

よい組合せの見つけ方:

食べたいものから順に食べた方がいい理由

藤井 海斗（国立情報学研究所）

よい組合せを見つける方法について研究する情報学の一分野

よい組合せを見つける方法について研究する情報学の一分野



ファミレス料理の
最高の組合せは？



ファミレスに来た人

よい組合せを見つける方法について研究する情報学の一分野



ファミレス料理の
最高の組合せは？



ファミレスに来た人

よい組合せを見つける方法について研究する情報学の一分野

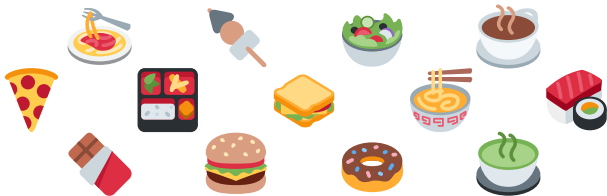


ファミレス料理の
最高の組合せは？



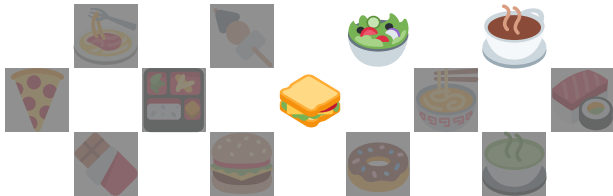
ファミレスに来た人

すべての組合せに「よさ」を表す点数を割り当てる関数



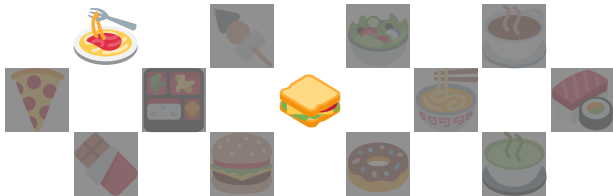
すべての組合せに「よさ」を表す点数を割り当てる関数

$$f(\{\text{salad}, \text{sandwich}, \text{coffee}\}) = 50$$



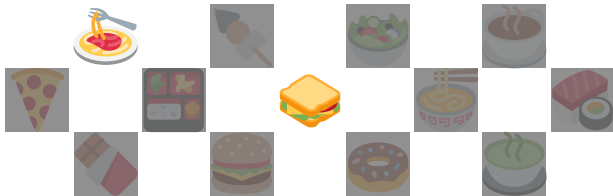
すべての組合せに「よさ」を表す点数を割り当てる関数

$$f(\{\text{🍝}, \text{🍔}\}) = 20$$



すべての組合せに「よさ」を表す点数を割り当てる関数

$$f(\{\text{🍝}, \text{🍔}\}) = 20$$



問

集合関数の値が最大となる組合せは？

素朴な方法

すべての組合せを一つずつ調べる

素朴な方法

すべての組合せを一つずつ調べる

- 料理数が 3 個のとき：組合せは全部で $2^3 = 8$ 種類

 ,  ,  の 3 料理のとき

$\{\}$, $\{\text{bread}\}$, $\{\text{orange}\}$, $\{\text{salad}\}$, $\{\text{bread}, \text{orange}\}$, $\{\text{bread}, \text{salad}\}$, $\{\text{orange}, \text{salad}\}$, $\{\text{bread}, \text{orange}, \text{salad}\}$ の 8 種類

素朴な方法

すべての組合せを一つずつ調べる

- 料理数が 3 個のとき：組合せは全部で $2^3 = 8$ 種類

 ,  ,  の 3 料理のとき

$\{\}, \{\text{bread}\}, \{\text{orange}\}, \{\text{salad}\}, \{\text{bread}, \text{orange}\}, \{\text{bread}, \text{salad}\}, \{\text{orange}, \text{salad}\}, \{\text{bread}, \text{orange}, \text{salad}\}$ の 8 種類

