

# 研修医配属における地域間格差を調整する制約のモンテカルロ木探索

板垣 圭知\*  
Keichi Itagaki

小宮山 純平†  
Junpei Komiyama

阿部 拳之‡  
Kenshi Abe

岩崎 敦\*  
Atsushi Iwasaki

## 1 はじめに

本研究では、初期設定として与えた枠のほかに追加で枠を与えることを許される二部マッチングを扱う。二部マッチング問題は研修医マッチングや学校選択制など社会の中で様々な問題に適用することができる汎用性の高い問題である。ここでは例として病院への研修医の配属を例にとり多対1マッチングを考える。このマッチングにおいて各病院は受け入れる研修医の数に対して制約を設け、また病院と研修医はお互いに対して選好順位を持つ。以上の条件の下で効率的なマッチングを求めることを考える。また、二部マッチングにおいては安定性という非常に重要な概念が存在する。安定マッチングとは、どの研修医や病院においてもそのマッチングから逸脱することにメリットはない、というようなマッチングのことである。安定マッチングを見つけることができる有名なアルゴリズムとして Deferred Acceptance (DA) アルゴリズムが広く知られている [1]。

しかし、あらかじめ設定された病院の受入上限がマッチング全体の効用を最大化するものとは限らない。そこで、ある一定の数だけ病院の受入上限を拡張することを考える。これによりマッチングにおけるあらゆる研修医の効用をいくぶんか向上することが可能であることが知られている [2]。

一方で、病院の数が増えるごとに計算量は指数的に増えていくため、各病院に対する拡張枠の割り振りの探索は一般に NP 困難な問題に帰着することが知られている。この問題を解くために UCT 探索 [3] (Upper Confident Tree 探索) という手法を導入する。この手法では拡張枠の割り振りの実行可能空間を木構造によって表現し、UCB と呼ばれる指標が最大になるものについてノードの探索を行っていく。これにより計算量の大幅な削減が可能となる。UCT は十分な時間があれば最適解を得ることができる一方、いつでも探索を停止して合理的によい解を得ることができる。最適解を得るための時間的なコストを考えると、このことは非常に重要である。

従来の研究では各病院の個別の上限のみを考慮して拡張枠を与える探索を行っていた。しかし、実際の研修医マッチングにおいては研修医が特定の地域に集中することを防ぎ、地方や離島に対して一定数の研修医が割り当てられるようにする

ために地域ごとに上限制約を課す。地域上限制約を設けることで安定性を満たすマッチングが存在しない場合があるが、安定性の条件を緩和して弱安定マッチングを定義すると、これを満たすマッチングは必ず存在する [4]。本研究では地域上限制約を課した状況において拡張枠を探索する問題を扱う。各病院に初期値として与える上限設定や拡張可能な枠の上限値の設定が、UCT 探索およびマッチング結果に及ぼす影響について考察する。

## 2 モデル

病院に対して地域上限を設けて研修医のマッチングを行う問題は  $(D, H, R, \succ, q_h, q_r)$  の組で定義される。  $D = \{d_1, d_2, \dots, d_n\}$  は研修医の集合、  $H = \{h_1, h_2, \dots, h_m\}$  は病院の集合、  $R = \{r_1, r_2, \dots, r_l\}$  は地域の集合であり、各地域は  $H$  の互いに素な部分集合として与えられる。なお、  $n = |D|$ ,  $m = |H|$ ,  $l = |R|$  とする。各研修医  $d$  は病院に対して厳密な選好順序  $\succ_d$  を持ち、病院についても同様に学生に対して個別の優先順位  $\succ_h$  を持つ。それぞれの病院の受入上限を  $q_h = \{q_{h_1}, q_{h_2}, \dots, q_{h_m}\}$  と表し、それぞれの地域の受入上限を  $q_r = \{q_{r_1}, q_{r_2}, \dots, q_{r_l}\}$  と表す。

マッチング  $\mu$  について定義する。研修医  $d \in D$  は最大で 1 つの辺を持つことができ、病院  $h \in H$  は最大で  $q_h$  個の辺を持つことができる。  $\mathcal{E}$  をあらゆる  $(d, h)$  のペアの集合としたとき、マッチング  $\mu$  は  $\mathcal{E}$  の部分集合として与えられ、あり得る全ての  $\mu$  の集合を  $\mathcal{M}$  と表す。また、  $\mu(d)$  と  $\mu(h)$  についてそれぞれ、  $d$  が配属された病院、  $h$  に配属された研修医の集合を表すこととする。

