Mechanism Design for School Choice with Soft Diversity Constraints

Haris Aziz. Serge Gaspers, Zhaohong Sun
UNSW Sydney, Australia and Data61, CSIRO, Australia
haziz@cse.unsw.edu.au, sergeg@cse.unsw.edu.au, zhaohong.sun@unsw.edu.au,

Abstract

In this paper, we study the controlled school choice problem where students may belong to overlapping types and schools have soft target quotas for each type. We consider how to define fairness properly and investigate which concepts are compatible with non-wastefulness. Then we present a class of algorithms that are fair for students of same types and non-wasteful by taking into account representation of combinations of type. We further prove that our algorithms are strategyproof for students and yield a stable outcome with respect to the induced quotas for type combinations. We experimentally compare our algorithms with two existing approaches in terms of achieving diversity goals and measuring welfare of students.

1 Introduction

Incorporating diversity constraints, fairness and transparency into systems and mechanisms is one of the prominent concerns in artificial intelligence. These concerns are also prevalent in matching markets where there has been increased attention to school choice models that take into account affirmative action and diversity concerns. One particular model of school choice [Abdulkadiroğlu and Sönmez, 2003] with diversity constraints is *controlled school choice*, in which students are associated with a set of types. These types capture traits such as being extra-talented or being from a disadvantaged group. Typically diversity is achieved by setting minimum and maximum target representation of students as has been done in deployed systems. These problems fall under the wider research agenda of matching with distribution constraints which is well studied in AI (see e.g., [Benabbou *et al.*, 2018; Goto *et al.*, 2016; Hamada *et al.*, 2017]).

If diversity constraints are considered as hard bounds, there may not exist an outcome that fulfills all minimum quotas, and a basic tension between fairness and non-wastefulness arises [Ehlers *et al.*, 2014]. Placing hard constraints on diversity constraints may be over-constraining and may put them in head-on conflict with school priorities or other merit consideration. Kojima [2012] show additional evidence that setting hard bounds can be counter-productive. Ehlers *et al.* [2014] remark in their influential paper on controlled school choice that treating quotas as hard bounds is "quite paternalistic in the sense that assignments are enforced independently of student preferences". There are challenges on the computational front as well: it is NP-complete to check whether there exists a feasible or stable matching for the school choice problem [Aziz *et al.*, 2019].

In view of these issues with hard bounds, the recent literature on controlled school choice treats diversity constraints as *soft bounds* which are soft goals that schools attempt to achieve [Hafalir *et al.*, 2013; Ehlers *et al.*, 2014; Kurata *et al.*, 2015, 2017; Gonczarowski *et al.*, 2019]. These quotas are often used to determine which types should be given higher precedence when school seats are scarce.

Most papers in controlled school choice assume that each student can belong to at most one type. In reality, students may satisfy multiple types. For example, a student could be both female and aboriginal. Kurata *et al.* [2015] were the first to investigate the setting where students are allowed to have overlapping types. They proposed an approach in which students and schools are required to reveal preferences and priorities over contracts that specify which particular type is being used for the match. Whereas it may be useful in some circumstances, it can also be problematic. It might invite collusion or bias depending on how preferences and priorities over contracts involving types are generated.

Gonczarowski *et al.* [2019] studied the Israeli "Mechinot" gap-year matching market with diversity goals for overlapping types and their proposed algorithm has been adapted since 2018. However, this algorithm is not strategy-proof for students and may not yield a stable matching. In addition, it does not eliminate justified envy among students who have exactly the same set of types.

In this paper, we study the controlled school choice problem where students may have overlapping types and diversity constraints are viewed as soft bounds. The research question we consider is *how to design mechanisms that cater to diversity objectives while still satisfying desirable fairness, non-wastefulness and strategy-proofness properties?*

Contributions We propose new fairness definitions for school choice with overlapping types that generalizes previous concept used for disjoint types. We show that in general fairness is incompatible with non-wastefulness even if there are only two types. We then present a class of algorithms *Generalized Deferred Acceptance for Combinations of Types*

(GDA-CT) that satisfy weaker version of fairness and non-wastefulness. Unlike the previous approach [Kurata *et al.*, 2015], that modifies the structure of preferences and priorities, we take an alternative route to overcome this incompatibility. The pivotal idea is to elinminate overlapping types by creating a new of set type combinations. Although there may be an exponential number of type combinations, we observe that the number of type combinations whose representation matters is always bounded by the number of students. Finally, we compare our solution with previous work by experimental simulation.