- 料理数が 10 個のとき：組合せは全部で $2^{10} = 1024$ 種類

素朴な方法

すべての組合せを一つずつ調べる

- 料理数が 3 個のとき：組合せは全部で $2^3 = 8$ 種類

 ,  ,  の 3 料理のとき

$\{\}, \{\text{bread}\}, \{\text{orange}\}, \{\text{salad}\}, \{\text{bread}, \text{orange}\}, \{\text{bread}, \text{salad}\}, \{\text{orange}, \text{salad}\}, \{\text{bread}, \text{orange}, \text{salad}\}$ の 8 種類

- 料理数が 10 個のとき：組合せは全部で $2^{10} = 1024$ 種類
- 料理数が 100 個のとき：組合せは全部で $2^{100} \approx 10^{30}$ 種類

素朴な方法

すべての組合せを一つずつ調べる

- 料理数が 3 個のとき：組合せは全部で $2^3 = 8$ 種類

 ,  ,  の 3 料理のとき

$\{\}, \{\text{bread}\}, \{\text{orange}\}, \{\text{salad}\}, \{\text{bread}, \text{orange}\}, \{\text{bread}, \text{salad}\}, \{\text{orange}, \text{salad}\}, \{\text{bread}, \text{orange}, \text{salad}\}$ の 8 種類

- 料理数が 10 個のとき：組合せは全部で $2^{10} = 1024$ 種類
- 料理数が 100 個のとき：組合せは全部で $2^{100} \approx 10^{30}$ 種類

料理数が増えるにしたがって組合せの数は天文学的に大きくなる

➡ もっと効率のよい探し方（アルゴリズム）が必要

関数値が最も増える要素を順に追加していくアルゴリズム



$$f(\{\text{🍜}\}) = 20$$

$$f(\{\text{🍷}\}) = 50 \leftarrow \text{最大}$$

$$f(\{\text{🥦}\}) = 10$$

⋮

$$f(\{\text{🍌}\}) = 5$$

$$f(\{\text{🍱}\}) = 10$$

問題設定：

3 つの料理を選ぶ

関数値が最も増える要素を順に追加していくアルゴリズム



$$f(\{\text{🍜}\}) = 20$$

$$f(\{\text{🍷}\}) = 50 \leftarrow \text{最大}$$

$$f(\{\text{🥦}\}) = 10$$

⋮

$$f(\{\text{🍌}\}) = 5$$

$$f(\{\text{🍱}\}) = 10$$

問題設定：

3 つの料理を選ぶ

関数値が最も増える要素を順に追加していくアルゴリズム



問題設定：

3 つの料理を選ぶ

$$f(\{\text{🍔}, \text{🍜}\}) = 55$$

$$f(\{\text{🍔}, \text{🥦}\}) = 60 \leftarrow \text{最大}$$

$$f(\{\text{🍔}, \text{🍕}\}) = 54$$

⋮

$$f(\{\text{🍔}, \text{🍌}\}) = 55$$

$$f(\{\text{🍔}, \text{🍙}\}) = 50$$

関数値が最も増える要素を順に追加していくアルゴリズム



問題設定：

3 つの料理を選ぶ

$$f(\{\text{Sandwich}, \text{Broccoli}, \text{Ramen}\}) = 61$$

$$f(\{\text{Sandwich}, \text{Broccoli}, \text{Pizza}\}) = 62$$

$$f(\{\text{Sandwich}, \text{Broccoli}, \text{Ice cream}\}) = 61$$

⋮

$$f(\{\text{Sandwich}, \text{Broccoli}, \text{Soup}\}) = 65 \leftarrow \text{最大}$$