**定義 1** (正当化される羨望). マッチング  $\mu$  において  $h \succ_d \mu(d)$  かつ  $d \succ_h d'$  のとき 研修医  $d$  は  $\mu(d') = h$  となるような研修医  $d'$  に対して正当化される羨望 (justifiable envy) をもつという。

どの研修医もほかの研修医に対して正当化される羨望をもたないとき、マッチング  $\mu$  は公平性 (fairness) を満たすという。

**定義 2** (空きシート要求). マッチング  $\mu$  において  $h \succ_d \mu(d)$ ,  $d \succ_h \emptyset$ ,  $|\mu(h)| < q_h$ ,  $|\mu(r) \cup \{d\}| \leq q_r$  であるとき、研修医  $d$  は地域  $r$  に属する病院  $h$  に対して空きシートを要求する (claiming an empty seat) という。

どの研修医も病院に対して空きシートを要求しないとき、マッチング  $\mu$  は非浪費性 (nonwasterfulness) を満たすという。

\* 電気通信大学大学院情報理工学研究所

† ニューヨーク大学スターンスクールオブビジネス

‡ 株式会社サイバーエージェント

**定義 3** (安定マッチング). マッチング  $\mu$  が公平かつ非浪費的であるとき、マッチング  $\mu$  は安定マッチングであるという。

しかし、安定マッチングは常に存在するとは限らない。例えば  $D = \{d_1, d_2\}$ ,  $H = \{h_1, h_2\}$ ,  $R = \{r\}$ ,  $r = \{r_1, r_2\}$ ,  $q_{h_1} = q_{h_2} = q_{r_1} = 1$  で、各研修医および病院の選好順序が以下のようになっているとす。

$$\begin{aligned} \succ_{h_1}: d_1 \succ d_2, \succ_{d_1}: h_2 \succ h_1, \\ \succ_{h_2}: d_2 \succ d_1, \succ_{d_2}: h_1 \succ h_2 \end{aligned}$$

このとき  $|\mu(r)| = 0$  とすると  $\mu$  は非浪費性を満たさない。 $|\mu(r)| = 1$  の場合について考える。 $\mu(h_1) = \{d_2\}$  とすると  $\mu(d_1) = \emptyset$  かつ  $d_1 \succ_{h_1} d_2$  であるから、 $d_1$  は  $d_2$  に対する羨望が正当化されるため、 $\mu$  は公平性を満たさない。 $\mu(h_2) = \{d_1\}$  とした場合についても同様である。 $\mu(h_1) = \{d_1\}$  とすると  $|\mu(h_2)| = 0 < q_{h_2}$  かつ  $h_2 \succ_{d_1} h_1$  であるから  $d_1$  は  $h_2$  に対して空きシートを要求する。 $\mu(h_2) = \{d_2\}$  とした場合についても同様である。よって、この条件の下では安定マッチングは存在しない。

そこで、弱安定マッチングについて考える。

**定義 4** (強空きシート要求). マッチング  $\mu$  において  $h \succ_d \mu(d)$ ,  $|\mu(h)| < q_h$ ,  $|\mu(r)| < q_r$  であるとき、研修医  $d$  は地域  $r$  に属する病院  $h$  の空きシートを強く要求するという。

どの研修医も病院に対して空きシートを強く要求しないとき、マッチング  $\mu$  は弱非浪費的であるという。

**定義 5** (弱安定マッチング). マッチング  $\mu$  が公平かつ弱非浪費的であるとき、マッチング  $\mu$  は弱安定マッチングであるという。

先の例に戻ると、 $\mu(h_1) = \{d_1\}$  としたとき、 $|\mu(r)| = q_r = 1$  となるため、空きシートを強く要求しない。 $\mu(h_2) = \{d_2\}$  とした場合についても同様である。このような  $\mu$  はいずれも弱安定マッチングとなる。

ここで、各病院が予算さえあれば受け入れ可能な研修医の数を  $\bar{q}_h$  とするとき、各病院が一時的な受入上限  $q_h$  と拡張上限  $t_h$  の和は  $\bar{q}_h$  を超えることはない。このため、 $q_h$  が正であるとき、地域上限の  $q_r$  を消費することになる。その病院に  $q_h$  人以上の割当がなくても、余った枠を同じ地域の他の病院に移すことはできないとする。このため、各地域  $r$  における受入上限と拡張上限の和は  $q_r$  を超えることはない。また、全病院における拡張上限の合計は  $B$  とする。これらを用いて、地域上限付きの個別上限拡張問題の実行可能空間を以下に示す。

$$P = \{(x, t) \in \{0, 1\}^{\mathcal{E}} \times \Theta \mid \sum_{h \in H} x_{dh} \leq 1 \forall d \in D, \\ \sum_{d \in D} x_{dh} \leq q_h + t_h \forall h \in H, \sum_{h \in R} q_h + t_h \leq q_r \forall r \in R\},$$

$$\Theta = \left\{ t \in \{0, 1, 2, \dots, B\}^H \mid q_h + t_h \leq \bar{q}_h \forall h, \sum_{h \in H} t_h \leq B \right\}$$

