Matching Market Design with Constraints

Haris Aziz,[1] **Péter Biró,**[2,3] **Makoto Yokoo**[4]
[1] UNSW Sydney
[2] Centre for Economic and Regional Studies, Hungary
[3] Corvinus University of Budapest
[4] Kyushu University, Japan
haris.aziz@unsw.edu.au, peter.biro@krtk.hu, yokoo@inf.kyushu-u.ac.jp

Abstract

Two-sided matching is an important research area that has had a major impact on the design of real-world matching markets. One consistent feature in many of the real-world applications is that they impose new feasibility constraints that lead to research challenges. We survey developments in the field of two-sided matching with various constraints, including those based on regions, diversity, multi-dimensional capacities, and matroids.

1 Introduction

Matching market design has provided a rich and fertile theory for designing centralized mechanisms for the ubiquitous issue of allocation. The research area has provided a ground for cross-fertilization of ideas, tools, and methods from economics (social choice, mechanism design etc.), computer science (theoretical computer science and artificial intelligence), mathematics (game theory and discrete applied mathematics), and operations research (in particular combinatorial optimization).

In matching markets, the central problem is to match agents to institutions (such as schools or hospitals). In these markets, not only do agents have preferences over institutions, but these institutions also have priorities over the agents (Roth and Sotomayor 1992). The goal is to match the agents to the institutions while keeping into consideration the preference and priorities.

In this survey, we consider matching market design under various constraints. In the seminal paper on this topic (Gale and Shapley 1962), there was already a practical constraint that was considered: the institutions had upper capacities on how many agents they can accommodate. As the theory of matching market design has developed, a consistent stream of results has focused on general two-sided matching models with various types of constraints (see, e.g., Kojima (2019)). Some of their results are aimed at abstraction, unification and generalization of existing methods for two-sided matching. Many results, however, are directly motivated by concrete problems that lead to new feasibility or distributional constraints to a standard two-sided matching problem. For example, there may be an upper bound on a set of urban hospitals to ensure there are enough doctors matched to rural hospitals. Another example is putting a lower bound on the number of underprivileged students at a university as an affirmative action policy. We survey some of the recent results and trends on two-sided matching with constraints. We cover both some of the active work in the economic theory community as well as a growing body of work on the topic within computer science.

2 Basic Model of Gale and Shapley

We first establish the base model of matching under preferences. The two-sided matching model typically has a set of agents N on one side and a set of institutions H on the other. Depending on the application context, the agents are students, doctors, or applicants. Similarly, depending on the application context, the institutions are schools, hospitals, or employers. The basic model has been referred to as the College Admissions model (Gale and Shapley 1962) or the Hospitals/Residents problem (Manlove 2013).

Each institution has an upper capacity on how many agents can be matched to it. Each agent has a preference relation over the institutions. Each institution has a priority relation over the agents. These preference and priority relations are typically ordinal rankings.

An outcome of a two-sided matching problem is an allocation or matching that matches each agent to either some institution or keeps her unmatched. When matching agents to institutions, it is natural to take the preferences and priorities into account. How well this information is taken into account is formalized through axioms that capture the desirability of the matching or the algorithm that finds the matching. Next, we discuss some of the most important axiomatic properties.

A minimal efficiency concept is *non-wastefulness* that requires that no agent i should prefer to be matched to an institution h that still has some vacant space. A widely-used fairness concept is *fairness* (Kamada and Kojima 2022) (also called *justified envy-freeness* (Abdulkadiroğlu and Sönmez 2003)) that requires that there should be no agent i who would prefer to be matched to some institution h and there is an agent matched to h who has lower priority than i in h's priority ranking. If non-wastefulness or fairness are violated, then the agent i and institution h with respect to which the violation occurs are viewed as forming a *blocking pair*. The combination of non-wastefulness and fairness (i.e., the lack of blocking) is also referred to as *stability* in the literature. Stability, however, is a loaded term and there are several variants and generalizations of stability for more complex models and many of them are simply referred to as stability or weak stability.

In two-sided matching, one of the most impactful results is that the Agent Proposing Deferred Acceptance Algorithm satisfies compelling properties. The algorithm returns a matching that is stable and agent-optimal (Gale and Shapley 1962). It is also polynomial-time and strategyproof (an agent cannot get a better outcome by misreporting its preferences). The algorithm works in the following way: agents propose to their most preferred institutions. Based on the proposals, the institutions choose the highest priority individuals up to the capacity. The rejected agents then propose to their next most preferred institution until they run out of options. For a history of the Deferred Acceptance (DA) algorithm, the reader is recommended to read the paper by Roth (2008). In the following section, we will present an extended version of the DA algorithm to accommodate our further generalized setting.