2 Preliminaries

To simplify the presentation, we only consider minimum quotas for the rest of the paper as was the focus of Kurata *et al.* [2015]. Our concepts and algorithms can be extended to cater to upper diversity quotas.

An instance I^T of the school choice problem with diversity constraints consists of a tuple $(S, C, q_C, T, \underline{\eta}, \mathcal{X}, \succ_S, \succ_C)$ where $S = \{s_1, ..., s_n\}$ and $C = \{c_1, ..., c_m\}$ denote the set of students and schools respectively. The capacity vector $q_C = (q_c)_{c\in C}$ gives a capacity q_c for each school c. The type space is denoted by $T = \{t_1, ..., t_k\}$. For each student s, we use $T(s) \subseteq T$ to represent the subset of types to which student s belongs. If $T(s) = \emptyset$, it indicates that student s does not have any privileged type. For each school c, we use $\underline{\eta}_c^t$ to represent the minimum quota for type t. Let $\underline{\eta}_c = (\underline{\eta}_c^t)_{t\in T}$ denote the type-specific minimum quota vector of school c and let $\underline{\eta}$ be a matrix consisting of all schools' type-specific minimum quotas.

Each contract $x = (s, c)$ consists of a student-school pair representing that student s is matched to school c. Let $\mathcal{X} \subseteq S \times C$ denote the set of available contracts. Given any $X \subseteq \mathcal{X}$, let $X_s = \{(s, c) \in X | c \in C\}$ be the set of contracts involving student s, let $X_c = \{(s, c) \in X | s \in S\}$ be the set of contracts involving school c and let $X_c^t = \{(s, c) \in X | s \in S, t \in T(s)\}$ be the set of contracts involving type t and school c.

Each student s has a strict preference ordering $\succ_s$ over $\mathcal{X}_s \cup \{\emptyset\}$ where $\emptyset$ is a null contract representing the option of being unmatched for student s. A contract (s, c) is *acceptable* to student s if $(s, c) \succ_s \emptyset$. Let $\succ_S = \{\succ_{s_1}, ..., \succ_{s_n}\}$ be the preference profile of all students S. Each school c has a strict priority ordering $\succ_c$ over $\mathcal{X}_c \cup \{\emptyset\}$ where $\emptyset$ represents the option of leaving seats vacant for school c. A contract (s, c) is *acceptable* to school c if $(s, c) \succ_c \emptyset$. Let $\succ_C = \{\succ_{c_1}, ..., \succ_{c_m}\}$ be the priority profile of all schools.

An outcome (or a matching) X is a subset of $\mathcal{X}$. An outcome X is *feasible* (under soft bounds) for I^T if i) each student s is matched with at most one school, i.e., $|X_s| \leq 1$, and ii) the number of students matched to each school c does not exceed its capacity, i.e., $|X_c| \leq q_c$.

A feasible outcome X is *individually rational* if each contract $(s, c) \in X$ is acceptable to both student s and school c. A feasible outcome X is *non-wasteful* if there is no student s and school c such that i) $(s, c) \succ_s X_s$ and $(s, c) \succ_c \emptyset$, and ii) $X \cup \{(s, c)\} \setminus X_s$ is feasible.

An *algorithm* takes an instance I^T as input and outputs a set of contracts. An algorithm is *strategy-proof* for students if there exists no student who can misreport his preferences to be matched with a better school.

Next we briefly introduce the generalized deferred acceptance (GDA) algorithm that provides the groundwork for all algorithms considered in this paper, which extends the classical deferred acceptance to the setting of matching with contracts, attributed to Hatfield and Milgrom [2005].