この実行可能空間において、研修医  $d$  が配属された病院  $h$

における  $d$  にとっての選好順位を  $rank_d(h)$  と表すとき、地域上制限約下での個別の受入上限拡張問題は以下のように定義される。

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{t}} \sum_{(d, h) \in \mathcal{E}} rank_d(h) x_{dh} \\ \text{s.t. } (q_h + t_h) \left( 1 - \sum_{h' \in S_{dh}} x_{dh'} \right) \leq \sum_{d' \in T_{dh}} x_{d'h} \forall (d, h) \in \mathcal{E}, \\ (\mathbf{x}, \mathbf{t}) \in \mathcal{P}. \end{aligned}$$

ここで  $S_{dh} = \{h' \in \mathcal{H} : rank_d(h') \leq rank_d(h)\}$  であり、研修医  $d$  にとって病院  $h$  以上の選好順位を持つ病院の集合を表し、同様に  $T_{dh} = \{d' \in \mathcal{D} : rank_h(d') < rank_h(d)\}$  であり、病院  $h$  にとって研修医  $d$  より選好順位が高い研修医の集合を表す。本研究では研修医の効用の最大化を目的関数としている。

### 3 UCT 探索

#### 3.1 UCT 探索の概略

本章では問題を解くためのアルゴリズムである、UCT 探索 [5] (Upper Confident Tree 探索) について解説する。UCT 探索は、モンテカルロ木探索 [6] と呼ばれる手法のうちの 1 つであり、実際にこの UCT 探索を活用することで囲碁のような完全情報ゲームにおいて先の行動を探索し、より有利になるような手を発見することが可能となっている [7] [8] [9]。UCT 探索は他にも様々な推薦手法に活用されている [10]。本研究では UCT を用いて  $v_i$  を評価値として持つノード  $i$  からなる木構造  $\mathcal{T}_{all}$  について探索し、その中から大域的最適となる拡張を表すノードを見つけることを試みる。UCT 探索では木構造を探索するためにいくつかのステップを踏む。 $n \in 1, 2, \dots, N$  をラウンド数とする。木全体  $\mathcal{T}_{all}$  は一般的に規模が大きくなるため、ラウンド  $n$  まで探索を行った部分木として  $\mathcal{T}(n)$  を定義する。current tree 内のそれぞれのノード  $i$  について  $(V_i, N_i) \in \mathbb{R}^2$  が存在し、 $V_i$  は報酬の合計、 $N_i$  はノード  $i$  が探索された回数である。 $\hat{\mu}_i := V_i/N_i$  は評価値  $v_i$  の推定値で、その標準偏差は  $\sqrt{1/N_i}$  に比例する。1 回のラウンドは Selection, Development, Simulation, Backpropagation の 4 つの過程からなる。

**Selection:** UCT は各イテレーションごとに根から木  $\mathcal{T}_s$  を探索する。各ノード  $i$  について子ノード  $c$  を以下に示す UCB が最大になるように選択していく。

$$UCB(c) := \hat{\mu}_c + C_p \sqrt{\frac{\log(N_i)}{N_c}},$$

第 1 項は過去の探索から得た報酬に基づく評価値の予測で、第 2 項は探索回数による評価値である。パラメータ  $C_p$  の値が大きくなるほど過去の探索結果よりシミュレーション回数に影響されるようになり、探索された回数が少ないノードから優先的に探索が行われるようになる。