3 Generalized Deferred Acceptance

The general model that we present here has three main extensions. First, we allow different contract terms to be used for an agent-institution pair, so for a contract $x = (i, h, c)$, c denotes the contractual relation, $i \in N$ is an agent, and $h \in H$ is an institution. For instance, in the college admission setting the contractual relation can mean the tuition and other terms, as in Hungary (Biró et al. 2010) and (Biró et al. 2020), or the number of years served, as in the US cadet-branch matching (Sönmez and Switzer 2013). Let $\mathcal{X}$ denote the set of all contracts, and for a set of contracts $X \subseteq \mathcal{X}$ let X_i and X_h denote the subsets of contracts restricted to i and h, respectively. Secondly, the institutions have choice functions Ch_h that for every set of contracts X, h selects a subset Y of X_h. Note that in some matching with contract models, e.g., in (Hatfield and Milgrom 2005), as the primitive of the model it is assumed that every institution h has strict preference over the sets of contracts involving h that determines its choice function. However, this is a simplification that does not even cover some basic settings that is present in some applications.[1] Thirdly, we shall define the collective choice function of the institutions Ch_H, that select $Ch_H(X) \subseteq X$ for every $X \subseteq \mathcal{X}$. Most of the models in the literature assume that Ch_H is simply the union of the individual choice functions Ch_h for all $h \in H$, however, this can be more complex in models with distributional constraints.[2] We will also consider a collective choice function Ch_N of the agents where $Ch_N(X) = \bigcup_{i \in N} Ch_i(X)$ for every $X \subseteq \mathcal{X}$. We summarize the main properties and results for this generalized setting.

The following properties are crucial for choice functions.

- Substitutability (SUB): $X' \subseteq X \subseteq \mathcal{X}$ implies $X' \setminus Ch_H(X') \subseteq X \setminus Ch_H(X)$, so if a contract is rejected from a set, then it is also rejected from an extended set.
- Irrelevance of Rejected Contracts (IRC): $Ch_H(X) \subseteq Y \subseteq X$ implies $Ch_H(X) = Ch_H(Y)$, so if a contract is selected from a set, then it will also be selected when some rejected contracts are removed from the choice set.
- Law of Aggregate Demand (LAD): $Y \subset X$ implies $|Ch_H(Y)| \leq |Ch_H(X)|$, so the number of selected contracts can only increase from an extended set of contracts.

Fleiner (2003) showed that for the general many-to-many matching model with substitutable choice functions on both sides, a so-called *three-stable* matching always exists by Tarski's fixpoint theorem. Three-stability means the existence of three sets X, A, B, subsets of $\mathcal{X}$, such that $X = A \cap B$, $\mathcal{X} = A \cup B$, and $Ch_H(A) = X = Ch_N(B)$.

Furthermore, Fleiner (2003) showed that the three-stable solutions form a lattice if the IRC property holds, where the two variants of the Generalized Deferred Acceptance algorithm produce the agent-optimal and institute-optimal solutions.

Algorithm 1 Generalized Deferred Acceptance (GDA)

Input: Instance I, Ch_N, Ch_H, a set of contracts $\mathcal{X}$
Output: An outcome $Z \subseteq \mathcal{X}$

```
1: Re ← ∅, Y ← 𝒳, Z ← ∅
2: while Y ≠ Z do
3:     Y ← Ch_N(𝒳 \ Re) {Agents select contracts.}
4:     Z ← Ch_H(Y) {Institutions select contracts.}
5:     Re ← Re ∪ (Y \ Z) {Update rejected contracts.}
6: end while
7: return Z
```

Note that when the IRC property holds, three-stability is equivalent to *stability* (Fleiner and Jankó 2014), where stability is defined as follows. An outcome $X \subseteq \mathcal{X}$ is *stable* if (i) $X = Ch_N(X) = Ch_H(X)$ and (ii) there exists no contract $x \in \mathcal{X} \setminus X$ such that $x \in Ch_H(X \cup \{x\})$ and $x \in Ch_N(X \cup \{x\})$ hold. However, without the IRC condition a stable matching may not exist even for substitutable choice functions in the many-to-one case, as Aygün and Sönmez (2013) demonstrated.