Given a set of acceptable contracts $X \subseteq \mathcal{X}$, let $Ch_s(X)$ denote the choice function of student s which selects her most preferred contract among the set of contracts X_s involving student s. Similarly, the choice function $Ch_c(X)$ of school c selects a subset of acceptable contracts from X_c. Note that the way to specify Ch_c is not unique and different implementations of the GDA algorithm vary on how to define the choice function of schools. Let Ch_S and Ch_C denote the choice function of all students and all schools respectively.

Input: A set of contracts $X \subseteq \mathcal{X}, Ch_S, Ch_C$
Output: An outcome $Y \subseteq X$

```
Y ← X, Z ← ∅, R ← ∅          % R: rejected contracts
while Y ≠ Z do
    Y ← Ch_S(X \ R), Z ← Ch_C(Y), R ← R ∪ (Y \ Z)
return Y
```

Algorithm 1: Generalized Deferred Acceptance (GDA)

3 Fairness

In this section, we discuss how to define fairness properly for school choice with soft diversity constraints and investigate whether there exists a reasonable fairness concept that is compatible with non-wastefulness.

The minimum requirement of fairness is that a feasible outcome should eliminate justified envy among students of same types, which has been considered in the seminal papers on school choice [Abdulkadiroğlu and Sönmez, 2003; Ehlers *et al.*, 2014]. Definition 1 extends this idea to the general case where overlapping types are possible.

Definition 1 (Fairness for same types). *Given an instance I^T and a feasible outcome X, student s has* justified envy *towards another student s' of same types if i) $(s, c) \succ_s \{X_s\}$, $(s', c) \in X$ and ii) $(s, c) \succ_c (s', c)$, $T(s) = T(s')$. An outcome is* fair for same types *if no student has justified envy towards another student of same types.*

The crux is how to measure justified envy among students of different types. Ehlers *et al.* [2014] proposed a natural idea of *dynamic priority* for the model where each student belongs to one type: Each school gives higher precedence to students whose types have not filled the minimum quotas and lower precedence to students whose types have reached the minimum quotas.

Definition 2 (Fairness across types). *Given an instance I^T in which each student belongs to exactly one type and a feasible outcome X, student s has* justified envy *towards another student s' of different type if i) $(s, c) \succ_s \{X_s\}$, $(s', c) \in X$*

and ii) one of the following cases holds, where $t = T(s)$ and $t' = T(s')$:

- *a)* $|X_c^t| < \underline{\eta}_c^t$ *and* $|X_c^{t'}| > \underline{\eta}_c^{t'}$*;*
- *b)* $|X_c^t| < \underline{\eta}_c^t$, $|X_c^{t'}| \leq \underline{\eta}_c^{t'}$ *and* $(s, c) \succ_c (s, c')$*;*
- *c)* $|X_c^t| \geq \underline{\eta}_c^t$, $|X_c^{t'}| > \underline{\eta}_c^{t'}$ *and* $(s, c) \succ_c (s, c')$*;*

An outcome is fair across types if no student has justified envy towards another student of different type.

Proposition 1 (Ehlers *et al.* [2014])**.** *When each student belongs to exactly one type, there always exists an outcome that is fair for same types, fair across types and non-wasteful.*

Before we proceed to the discussion on fairness across types where overlapping types are allowed, we first propose two functions to facilitate the representation of fairness by merging Definition 1 and Definition 2 in an equivalent but concise way.[1]

Given an instance I^T and a feasible outcome X, the function $f(X_c, t)$ specifies the status of type t at school c, depending on the number of contracts involving type t that have already been assigned to school c.

$$f(X_c, t) = \begin{cases} 1 & \text{if } |X_c^t| < \underline{\eta}_c^t \\ 0 & \text{if } |X_c^t| \geq \underline{\eta}_c^t \text{ or } t = \emptyset \end{cases} \tag{1}$$

It returns 1 if type t is undersubscribed, and 0 otherwise. Note that when $f(X_c, t) = 1$, it is still possible to add one more contract involving type t without exceeding the minimum quota $\underline{\eta}_c^t$ of type t.