$$f(\{\text{Sandwich}, \text{Broccoli}, \text{Onigiri}\}) = 60$$

関数値が最も増える要素を順に追加していくアルゴリズム



問題設定：

3 つの料理を選ぶ

問

どんな集合関数に対して貪欲法がうまくいくのか？

手元にある要素が増えると新しい要素による増分が減る

手元にある要素が増えると新しい要素による増分が減る

$$f(\{\text{🍕}\}) - f(\{\})$$



何も持っていないときに

🍕 をもらった場合の増分

手元にある要素が増えると新しい要素による増分が減る

$$f(\{\text{🍕}\}) - f(\{\})$$

何も持っていないときに

🍕 をもらった場合の増分

$$f(\{\text{🍕}, \text{🍷}\}) - f(\{\text{🍷}\})$$

🍷 を持っているときに

🍕 をもらった場合の増分

手元にある要素が増えると新しい要素による増分が減る

$$f(\{\text{🍕}\}) - f(\{\})$$

何も持っていないときに

🍕 をもらった場合の増分

$\geq$

$$f(\{\text{🍕}, \text{🍷}\}) - f(\{\text{🍷}\})$$

🍷 を持っているときに

🍕 をもらった場合の増分

手元にある要素が増えると新しい要素による増分が減る

$$f(\{\text{🍕}\}) - f(\{\})$$

何も持っていないときに
🍕 をもらった場合の増分

$\geq$

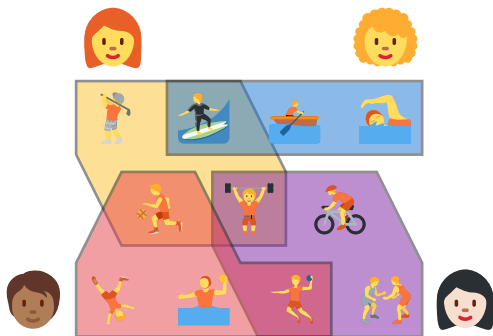
$$f(\{\text{🍕}, \text{🍷}\}) - f(\{\text{🍷}\})$$

🍷 を持っているときに
🍕 をもらった場合の増分

定理 (informal)

劣モジュラ性が成り立っていれば、貪欲法でよい近似解を見つけられる

スポーツ大会のために、たくさんの競技に対応できるチームを作りたい



各候補にいくつか得意競技がある

関数値はチームで対応できる競技の数

$$f(\{\text{👧}, \text{👦}\}) = 7$$

命題

被覆関数は劣モジュラ

→ 貪欲法でよい近似

スポーツ大会のために、たくさんの競技に対応できるチームを作りたい

				
	6	10	4	7
	4	8	9	2
	7	5	9	8
	5	7	6	4

各競技における点数がわかっている
関数値は各競技の最高点の和

命題

施設配置関数は劣モジュラ

→ 貪欲法でよい近似

スポーツ大会のために、たくさんの競技に対応できるチームを作りたい

				
	6	10	4	7
	4	8	9	2
	7	5	9	8
	5	7	6	4
	7	10	9	8

各競技における点数がわかっている
関数値は各競技の最高点の和

$$f(\{\text{👩}, \text{👨}\}) = 7 + 10 + 9 + 8 = 34$$

命題

施設配置関数は劣モジュラ

➡ 貪欲法でよい近似

どの部屋を内見すべきか？ : 部屋の点数は内見するまでわからない

1



確率 $\frac{1}{3}$ で 84 点

確率 $\frac{2}{3}$ で 0 点

2



確率 $\frac{1}{2}$ で 66 点

確率 $\frac{1}{2}$ で 0 点

3



確率 $\frac{3}{5}$ で 50 点

確率 $\frac{2}{5}$ で 0 点

制約

内見できるのは 2 部屋まで

目標

内見したうちで最良の部屋の点数の期待値

細かい注意：

それぞれの点数は独立に決まると仮定

あらかじめどの 2 部屋を内見するか決めておく



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あらかじめどの 2 部屋を内見するか決めておく

● $X = \{\text{🏯}, \text{🏠}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{1}{2} \times 66 = 50$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あらかじめどの 2 部屋を内見するか決めておく

● $X = \{\text{🏯}, \text{🏠}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{1}{2} \times 66 = 50$$