Hatfield and Milgrom (2005) considered a restricted many-to-one setting, where they assumed strict preferences for the institutions (in which case the IRC property holds automatically). Furthermore they assumed that the choice

[1] As an example we mention the choice function of the Hungarian universities, where ties are present due to equal scores. When two students are applying for the last seat at a program the decision is to reject them both, however, if only one of these students are applying then she is admitted. In Chile, both of them are admitted even if the quota is violated. See more about these choice functions in (Biró and Kiselgof 2015) and (Fleiner and Jankó 2014).

[2] An example again from the Hungarian university admissions, where common quotas are applied for sets of programs on the top of the individual capacities of the programs, where the programs considered share the same ranking over the students (Biró et al. 2010).

functions are *unitary*, meaning that each hospital selects at most one contract by every agent.[3] Using the same proof technique as Fleiner (2003) they showed that if the SUB and LAD properties hold, then GDA produces an agent-optimal stable matching, and as a new result they also showed that this mechanism is strategyproof for the agents.

The above described general framework with contracts is useful to derive similar results in the context of distributional constraints using the variants of the DA algorithm. However, in most papers in this literature the primitives are simpler, based on the Gale-Shapley model with no contracts. Every agent i has a preference order $\succ_i$ over the institutions and her choice function Ch_i simply selects the best institution from a set offered. Similarly, every institution h has a strict order over the acceptable agents $\succ_h$, together with a quota q_h and from a set of agents Ch_h selects the quota-many best agents.

Example 1 (Illustration of Matching with Contracts)**.** *We illustrate our key concepts and the GDA algorithm on a simple market with three agents $1, 2, 3$ and three insititutions h_1, h_2, h_3 with quota 1 each. Suppose the preferences of agents and the priorities of the institutions are represented in the ordered list in decreasing order of preference/priority.*

$$
\begin{array}{ll}
1 : h_2, h_1, h_3 & h_1 : 1, 3, 2 \\
2 : h_1, h_2, h_3 & h_2 : 2, 1, 3 \\
3 : h_1, h_2, h_3 & h_3 : 2, 3
\end{array}
$$

Note that agent 1 is not specified in the priority list of the h_3 which means that agent 1 cannot be matched to h_3 (for example it does not qualify). The set of valid contracts $\mathcal{X}$ is

$$(1, h_1), (1, h_2), (2, h_1), (2, h_2), (2, h_3), (3, h_1), (3, h_2), (3, h_3).$$

Note that for each pair (i, h), there are no two different possible contracts (i, h, c') and (i, h, c'') so we have ignored the third tuple corresponding to the contractual relation. In this simple example, the choice functions of each agent and institution are directly specified by the preference/priority lists. For example the choice function Ch_{h_3} takes as input a set of agents and selects the highest priority acceptable agent.

The outcome of running GDA is $\{(1, h_1), (2, h_2), (3, h_3)\}$.

When additional feasibility constraints are imposed in two-sided matching, several challenges come up (Kojima 2019). A stable matching or even a feasible matching may not exist and even checking the existence of such a matching can be computationally intractable. In the subsequent sections, we discuss various types of constraints that have been considered in the context of two-sided matching under preferences.

4 Lower Quotas

The challenge of feasibility is immediately seen if we impose *lower* capacities at institutions besides the upper capacities. In college admission or course allocation setting, the lower quota is considered as a minimum number of students needed to open the program. Stability is defined in a standard way for open programs, and for each closed program, the natural non-wastefulness condition is that the number of students preferring this program to their matches cannot reach the lower quota. Biró et al. (2010) conducted a detailed computational complexity analysis of this problem. They showed that a stable matching may not exist and proved that the problem of determining whether a stable matching exists is NP-hard in general. A follow-up paper has been recently written by Boehmer and Heeger (2020).

The other main case is motivated by the resident allocation setting, where the lower quotas have to be obeyed for all the hospitals in order to ensure sufficiently many doctors also at the unpopular (rural) places. This model was first proposed in (Kamada and Kojima 2015), who analysed the Japanese practice of putting artificial upper quotas for popular places to enforce the required lower quotas in the country. Fragiadakis et al. (2016) developed two strategyproof mechanisms: one satisfies fairness and the other one satisfies non-wastefulness (note that no mechanism satisfies fairness and non-wastefulness simultaneously in this setting). More general cases with regional lower quotas have also been studied (Fleiner and Kamiyama 2016; Yokoi 2017; Goto et al. 2016).