**Development:** Scurrent tree 外のノード  $k$  にたどり着いたら、current tree にノード  $k$  を追加する。

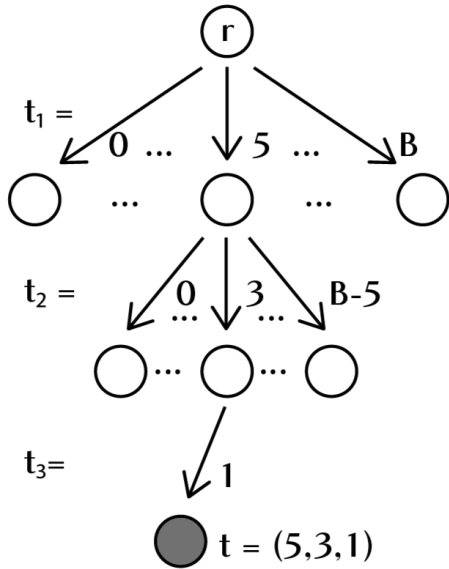


図 1: バッチ木

**Simulation:** ノード  $k$  から葉ノード  $l$  にたどり着くまでランダムな探索を行う。

**Backpropagation:** このステップでは  $v_l$  を葉ノード  $l$  の評価値として、これまでのステップで探索してきた各ノード  $i$  の値を次のように更新する。

$$V_i \leftarrow V_i + v_l, N_i \leftarrow N_i + 1.$$

以上の 4 ステップが一般的な UCT 探索の過程である。

### 3.2 木構造による拡張の表現

この節では、拡張  $t \in \Theta$  と木構造のノードとの関係について解説する。式 (2) によって定義される上限拡張問題において拡張を表現するベクトル  $t$  が求まれば、DA は枠の配分の下で最適なマッチングを求めることができる。そこで、各ノードがベクトル  $t$  に対応ノードを持つ木を構成することを考える。この木構造における葉ノードは、全体での拡張の上限  $\sum_{h \in \mathcal{H}} t_h = B$  までの拡張と対応している。また各ノード  $i$  は、そのノードに対応する  $t$  による拡張を行うことで求めた上限制約において、DA アルゴリズムを実行して得られるマッチングにおける評価値  $v_i$  を持つ。

本研究ではバッチ木 (BT) 図 1 によって拡張を木構造で表現する。この構造において木の高さはマッチングに参加する病院の数と等しくなっており、それぞれの階層において各病院に割り当てる拡張枠をノードで表現している。BT による表現では各病院に対して優先度を計算し、優先度の高い病院から浅い階層に対応させている。このようにすることで人気のある病院から拡張枠を割り当て、最適解の木における深さを減らしている。優先度の計算方法に関しては次節で解説する。

### 3.3 病院の優先度

BT における優先度として、2 つの計算方法を考える。1 つ目は、病院の人気 (popularity) に基づいて優先順位をつける方

法である [11].

$$\text{Popularity}(h) := \sum_{d \in \mathcal{D}} \text{rank}_d(h).$$

この指標は研修医の病院  $h$  への評価を表している。なお、人気が高い病院ほど選好順位の値は小さくなるため、Popularity の値が小さい病院から優先度が高いとみなしている。

2 つ目は、潜在的な羨望 (potential envies) を指標として扱うものである。

$$\text{Envy}(h) := \sum_{d \in \mathcal{D}} \mathbf{1}[\text{rank}_d(h) < \text{rank}_d(\mu(d))]$$

この指標は、マッチング  $\mu$  において割り当てられた病院よりも病院  $h$  の方を好む研修医の数を表しており、マッチング  $\mu$  において病院  $h$  に割り当てられた研修医に対して他の研修医がどれだけ羨望をもつかという指標として解釈される。なお、 $\mu$  は、UCT 探索を行う前の個別上限において DA アルゴリズムを実行することで求める。

## 4 実行可能空間の構造

本節では、地域上限制約下での受入上限拡張問題の構造を理解するため、簡単な具体例を概説する。研修医の数を 128 人に対して、16 個の病院が 4 つの地域に 4 つずつ属しているとする。それぞれの地域上限を 36 とし、各病院の物理的な受入上限  $\bar{q}_h$  を 10 とする。簡単のため、各研修医の病院に対する選好順序は全て等しいと仮定する。

まず、従来アルゴリズムである Flexible DA (FDA) がどのようなマッチングを与えるかを考える。FDA は  $\bar{q}_h = 10, q_h = 0, B = 128$  で表される  $P$  に含まれる空間から拡張上限のベクトル

$$t^{\text{FDA}} = (t_h^{\text{FDA}})_{h \in \mathcal{H}} = (9, 9, 9, 9, \dots, 10, 10, 0, 0)$$

を得る。これは人気のある病院から拡張上限 9 を割り当てていくことで、地域上限  $q_r = 36$  を満たすようにする。3 つの地域に 36 人ずつの枠を割り当てた後、残った予算 20 を使って 4 つ目の地域に属する 2 つの病院にそれぞれ 10 人分の枠を割り当てる。この拡張上限を使って DA を実行すると、得たマッチングが与える研修医の選好順位の合計は 844 になる。一方で、この選好順位の合計を最小化する拡張上限のベクトルは

$$t^* = (t_h^*)_{h \in \mathcal{H}} = (10, 10, 10, 6, \dots, 10, 10, 0, 0)$$

となる。ここでは、最初の 3 つの地域に拡張条件を (10, 10, 10, 6) ずつ割り当てた後、4 つ目の地域に (10, 10, 0, 0) を割り当てる。その結果としての選好順位の合計は 826 になり、FDA が与えるマッチングより改善することがわかる。