● $X = \{\text{🏯}, \text{🏡}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{3}{5} \times 50 = 48$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あらかじめどの 2 部屋を内見するか決めておく

● $X = \{\text{🏯}, \text{🏠}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{1}{2} \times 66 = 50$$

● $X = \{\text{🏯}, \text{🏡}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{3}{5} \times 50 = 48$$

● $X = \{\text{🏠}, \text{🏡}\}$

$$\frac{1}{2} \times 66 + \frac{1}{2} \times \frac{3}{5} \times 50 = 48$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あらかじめどの 2 部屋を内見するか決めておく

● $X = \{\text{🏯}, \text{🏠}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{1}{2} \times 66 = 50$$

● $X = \{\text{🏯}, \text{🏡}\}$

$$\frac{1}{3} \times 84 + \frac{2}{3} \times \frac{3}{5} \times 50 = 48$$

● $X = \{\text{🏠}, \text{🏡}\}$

$$\frac{1}{2} \times 66 + \frac{1}{2} \times \frac{3}{5} \times 50 = 48$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



確率 $\frac{1}{3}$ で 84

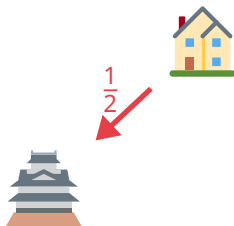


確率 $\frac{1}{2}$ で 66



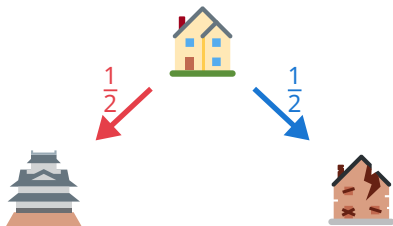
確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



	確率 $\frac{1}{3}$ で 84
	確率 $\frac{1}{2}$ で 66
	確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



確率 $\frac{1}{3}$ で 84

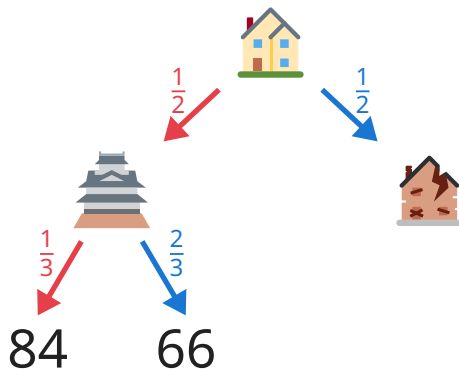


確率 $\frac{1}{2}$ で 66



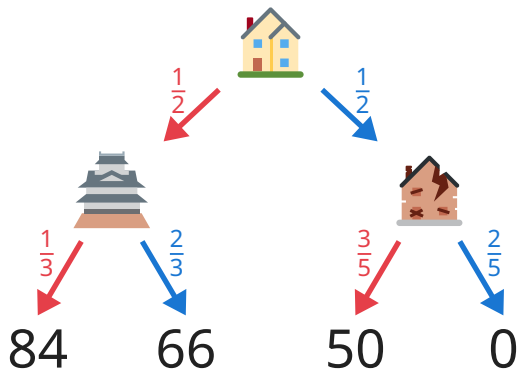
確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



	確率 $\frac{1}{3}$ で 84
	確率 $\frac{1}{2}$ で 66
	確率 $\frac{3}{5}$ で 50

1 部屋目を内見してから 2 部屋目を決める



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

$$\frac{1}{6} \times 84 + \frac{1}{3} \times 66 + \frac{3}{10} \times 50 = 51$$

非適応的な方法より改善

増分の期待値が最大の部屋を順に内見していく

増分の期待値はそれぞれ



$$\frac{1}{3} \times 84 = 28$$



$$\frac{1}{2} \times 66 = 33 \leftarrow \text{最大}$$



$$\frac{3}{5} \times 50 = 30$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