5 Regional Constraints

Regional constraints came up in the context of residential allocation, i.e., matching medical residents (doctors) to hospitals. The background was that in order to ensure that sufficient doctors work in rural regions, an upper capacity was placed on urban regions. In general, regional constraints place subsets of institutions such as hospitals into regions and then place upper capacities on the regions. Regional constraints have also been referred to as common quotas (Biró et al. 2010).

There are various levels of regional constraints: (1) disjoint (2) nested (in which for any two regions, either they are disjoint or one is contained in the other), and (3) arbitrary. For the disjoint case, we assume there exists a set of regions R, which is a partition of institutes H. Each region $r \in R$ has its upper quota q_r. In addition to the constraints for the basic model, we require for any $X' \subseteq \mathcal{X}$, if $X = Ch_H(X')$, then $\sum_{h \in r} |X_h| \leq q_r$ holds. Note that with regional constraints, we are no longer able to represent choice function Ch_H as a union of individual choice functions Ch_h for $h \in H$. The next example shows that regional quotas can lead to the non-existence of stable matchings.

Example 2 (A stable matching may not exist under regional constraints)**.** *Suppose the preferences of agents and the priorities of the institutions are represented in the ordered list in decreasing order of preference/priority.*

$$
\begin{array}{ll}
1 : h_1, h_2 & h_1 : 2, 1 \\
2 : h_2, h_1 & h_2 : 1, 2
\end{array}
$$

Suppose both h_1 and h_2 are in a region that has an upper quota of 1. Any feasible matching only matches one agent

[3] In a recent paper, Hatfield and Kominers (2016) showed that indeed this unitary assumption in (Hatfield and Milgrom 2005) was unnecessary.

with one of the instititutions. Consider $\{(1, h_1)\}$. *The matching is not stable as both* h_1 *and* 2 *would prefer to form a contract* $(2, h_2)$. *Consider* $\{(1, h_2)\}$. *The matching is not stable as both* h_1 *and 1 would prefer to form a contract* $(1, h_1)$. *By a symmetric argument no matching that matches agent 2 is stable.*

If we remove the regional quota constraint but add a third institution h_3 *with a lower quota constraint of 1 that all agents prefer the least, a similar situation arises.*

The hospital-resident matching problem with regional constraints has been studied in several papers (Kamada and Kojima 2015; Biró et al. 2010; Goto et al. 2016; Aziz et al. 2019). Kamada and Kojima (2017a) focused on the case of disjoint regions and showed that even for this restriction, a natural notion of stability which they refer to as strong stability is not guaranteed to be satisfied by at least some matching. Aziz, Baychkov, and Biró (2020) showed that even for disjoint regions, checking whether a strongly stable matching exists or not is NP-hard. Kamada and Kojima (2017a) proposed a weak stability concept for the setting. We call a matching X *weakly stable* if, for any blocking pair (i, h) for matching X, the following two conditions are satisfied. (i) $j \succ_h i$ for all applicants s.t. $(j, h) \in X_h$ (ii) $X \cup \{(i, h)\}$ is not feasible.

We can construct Ch_H such that it satisfies the SUB condition as well as required constraints, as discussed in Kojima, Tamura, and Yokoo (2018). For the simplest case, assume each region r has a priority order among all contracts related to the institutions in r. $Ch_H(X)$ chooses at most q_r contracts from $\bigcup_{h \in H} X_h$ as a whole, and at most q_h contracts from X_h, according to the priority. It is easy to verify such a choice function satisfies the SUB, IRC, and LAD conditions. By using this choice function, Algorithm 1 satisfies strategyproofness, and the obtained matching satisfies the stability defined in Section 3 as well as weak stability.

Kamada and Kojima (2017a) also presented an algorithm to compute a weakly stable matching for the problem for any class of region constraints. The algorithm is not strategyproof and in fact, it is an open problem whether weak stability and strategyproofness are compatible.

For a general model of matching with regional constraints in which regions also have priorities over the distributions of agents within the sub-regions within the region, Kamada and Kojima (2018) considered an intermediate notion of stability that they refer to as stability. They show that a matching guaranteeing the notion is guaranteed to exist if and only if the set of region constraints forms a nested structure. Biró et al. (2010) and Goto et al. (2016) also presented positive existence and computational results for nested constraints. Goto et al. (2016) allow both lower and upper quotas on the regions.