これらをベンチマークとして、 $\bar{q}_h$  と  $B$  を変化させたとき何が起きるかを概説する。先に述べた通り、 $q_h = 0, B = 128$  としたとき、全ての実行可能なマッチングを含む  $P$  となっているので、木がうまく展開できれば最適な拡張上限のベクトル  $t^*$  を得る。しかし、現状の実装では人気のある地域や病院に割り当てる拡張上限が一度固定すると、そこまで木をさかのぼっての探索が行われなくなる。実際、もっとも人気のある地域

には, (10, 8, 10, 8) といった拡張上限が割り当てられる。

次に, 探索空間を限定するため,  $q_h$  を 0 から 8 に増やしたケースを考えてみよう。このとき, 各病院に割り当てることができる拡張上限は高々 2 になり, 各地域に割り当てることができる拡張上限は高々 4 にしかならない。よって, 128 人を割り当てするのに必要な予算  $B$  は 16 でよくなる。その結果, 探索空間  $P$  はかなり限定されるので, 最適解の探索が容易になると予想される。しかし,  $q_h = 8$  であるので, 人気のある地域に割り当てられる上限は最終的に (10, 10, 8, 8) にしかならない。その地域でもっとも人気がない病院の枠を 8 から減らすことができないため, 最適な上限を与える (10, 10, 10, 6) を作るができない。

では, もう少し  $q_h$  を減らして 6 とした場合を考えよう。このとき, 各病院に割り当てることができる拡張上限は 4 になり, 各地域に割り当てることができる拡張上限は 12 になる。その結果,  $B = 32$  であれば 128 人全員を割り当てることができる。このとき, もっとも人気の高い地域には, (10, 10, 10, 6) の枠を割り当てることができる。しかし, 3つの地域にこうなるように拡張上限を割り当てた時点で, 予算を使い切ってしまう, 4つ目の地域には (6, 6, 6, 6) しか割り当てることができない。したがって残った研修医 20 人は (6, 6, 6, 2) となるように割り当てられてしまい, 最適な割当である (10, 10, 0, 0) を実現できない。したがって,  $q_h = 6, B = 32$  としたときも最適解である  $t^*$  を含まない。

ここまで見てきたように最適解を必ず含むように  $P$  を設定するには  $q_h$  を 0 にして予算  $B$  を大きく設定しなければならない。しかし, そうすると探索空間  $P$  が大きくなりうまく探索木が展開できなくなる恐れがある。 $q_h$  を非負に設定することで限定した探索空間での最適解は求まりやすくなるが, そもそも探索空間に真の最適解が含まれなくなる恐れがある。このように  $q_h$  は小さく  $B$  は大きく設定するのが最適解を得る上では理想ではあるが, 実行時間が現実的でなかったり, うまく探索木を展開しきれないことがある。こうしたトレードオフがどのように発生するかを次節で検証する。

## 5 評価

本章では, 地域上限付きの個別上限拡張問題のための提案アルゴリズムを実験的に検証する。ベンチマークとして地域上限制約を扱うための DA アルゴリズムの改良である Flexible DA (FDA) [12] を用いる。実験を通して, 研修医の人数  $D$  を 128, 各病院の物理的な受入上限  $\bar{q}_h$  を 10 に固定する。病院の数  $\mathcal{H}$  は 8 もしくは 16 とし, 1つの地域に属する病院の数を 4 に固定した。すなわち, 病院の数を 8 としたときの地域数は 2, 16 としたときの地域数を 4 とする。各地域に課す上限  $q_r$  は,  $\mathcal{H} = 8$  のとき, 18 とし,  $\mathcal{H} = 16$  のとき, 36 とした。4 章で述べたとおり, 個別上限拡張問題の構造は, 各病院に事前に与える枠  $q_h$  と予算  $B$  によって変化する。FDA を用いるとき,  $q_h = 0$  および  $B$  を十分大きくとったときの実行可能空間からマッチングを見つけ出していることになる。つま