増分の期待値が最大の部屋を順に内見していく



増分の期待値はそれぞれ



$$\frac{1}{3} \times 84 = 28$$



$$\frac{1}{2} \times 66 = 33 \leftarrow \text{最大}$$



$$\frac{3}{5} \times 50 = 30$$



確率 $\frac{1}{3}$ で 84

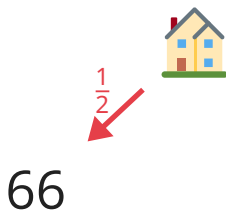


確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

増分の期待値が最大の部屋を順に内見していく



増分の期待値はそれぞれ



$$\frac{1}{3} \times (84 - 66) = 6 \leftarrow \text{最大}$$



$$\frac{3}{5} \times 0 = 0$$



確率 $\frac{1}{3}$ で 84

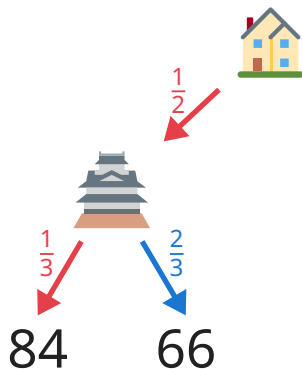


確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

増分の期待値が最大の部屋を順に内見していく



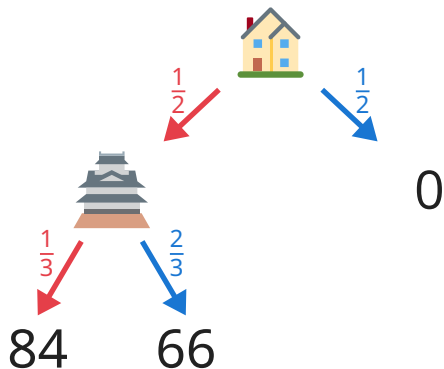
増分の期待値はそれぞれ

 $\frac{1}{3} \times (84 - 66) = 6 \leftarrow \text{最大}$

 $\frac{3}{5} \times 0 = 0$

	確率 $\frac{1}{3}$ で 84
	確率 $\frac{1}{2}$ で 66
	確率 $\frac{3}{5}$ で 50

増分の期待値が最大の部屋を順に内見していく



増分の期待値はそれぞれ



$$\frac{1}{3} \times 84 = 28$$



$$\frac{3}{5} \times 50 = 30 \leftarrow \text{最大}$$



確率 $\frac{1}{3}$ で 84

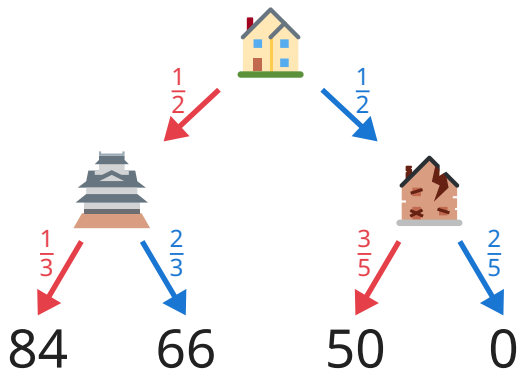


確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

増分の期待値が最大の部屋を順に内見していく



増分の期待値はそれぞれ



$$\frac{1}{3} \times 84 = 28$$



$$\frac{3}{5} \times 50 = 30 \leftarrow \text{最大}$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あとになるほど新しい要素による**増分の期待値**が減る



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あとになるほど新しい要素による**増分の期待値**が減る