6 Diversity Constraints

Diversity concerns are prevalent in many decision making problems including that of two-sided matching. Each agent in a two-sided matching problem may have certain types that indicate they have special attributes (such as being highly talented) or satisfy some affirmative action categories (such as being a historically disadvantaged group). Each institution may have its own diversity constraints and may make selection decisions based on a combination of its priority list over agents as well as information about the types satisfied by the agents.

A standard technique to address diversity constraints is to impose lower and upper quotas on the number of members of given types at the institution. These quotas may be treated as hard constraints or soft constraints (goals). For example, assume agents are divided into disjoint types $T = \{t_1, \ldots, t_k\}$, and institution h imposes a hard type-specific upper quota $q_{h,t}$ for each type $t \in T$ (as well as its overall upper quota q_h). Then, we require for any $X' \subseteq \mathcal{X}$, if $X = Ch_H(X')$, then $|X_{h,t}| \leq q_{h,t}$ holds, where $X_{h,t} = \{(i, h) \mid (i, h) \in X, i \text{ belongs to type } t\}$.

Ehlers et al. (2014a) wrote one of the most influential papers on the topic where they considered the impact of soft and hard diversity constraints. Echenique and Yenmez (2015) also assumed that each student has one type and examined the structure of choice functions that satisfy the SUB condition. Next, we illustrate how one of the choice functions proposed by Ehlers et al. (2014a) works to select a set of agents for an institution by treating the quotas as soft constraints.

Example 3 (Illustration of a choice function with diversity constraints)**.** *Suppose there is one institution* h *with upper quota* $q_h = 3$ *priorities:* $h : 1, 2, 3, 4, 5$. *Suppose there is a minority type* t *such that agents* $3, 4, 5$ *satisfy. Suppose* h *also has type-specific quotas that capture diversity goals: a lower quota of 1 and upper quota of 2 for* t.

Then, Ch_h *can be defined as follows. Firstly, we try to select the highest priority agents of a type that has not reached the lower quota. So agent* 3 *is selected. Then, we try to select the highest priority agents of a type that has not reached the upper quota, so* 4 *is selected. After that, we select the highest priority agents if there is space, so agent 1 is selected. Therefore,* $Ch_h(\{1, 2, 3, 4, 5\}) = \{1, 3, 4\}$. *Note that if diversity constraints are ignored,* h *would use the priority relation to select* $\{1, 2, 3\}$.

Kurata et al. (2017) were the first to consider the setting in which each student has multiple types and each school imposes soft quotas on each type. There are several subsequent works with similar approaches. For example, Aygün and Turhan (2020) and Correa et al. (2019) focus on a model in which student are allowed to have multiple types and colleges divide the seats into groups. Aygün and Bó (2020) study the college admission with multidimensional privileges where each student may have multiple types and each seat within the same school may have different priorities over students. One of the main routes to achieving desiderata such as stability and strategyproofness in this domain is to design suitable choice functions of institutions that satisfy the SUB and LAD conditions. By doing so, the machinery of GDA (Algorithm 1) and the general results of Hatfield and Milgrom (2005) can be applied in these contexts.

Gonczarowski et al. (2019) proposed an algorithm for the Israeli "Mechinot" gap-year problem in which each student has multiple types and schools impose soft lower quotas and hard upper quotas. There has also been some recent work on matchings with constraints on the ratios of types (Nguyen and Vohra 2019).

Matching with diversity has also been examined from a complexity perspective (Biró et al. 2010; Chen, Ganian, and Hamm 2020; Huang 2010; Aziz et al. 2019). Aziz et al. (2019) and Sun (2020) showed that under hard constraints, certain problems for diversity constraints can be reduced to problems under general regional constraints.

In recent years, a new application where matching must take into account representation of various type agents is in healthcare rationing. In this application, the types of agents correspond to important categories such as being a frontline worker, being elderly, or being extra vulnerable (see, e.g., (Aziz and Brandl 2021; Pathak et al. 2020)).

7 Multi-dimensional Capacity Constraints and Agent Sizes

Matching with Sizes

When the agents have sizes, then a stable matching may not exist, and the corresponding decision problem is NP-hard, as McDermid and Manlove (2010) showed for the case of sizes one and two, motivated by the resident allocation problem with couples. In this model, the couples are accepting joint offers only at the hospitals occupying pairs of positions. However, we note that the usual model setting and practice allow also the couples to apply for positions in different hospitals, for a survey see (Biró and Klijn 2013). The stable matching problem with sizes was also studied in the context of *unsplittable flow* problem, motivated by job scheduling to machines (Dean, Goemans, and Immorlica 2006; Cseh and Dean 2016).