り, 各病院の個別上限を  $\bar{q}_h$ , 地域上限  $q_r$  を 18 もしくは 36 とした設定において FDA を実行することで, 弱安定マッチングを求める。 $\mathcal{H} = 8$  のとき, 地域上限  $q_r$  を 18 としているため, 全体で病院に配属できる研修医は高々 36 人であるので, どの病院にも配属されない研修医が多く発生する競争的な設定になっている。一方で,  $\mathcal{H} = 16$  のとき, 地域数 4, 地域上限  $q_r$  を 36 としているため, 合計 144 人を配属できるため, かなり競争が緩和されることになる。

UCT 探索を実行するには, 3.3 節で述べた通り, バッチ木を生成する際に病院への優先度を設定しなければならない。random, envy, popularity の 3つの優先度に対応して, 探索アルゴリズムを BT-R, BT-E, BT-P とする。このとき, ロールアウト回数は  $B \times 100$  回, 探索回数と知識利用のバランスをとるパラメータである  $C_p$  の値は 0.1 とし, 20 回の探索による結果の平均値を用いて評価する。

研修医の選好の偏りを調整するパラメータとして  $\alpha \in [0.0, 1.0]$  を導入する。 $\alpha$  の値が大きいほど研修医の選好は偏り, 1.0 では全員が同じ選好順序を持つようになる。各病院に対する評価値はベクトルで与える。全ての研修医で共通のベクトル  $u_c$  と各研修医で個別に持つベクトル  $u_d$  を  $[0.0, 1.0]$  をそれぞれ一様分布により生成する。各病院の評価値を  $\alpha u_c + (1 - \alpha) u_d$  で与え, この値が大きいほど選好順位が高いとしている [13]。

表 1 に探索の実行結果として, 求めた個別上限が与えるマッチングにおいて, 空きシートを要求する研修医の割合を示す。黄色で網掛けをした数字はその設定でもっともよい結果を与えていることを示す。まず, 優先度の設定が結果に与える影響について述べる。どの実験設定において, BT-R は BT-E と BT-P と比較して空きシートを要求する研修医の数が多い一方で, BT-E と BT-P の間にはあまり差がみられない。研修医の選好が比較的近い ( $\alpha = 0.8$ ) ときには, 同じ割合を示すこともある。また, 病院の数を増やしたとき, BT-R の結果が他と比べて極端に悪くなる。これは, 病院の数が少ないときは, UCT 探索が十分に木を展開することができる程度の実行可能空間であったのに対して, 病院の数が多くなったとき, 十分に木を展開しきれないほどに実行可能空間が大きくなり, もっとも人気のある病院への枠を探索する前に与えられた予算を使い切ってしまう最適解に到達しないことが起こった。以上を踏まえて, 以降では BT-E に着目して評価する。

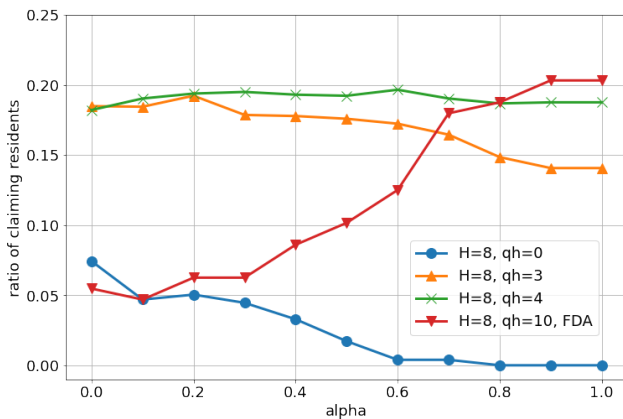
図 2a および 2b に, 研修医の選好の相関  $\alpha$  に対してどのように空きシートを要求する研修医の割合が変化するかを示す。基本的には FDA は選好の相関が小さいほど空きシート要求の割合は小さくなり, 相関が高くなるにつれて, 研修医の希望に合わせた調整がうまく行かず空きシート要求の割合が大きくなる。 $\mathcal{H} = 8$  のとき, 研修医の数に比べて予算が少ないため,  $q_h = 3$  および  $q_h = 4$  では, 人気のある病院の枠を十分に拡張できないため, 相関の大きさに関わらず, 高い空きシート要求が発生する。選好の相関が 0.8 以上でなければ FDA よりよい解に到達しない。 $q_h = 0$  のとき, 実行可能空間は広くなるも