We use a function $g(X_c, t, t')$ to compare the status of two types t and t' at school c in the outcome X.

$$g(X_c, t, t') = f(X_c, t) - f(X_c, t'). \tag{2}$$

The following Defintion 3 serves as a vivid illustration of how to employ functions f and g to simplify the representation of fairness.

Definition 3 (Fairness for distinct type)**.** *Given an instance I^T in which each student has a distinct type and a feasible outcome X, student s has* justified envy *towards another student s' if i) $(s, c) \succ_s \{X_s\}$, $(s', c) \in X$ and ii) for outcome $X' = X \setminus \{(s', c)\}$, one of the following two cases holds, where $t = T(s)$ and $t' = T(s')$:*

- *a)* $g(X_c', t, t') > 0$*;*
- *b)* $g(X_c', t, t') = 0$ *and* $(s, c) \succ_c (s, c')$*.*

An outcome is fair for distinct type if it admits no justified envy.

Theorem 1. *When each student belongs to exactly one type, Definition 3 is equivalent to the combination of Definition 1 and Definition 2.*

Note that Definition 2 becomes more complicated when considering both minimum and maximum quotas, while we do not need to make any change to Definition 3.

Next we propose a new fairness concept for overlapping types which collapses to Definition 3 when each student belongs to one type.

[1] Ehlers *et al.* [2014] proposed a fairness concept that merges Definition 1 and Definition 2 by enumerating all circumstances.

Definition 4 (Fairness)**.** *Given a feasible outcome X for instance I^T, student s has* justified envy *towards student s' if i) $(s, c) \succ_s X_s$, $(s, c) \succ_c \emptyset$ and $(s', c) \in X$ and ii) for outcome $X' = X \setminus \{(s', c)\}$, one of the two cases holds:*

- *(a) for every two types $t \in T(s) \setminus T(s')$ and $t' \in T(s') \setminus T(s)$, we have that $g(X_c', t, t') \geq 0$; and there exist two types $t \in T(s) \setminus T(s')$ and $t' \in T(s') \setminus T(s)$ such that $g(X_c', t, t') > 0$;*
- *(b) for every two types $t \in T(s) \setminus T(s')$ and $t' \in T(s') \setminus T(s)$ we have that $g(X_c', t, t') = 0$; and $(s, c) \succ_c (s', c)$.*

A feasible outcome X is fair if it admits no justified envy.

Proposition 1 does not hold for the general model when overlapping types are allowed.

Theorem 2. *The set of fair and non-wasteful outcomes could be empty when overlapping types are allowed even if there are only two types.*

In contrast to the impossibility result in Proposition 2, fairness for same types is compatible with non-wastefulness.

Theorem 3. *There always exists an outcome which satisfies fairness for same types and non-wastefulness.*

4 A Class of New Algorithms GDA-CT

In this section, we propose a class of algorithms Generalized Deferred Acceptance for Combinations of Types (GDA-CT) that are fair for same types and non-wasteful. The general idea is to eliminate overlapping types by creating a new set U corresponding to type combinations of T so that each student belongs to exactly one type combination. Then we create new quotas for type combinations U and incorporate the induced quotas for type combinations into the choice function of schools. We employ the GDA algorithm with the new choice function to determine the outcome. All these procedures consist of our new class of algorithm, generalized deferred acceptance for combinations of types (GDA-CT).

Input: $I^T{=}(S, C, q_C, T, \underline{\eta}, \mathcal{X}, \succ_S, \succ_C)$
Output: $X \subseteq \mathcal{X}$

1: Create a set of type combinations U from types T.
2: Convert I^T into an instance $I^U{=}(S, C, q_C, U, \underline{\delta^U}, \mathcal{X}, \succ_S, \succ_C)$ by replacing T with U and by replacing $\underline{\eta}$ with $\underline{\delta}$.
3: Incorporate quotas for type combinations into choice function Ch_c^{CT}.
4: Run GDA with choice function Ch_c^{CT}.