Refugee Constraints

Next, we overview multi-dimensional upper capacity constraints inspired by the problem of matching refugee families. Delacrétaz, Kominers, and Teytelboym (2019) formalized refugee allocation as a centralized matching market design problem. A locality can feasibly host a set of families if it can meet the multi-dimensional requirements of the families that could involve services such as hospital beds, children's day care, special medical services, etc. These feasibility requirements can be captured by multi-dimensional feasibility constraints.

Example 4 (multi-dimensional capacity constraints)**.** *Consider three families represented by three agents. Family 1 is an elderly couple that requires 1 room. Family 2 is a family of four that requires 2 rooms and 2 school seats. Family 3 is a family of a single mother and her daughter who require 1 room and 1 school seat. The consumption requirements of the families can be represented by vectors* $(1, 0)$*,* $(2, 2)$*, and* $(1, 1)$ *where the first coordinate represents the required number of rooms and the second coordinate represents the required number of school seats. If a council has 2 available rooms and 2 free school seats, then it cannot accommodate any other family if it accepts Family 2.*

Under multi-dimensional sizes and capacity constraints, the synergies between the agents leads to the choice functions of institutions not satisfying the SUB condition. Delacrétaz, Kominers, and Teytelboym (2019) pointed out that for the refugee allocation problem, the standard stability concept may lead to non-existence of a stable matching.

Aziz et al. (2018) considered a stability concept that is also a natural weakening of stability for two-sided matching and proposed an algorithm to achieve the property. The concept requires that only those agents can have justified envy for another agent if the former require weakly less resources in each category. Aziz et al. also studied the computational complexity of computing matchings satisfying various notions of stability.

Andersson and Ehlers (2016) focused on a restricted version of the refugee allocation problem with unidimensional service demands and capacity vectors but with a feature that captures language compatibility of families and hosts. For this setting, they presented an algorithm that finds a stable maximum matching.

Budgetary Capacity Constraints

Next, we consider related models in which budgetary constraints are involved that are essentially single-dimensional constraints. Ismaili et al. (2019) considered the case where hospitals hire doctors. To hire a doctor, a hospital need to pay a certain wage, which can vary according to the expertise of the doctor. A hospital wants to hire the highest priority doctors within a given budget limit. Assume each contract is represented as (i, h, w), where w is the wage for agent i paid by institution h. We require for any $X' \subseteq \mathcal{X}$, if $X = Ch_H(X')$, then $\sum_{(i,h,w)\in X_h} w$ is at most the budget limit of h. It is easy to see that the SUB condition can be violated; if a hospital prefers hiring two low-wage doctors over hiring one high-wage doctor, two low-wage doctors can be *complementary*. Ismaili et al. (2019) presented several impossibility results for the general case, as well as positive results for some special cases. There also exist works that deal with budget constraints where hospitals have additive utilities (Kawase and Iwasaki 2017, 2018).

Kawase and Iwasaki (2020) considered a model in which institutions have cardinal preferences rather than ordinal. They focus on three types of cardinal preferences: cardinality-based (utility is equal to the size of the match), additive, and submodular. The choice functions of the institutions maximize total utility subject to various kinds of packing constraints such as capacity (upper quota), an intersection of multiple matroids, or a multi-dimensional knapsack.[4] They considered the problems of checking existence of a feasible matching as well as a stable matching. The cardinal preference allows them to also explore approximately stable outcomes.

8 Matroidal Constraints

In the classical stable matching problem, each institution has a capacity constraint. These constraints can be generalized

[4]Note that the constraints involved in refugee matching markets are also based on multi-dimensional knapsacks.

to matroidal constraints. In addition to capacity constraints, matroids also capture more complex constraints including hierarchical or nested capacity constraints that are discussed in the section on regional constraints.