表 1: 空きシートを要求する研修医の割合

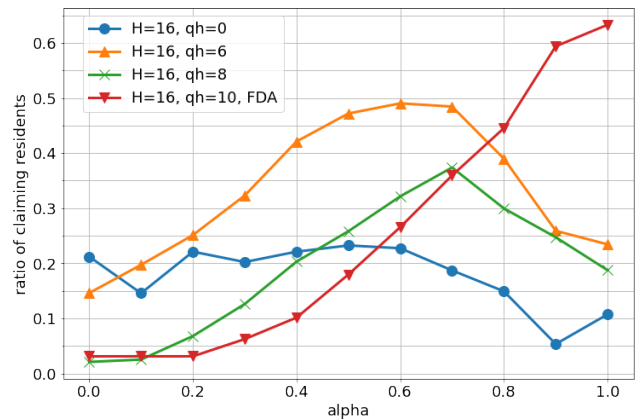
$\mathcal{H}$	$q_h$	B	$\alpha$	FDA	BT-R	BT-E	BT-P
8	0	36	0.0	0.047	0.137	0.074	0.056
8	3	12	0.0	-	0.184	0.185	0.183
8	4	4	0.0	-	0.182	0.182	0.182
8	0	36	0.4	0.047	0.067	0.033	0.041
8	3	12	0.4	-	0.182	0.178	0.175
8	4	4	0.4	-	0.194	0.193	0.193
8	0	36	0.8	0.195	0.096	0.000	0.000
8	3	12	0.8	-	0.155	0.148	0.148
8	4	4	0.8	-	0.187	0.187	0.187
16	0	128	0.0	0.016	0.234	0.212	0.207
16	6	32	0.0	-	0.220	0.146	0.139
16	8	16	0.0	-	0.019	0.021	0.020
16	0	128	0.4	0.133	0.499	0.221	0.280
16	6	32	0.4	-	0.678	0.421	0.435
16	8	16	0.4	-	0.222	0.203	0.177
16	0	128	0.8	0.492	0.468	0.149	0.129
16	6	32	0.8	-	0.662	0.390	0.384
16	8	16	0.8	-	0.388	0.300	0.311

表 2: 実行時間

$\mathcal{H}$	$q_h$	B	$\alpha$	FDA	BT-R	BT-E	BT-P
8	0	36	0.0	0.0	35.0	34.9	34.0
8	3	12	0.0	-	8.4	8.4	8.3
8	4	4	0.0	-	3.4	3.4	3.4
8	0	36	0.4	0.0	33.6	31.0	30.4
8	3	12	0.4	-	8.7	8.7	8.7
8	4	4	0.4	-	3.5	3.4	3.4
8	0	36	0.8	0.0	24.6	21.7	21.6
8	3	12	0.8	-	6.9	6.9	6.8
8	4	4	0.8	-	3.0	2.9	2.9
16	0	128	0.0	0.0	313.4	345.3	317.7
16	6	32	0.0	-	58.7	51.3	56.9
16	8	16	0.0	-	12.8	13.1	12.4
16	0	128	0.4	0.0	222.7	253.1	231.7
16	6	32	0.4	-	64.9	61.4	66.4
16	8	16	0.4	-	28.1	23.6	26.6
16	0	128	0.8	0.0	130.1	146.0	119.0
16	6	32	0.8	-	52.3	45.2	49.2
16	8	16	0.8	-	28.5	24.1	27.9



(a)  $\mathcal{H} = 8$



(b)  $\mathcal{H} = 16$

図 2: 空きシートを要求する研修医の割合

この、十分に探索できる大きさであったため、よい解を見つけることができ、 $\alpha = 0.4$  の時点で FDA を上回り、 $\alpha = 0.8$  以上の場合において必ず最適解を求めることができています。ただし  $\alpha = 0.0$  になると FDA より空きシート要求が多い結果になっている。これは、研修医の選好がバラバラであるためどの病院に枠を割り当てても研修医の効用に大きな変化が現れず、ノードの評価値に差をつけることが難しくなっていることが原因であると考えている。

一方で、 $\mathcal{H} = 16$  のとき、全体的に実行可能空間が大きくなり、最適解にたどりつくことが難しくなるため、選好の相関が十分大きくないと FDA の結果を改善しないことがわかった。

$q_h = 8$  のとき、 $\alpha$  は 0.7 以上でないと FDA を上回らない。しかし、 $\alpha = 0$  のとき、FDA と同等の結果を得るが、これは研修医の選好が分散した結果、あらかじめ配分していた  $q_h$  をうまく埋めることができたためと考えられる。 $q_h = 6$  のとき、 $\alpha$  は 0.8 以上でないと FDA を上回らないが、 $q_h = 8$  のときと比べて空きシート要求が増加している。こちらは、 $q_h = 8$  に比べて  $q_h = 6$  で大きくなった実行可能空間を十分に探索できなくなったことが影響していると考えられる。最後に、 $q_h = 0$  のとき、 $\alpha$  は 0.6 以上で FDA を上回る結果となった。実行可能空間は最初の 2 つと比較して大きくなっているものの、事前に配分した無駄が少なくなる分、良い結果が出たと考えら