Algorithm 2: GDA-CT

We use the minimum targets for the type combinations to define our choice function Ch_c^U for each school c as described in Algorithm 3. Given a set of contracts X, the choice function Ch_c^U traverses the set of contracts X_c involving school c twice: in the first round, it selects a set of contracts without exceeding the minimum quotas for type combinations and the capacity q_c of school c; in the second round, it selects a set of contracts without exceeding the capacity only. Note that if each student belongs to at most one type in I^T, then I^U is equivalent to I^T. In that case, the choice function defined in

Algorithm 3 is equivalent to the choice function defined by Ehlers *et al.* [2014]. The way to determine quotas for type

Input: An instance I^U, a set of contracts X
Output: A set of contracts $Y \subseteq X$

$Y \leftarrow \emptyset$
for $x = (s, c) \in X$ in descending ordering of $\succ_c$ **do**
  if $|Y_c| < q_c$ **and** $|Y_c^u| < \underline{\delta}_c^u$ with $u = U(s)$ **then**
    $Y \leftarrow Y \cup \{x\}, X \leftarrow X \setminus \{x\}$
while $|Y| < q_c$ **and** $|X_c| > 0$ **do**
  Select $x \in X$ with highest priority based on $\succ_c$
  $Y \leftarrow Y \cup \{x\}, X \leftarrow X \setminus \{x\}$
return Y

Algorithm 3: Choice function Ch_c^{CT} of school c

combinations is not unique. For instance, we can invoke linear programming to divide minimum quotas for types T into minimum quotas for type combinations U with the following linear programming. And we refer to the GDA alorithm as GDA-CT-LP that makes use of linear programming to generate quotas.

$$\min \quad \sum\nolimits_{u \in U} \underline{\delta}_c^u \tag{3}$$

$$\sum\nolimits_{u \in U^t} \underline{\delta}_c^u \geq \underline{\eta}_c^t, \quad \forall c \in C, \forall t \in T \tag{4}$$

$$\underline{\delta}_c^u \geq 0, \quad \forall u \in U \tag{5}$$

$$\underline{\delta}_c^u \times |S^v| = \underline{\delta}_c^v \times |S^u|, \quad \forall c \in C, \forall u, v \in U \tag{6}$$

5 Experiments

In this section, we evaluate our algorithm with existing approaching by experimental simulation. We first describe two existing algorithms for school choice with soft diversity constraints [Kurata *et al.*, 2017; Gonczarowski *et al.*, 2019].

Gonczarowski *et al.* [2019] proposed a choice function Ch_c^{PMA} to handle soft diversity constraints as follows. It traverses the set of contracts X from highest priority to lowest priority twice: in the first round, a contract $x = (s, c)$ is selected if the capacity q_c of school c is not reached and some type related to student s has not met its minimum quota at school c. In the second round, it selects a set of contracts without exceeding the capacity only.

However, this algorithm is not strategy-proof for students and does not yield a stable outcome. It does not eliminate justified envy among students who have the same type combinations, which is a weaker requirement than fairness.

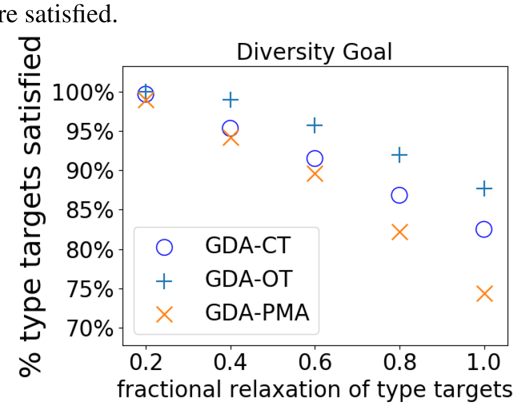
Kurata *et al.* [2015, 2017] proposed another solution by modifing the structure of contracts: students and schools need to specify which particular type is being used in the contract. They assume a student only consumes one unit of some type rather than one unit of all types she belongs to during the process of algorithm. Their choice function of school c works in two rounds: in the first round, for each type t, school c chooses the top students according to the priority ordering until it reaches the minimum quota of type t. In the second round, school c selects the highest priority students without exceeding the capacity.

