゴリズム：安定性 (2)

証明 (続き)：
- つまり，アルゴリズムの出力において，
  w のパートナーは w であるか，m よりも w に好まれる男性である
- m' は m よりも w に好まれる男性ではないので，矛盾 □

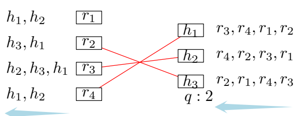
結論 (Gale, Shapley '62)

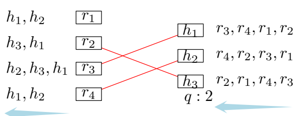
安定結婚問題において，どんな場合でも，安定な割当が存在する

研修医配属問題に 受入保留アルゴリズムを 適用してみる



安定な割当は一意ではない





性質 (証明はしないが難しくない)

研修医配属問題において，
どの安定な割当においても，割り当てられる研修医数は同じ

性質：より強いことが言える (田舎の病院定理)

研修医配属問題において，
- どの安定な割当でも，配属病院が決まる研修医は変わらない
- どの安定な割当でも，病院に配属される研修医数は変わらない

安定な配属割当：まとめ

定義：安定な配属割当

研修医の配属割当が **安定** であるとは，
それがブロッキング・ペアを持たないこと

- 安定な配属割当が望ましい
- 安定な配属割当の中で，研修医ができるだけ多く配属される
  ものが望ましい

根本的な問題

安定な配属割当はどんな場合でも存在するのか？

解決

- 安定な配属割当はどんな場合でも存在する
  (受入保留アルゴリズムで簡単に 1 つ見つけられる)
- どの安定な配属割当でも，配属される研修医の数は変わらない

臨床研修病院選考：補足 (1)

**日本の 2004 年に始まった医師臨床研修制度では，
受入保留アルゴリズムを使っている (ようである)**

- 参考：医師臨床研修マッチング協議会 `https://www.jrmp.jp/`

諸外国では既に使われていた

- 例：アメリカ NMRP (1952 年より)

**諸大学では，学科配属，研究室配属で
受入保留アルゴリズム (の変種) を使っている (ようである)**

- 例：お茶の水大学理学部情報科学科
  `https://www.is.ocha.ac.jp/~kudo/lab_algo.html`
- 例：慶應義塾大学理工学部

臨床研修病院選考：補足 (2)

安定な割当に関する研究は次の論文が発端となっている

- D. Gale and L. S. Shapley, College Admissions and the Stability of Marriage. American Mathematical Monthly 69 (1962) 9–14

$\rightsquigarrow$ L. Shapley はノーベル経済学賞 (2012 年) を受賞

その後，研究が進んでいる

- 1 対 1 の割当：安定結婚問題，...
- 多対 1 の割当：研修医配属問題，学校選択問題，...
- 多対多の割当：研修医配属問題 (昔のイギリス)，...
- 非二部グラフでの割当：安定ルームメイト問題，...
- ⋮

目次

今日のまとめ

今日の目標

マッチングを使って問題を解決する例を見る

- XML のパターン・マッチング
- ドナー交換腎移植
- 臨床研修病院選考

これから，皆さんに考えてほしいこと

既存のモデル化・応用例を知ること以上に
自分なりのモデル化・応用例を考えること