れる。また、 $\alpha$ が0.4以下のとき、 $q_h = 8$ に比べて空きシート要求が多くなるのは、ノードの評価値に差をつけることができないことが原因であると考えられる。

表2にはアルゴリズムの実行時間を示す。 $\mathcal{H} = 8$ の実行可能空間では探索量の多い $B = 36$ でも35秒程度であったが、 $\mathcal{H} = 16$ では $\mathcal{H} = 8$ のときと比べて全体的に実行時間が大幅に長くなり、探索量の多い $B = 128$ においては6分近くかかった。現状、FDAとBTでの実行時間しか計測していないが、今後は先行研究の近似アルゴリズム[14]を実装しベンチマークに追加する予定である。

## 6 おわりに

本論文は、地域上限付き制約付きマッチング問題において、できるだけ効率的な弱安定マッチングをモンテカルロ木探索で求める手法を提案、検討した。ここでは、地域上限を満たすように個別上限を調整し、その調整した個別上限のもとでDAアルゴリズムを用いるアプローチをとった。地域上限を直接扱えるよう調整されたFDAをベンチマークに評価実験を行った結果、研修医の選好の相関が十分大きいとき、FDAが与えるマッチングより効率的な弱安定マッチングを求めることに成功した。今後は、階層的な地域上限制約といった複雑な制約を扱う方法や、地域ごとの配属の公平性といった、研修医の効用の最大化とは異なる目的関数を吟味することを考えている。

## 参考文献

- [1] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, Vol. 69, No. 1, pp. 9–15, 1962.
- [2] Alvin E Roth and Marilda Sotomayor. A study in game-theoretic modeling and analysis. *Econometric Society Monographs*, Vol. 18, , 1990.
- [3] Kenshi Abe, Junpei Komiyama, and Atsushi Iwasaki. Anytime capacity expansion in medical residency match by monte carlo tree search. In *IJCAI*, pp. 3–9, 2022.
- [4] Atsushi Iwasaki Yujiro Kawasaki Yosuke Yasuda Goto, Masahiro and Makoto Yokoo. Improving fairness and efficiency in matching with distributional constraints: An alternative solution for the japanese medical residency match. *Munich Personal RePEc Archive, Paper No. 53409.*, 2014.
- [5] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *Machine Learning: ECML 2006: 17th European Conference on Machine Learning Berlin, Germany, September 18-22, 2006 Proceedings 17*, pp. 282–293. Springer, 2006.
- [6] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, Vol. 4, No. 1, pp. 1–43, 2012.
- [7] Yizao Wang and Sylvain Gelly. Modifications of uct and sequence-like simulations for monte-carlo go. In *2007 IEEE Symposium on Computational Intelligence and Games*, pp. 175–182. IEEE, 2007.
- [8] Kazuki Yoshizoe, Akihiro Kishimoto, Tomoyuki Kaneko, Haruhiro Yoshimoto, and Yutaka Ishikawa. Scalable distributed monte-carlo tree search. In *Proceedings of the International Symposium on Combinatorial Search*, Vol. 2, pp. 180–187, 2011.
- [9] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, Vol. 529, No. 7587, pp. 484–489, 2016.
- [10] Guillaume Bosc, Jean-François Boulicaut, Chedy Raïssi, and Mehdi Kaytoue. Anytime discovery of a diverse set of patterns with monte carlo tree search. *Data mining and knowledge discovery*, Vol. 32, pp. 604–650, 2018.
- [11] Hervé Moulin. Social choice. *Handbook of game theory with economic applications*, Vol. 2, pp. 1091–1125, 1994.
- [12] Yuichiro Kamada and Fuhito Kojima. Efficient matching under distributional constraints: Theory and applications. *American Economic Review*, Vol. 105, No. 1, pp. 67–99, 2015.
- [13] Masahiro Goto, Atsushi Iwasaki, Yujiro Kawasaki, Ryoji Kurata, Yosuke Yasuda, and Makoto Yokoo. Strategyproof matching with regional minimum and maximum quotas. *Artificial intelligence*, Vol. 235, pp. 40–57, 2016.
- [14] Federico Bobbio, Margarida Carvalho, Andrea Lodi, and Alfredo Torrico. Capacity expansion in the college admission problem. *arXiv preprint arXiv:2110.00734*, 2021.