Whereas it may be useful in some circumstances, it can also be problematic: (1) students may not care about which privilege type they were granted an admission as along as they obtained a school seat; (2) students may be averse to reveal their contract explicitly corresponding to some type; and (3) algorithms based on these approaches are susceptible to collusion or bias depending on how preferences and priorities over contracts involving types are generated.

Setup of Experiments We consider a market consisting of 5000 students and 50 schools with capacity 100 each. The number of types vary in the range $[2, 4, 6, 8]$. We assume that the distributions of types is mutually independent, thus we can calculate the percentage of students of different type combinations by the product of the probability of belonging to each type.

We impose the same minimum vector $\underline{\eta}_c$ to each school c. The minimum target $\underline{\eta}_c^t$ for type t at school c is determined by $\underline{\eta}_c^t = |S^t|/|C| * \alpha$, where $|S^t|/|C|$ is the number of students with type t divided by the number of schools and α is a constant in $[0, 1]$. In this experiment, we choose α to be 0.9, since it is easy to fulfill most of the minimum targets when the constant α is small. We employ Mallows model to generate preferences for students and the priority of each school is equiprobably created.

We measure the performance of algorithms by comparing the percentage of types that satisfy different fractional relaxation of targets. In Figure, the x-axis denotes the fractional relaxation of type targets at schools and the y-axis denotes the percentage of types whose fractional relaxation of targets are satisfied.



For instance, the circle located at $(0.6, 93\%)$ indicates that in the outcome returned by our GDA-CT-LP algorithm, 93% of all types at all schools could satisfy 0.6 fraction of the minimum target.

In summary, the GDA-PMA algorithm outperforms the other two algorithms consistently in terms of achieving diversity goals. However, the number of students who have justified envy towards student of same types is also obvious. The GDA-CT-LP algorithm performs slightly worse than GDA-PMA, but better than GDA-OT. In addition, different ways to break ties among contracts involving types will make a difference to the outcomes yield by GDA-OT: around 5% of total students will be matched to different schools.

References

A. Abdulkadiroğlu and T. Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

H. Aziz, S. Gaspers, Z. Sun, and T. Walsh. From matching with diversity constraints to matching with regional quotas. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 377–385. International Foundation for Autonomous Agents and Multiagent Systems, 2019.

N. Benabbou, M. Chakraborty, X. Ho, J. Sliwinski, and Y. Zick. Diversity constraints in public housing allocation. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 973–981, 2018.

L. Ehlers, I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683, 2014.

Y. A. Gonczarowski, N. Nisan, L. Kovalio, and A. Romm. Matching for the israeli "mechinot" gap year: Handling rich diversity requirements. pages 321–321, 2019.

M. Goto, A. Iwasaki, Y. Kawasaki, R. Kurata, Y. Yasuda, and M. Yokoo. Strategyproof matching with regional minimum and maximum quotas. *Artificial intelligence*, 235:40–57, 2016.

I. E. Hafalir, M. B. Yenmez, and M.A. Yildirim. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363, 2013.

N. Hamada, C. Hsu, R. Kurata, T. Suzuki, S. Ueda, and M. Yokooo. Strategy-proof school choice mechanisms with minimum quotas and initial endowments. *Artificial Intelligence*, 249:47–71, 2017.

J. W. Hatfield and P. R. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005.

F. Kojima. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, 75(2):685–693, 2012.

R. Kurata, N. Hamada, A. Iwasaki, and M. Yokoo. Controlled school choice with soft bounds and overlapping types. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 951–957, 2015.

R. Kurata, N. Hamada, A. Iwasaki, and M. Yokoo. Controlled school choice with soft bounds and overlapping types. *Journal of Artificial Intelligence Research*, 58:153–184, 2017.