$$\Delta(\text{🏰} \mid \{\})$$

最初に🏰を
内見する場合の
増分の期待値

$$\frac{1}{3} \times 84 = 28$$



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

あとになるほど新しい要素による**増分の期待値**が減る



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

$$\Delta(\text{🏯} \mid \{\})$$

最初に🏯を
内見する場合の
増分の期待値

$$\frac{1}{3} \times 84 = 28$$

$$\Delta(\text{🏯} \mid \{\text{🏠}\})$$

🏠を内見したあとに
🏯を内見する場合の
増分の期待値

$$\frac{1}{3} \times (84 - 66) = 6$$

あとになるほど新しい要素による**増分の期待値**が減る



確率 $\frac{1}{3}$ で 84



確率 $\frac{1}{2}$ で 66



確率 $\frac{3}{5}$ で 50

$$\Delta(\text{🏠} | \{\})$$

最初に🏠を
内見する場合の
増分の期待値

$$\frac{1}{3} \times 84 = 28$$

$\geq$

$$\Delta(\text{🏠} | \{\text{🏠}\})$$

🏠を内見したあとに
🏠を内見する場合の
増分の期待値

$$\frac{1}{3} \times (84 - 66) = 6$$

あとになるほど新しい要素による**増分の期待値**が減る

	確率 $\frac{1}{3}$ で 84
	確率 $\frac{1}{2}$ で 66
	確率 $\frac{3}{5}$ で 50

$$\Delta(\text{house} \mid \{\})$$

最初に  を
内見する場合の
増分の期待値

$\geq$

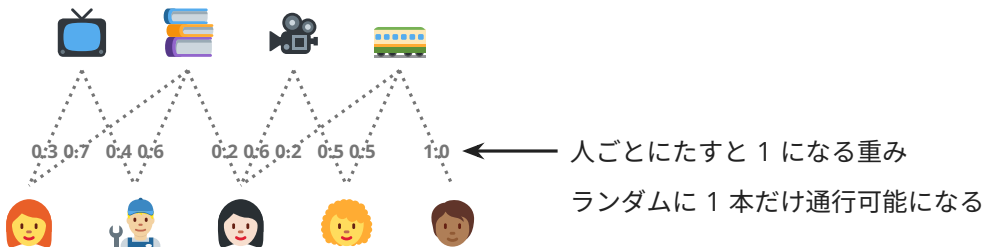
$$\Delta(\text{house} \mid \{\text{house}\})$$

 を内見したあとに
 を内見する場合の
増分の期待値

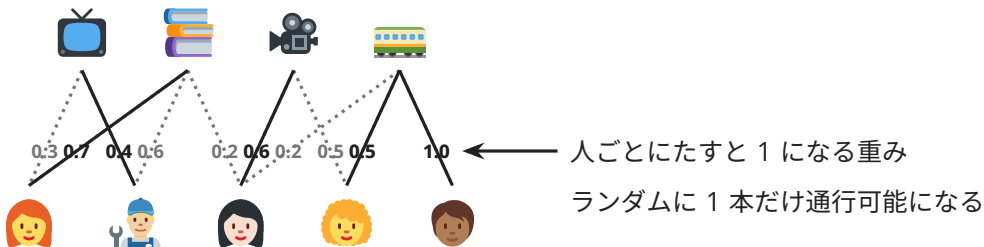
定理 (informal)

適応的劣モジュラ性が成り立っていれば、適応的貪欲法はよい近似戦略

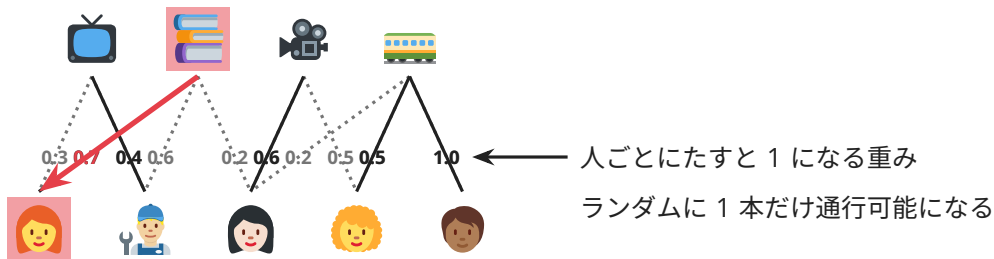
媒体の集合を選んでできるだけたくさんの人に情報を広めたい



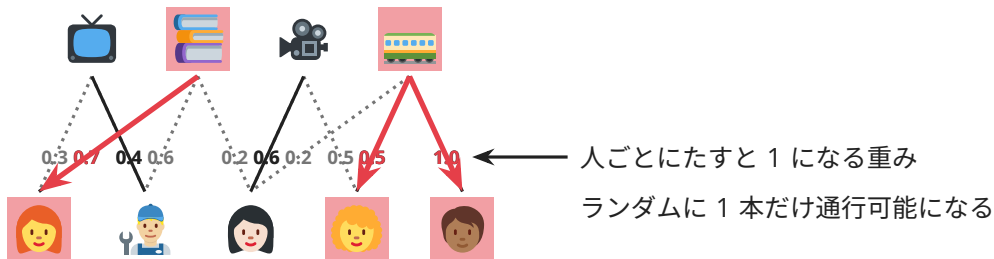
媒体の集合を選んでできるだけたくさんの人に情報を広めたい



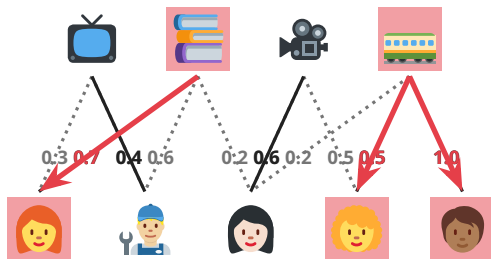
媒体の集合を選んでできるだけたくさんの人に情報を広めたい



媒体の集合を選んでできるだけたくさんの人に情報を広めたい



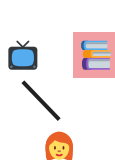
媒体の集合を選んでできるだけたくさんの人に情報を広めたい



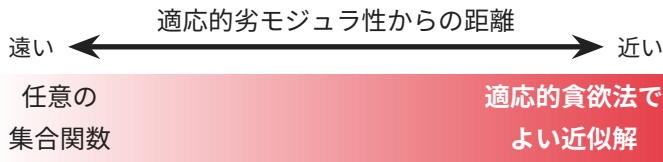
適応的劣モジュラ性は成り立たない



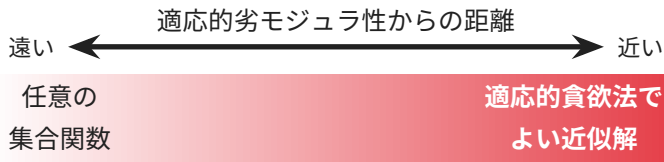
における TV の
増分の期待値：**0.5**



における TV の
増分の期待値：**1**

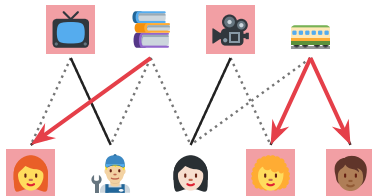


適応的劣モジュラ性に近ければ、適応的貪欲法がよい近似戦略となる



適応的劣モジュラ性に近ければ、適応的貪欲法がよい近似戦略となる

応用 1：適応的影響最大化



応用 2：適応的変数選択

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- 1 相馬輔、藤井海斗、宮内敦史
『組合せ最適化から機械学習へ：劣モジュラ最適化とグラフマイニング』サイエンス社
貪欲法、最大値探索問題を含む、さまざまな劣モジュラ最大化のアルゴリズムと応用例を解説
 - 2 河原吉伸、永野清仁『劣モジュラ最適化と機械学習』講談社
劣モジュラ最小化とその応用に関してはこちらを参照
 - 3 Kaito Fujii, Shinsaku Sakaue.
Beyond adaptive submodularity: Approximation guarantees of greedy policy with adaptive submodularity ratio.
ICML 2019, pp. 2042–2051.
二部グラフにおける適応的影響最大化で貪欲法がよい近似を得ることを証明した論文
- ウェブサイト：<https://fujiik.github.io>