For a finite set of contracts $\mathcal{X}$, let $\mathcal{F}$ denote a family of subsets of $\mathcal{X}$, each of which is feasible according to the constraints. We say a pair $(\mathcal{X}, \mathcal{F})$ is a *matroid* if it satisfies the following conditions. (i) $\emptyset \in \mathcal{F}$ (ii) If $X' \in \mathcal{F}$ and $X'' \subseteq X'$, then $X'' \in \mathcal{F}$. (iii) If $X', X'' \in \mathcal{F}$ and $|X'| > |X''|$, then there exists some $x \in X' \setminus X''$ such that $X'' \cup \{x\} \in \mathcal{F}$. It is easy to verify that for the standard model, where only constraints are the upper quota of each institute, the family of feasible contracts related to each institute constitutes a matroid. Also, for the case of disjoint regions, the family of feasible contracts related to each region constitutes a matroid. Matroidal constraints have been examined in a series of papers (see, e.g., Fleiner (2001), Fleiner and Kamiyama (2016), Kamiyama (2019b), Kamiyama (2019a), (Kamiyama 2020a)). Yokoi (2019) discussed that if the feasibility constraints are madroidal and if there is a total order on the individuals, then the greedy algorithm of selecting agents while not violating feasibility constraints gives rise to a choice function that satisfies the SUB condition.

Note the matroidal feasibility constraints require that the empty set must be feasible. Hence, feasibility constraints that involve lower quotas cannot be directly captured by matroidal constraints. However, both Fleiner and Kamiyama (2016) and Yokoi (2017) consider a generalized approach based on matroids that captures stability and fairness, respectively, for lower quotas with matroid feasibility constraints.

Kojima, Tamura, and Yokoo (2018) showed that (i) Algorithm 1 is strategyproof for agents, (ii) the resulting matching is stable (according to the definition in Section 3) and optimal for each agent among all stable matchings, and (iii) the time complexity of the algorithm is proportional to the square of the number of possible contracts, assuming Ch_H is defined based on matroidal constraints (as well as institutions' aggregated preferences have some simple structures, e.g., maximizing the sum of values associated with individual contracts).[5]

Kojima, Tamura, and Yokoo (2018) showed that a wide variety of constraints, including nested regions (Biró et al. 2010; Goto et al. 2016), diversity requirements in school choice with soft constraints (Ehlers et al. 2014b), the student-project allocation problem (Abraham, Irving, and Manlove 2007), and the cadet-branch matching problem (Sönmez and Switzer 2013), can be represented as matroidal constraints and existing mechanisms/algorithms that had been developed independently/separately can be represented as a unified mechanism/algorithm described as Algorithm 1.

[5]More precisely, Kojima, Tamura, and Yokoo (2018) assume that Ch_H is represented as a maximizer of a given function f that satisfies the condition called M$^\natural$-concavity. The fact that f is an M$^\natural$-concave function implies that hard constraints are represented as matroidal constraints.

9 Heredity Constraints

In this section, we consider feasibility constraints that satisfy the heredity property where the feasibility of a matching is monotone in the number of agents matched at the institutions. The heredity property requires that if a matching is feasible, then any matching in which the numbers of agents matched to each institution weakly decreases, is feasible as well. Heredity constraints are more general than imposing upper bounds on regions/subsets of institutions. The reason is that if the number of agents at institutions weakly decreases, the constraints imposed by imposing upper quotas on regions are still satisfied. However, heredity constraints do not capture certain types of diversity constraints: decreasing the matches of institutions may make the representation requirement of some type of agents to be lower than the required level. Heredity constraints have been considered in some recent papers (Kamada and Kojima 2017b; Goto et al. 2017; Aziz, Baychkov, and Biró 2020; Kamada and Kojima 2022).

Adaptive Deferred Acceptance

Goto et al. (2017) proposed a general algorithm called *Adaptive Deferred Acceptance (ADA)* that can be applied to any heredity constraints. This algorithm satisfies strategyproofness, non-wastefulness, and a fairness property (that is weaker than the weak stability concept of Kamada and Kojima (2017b)). Next, we provide some high level ideas behind ADA. ADA utilizes a serial order among agents called *master-list*. ADA works by repeatedly calling the standard Deferred Acceptance algorithm by adding agents one by one according to the master-list. In ADA, an institution becomes forbidden if it cannot accommodate any additional agent without some feasibility constraint being violated. In each stage, there are multiple rounds, each adding students according to the master-list as long as no institution becomes forbidden. When an institution becomes forbidden, the algorithm finalizes the current matching and moves to the next stage where the upper quotas and the distributional constraints are updated by taking into account all the contracts that have been finalized so far. The formal description of ADA is given as Algorithm 2.

Cutoff Stable Algorithm

Aziz, Baychkov, and Biró (2021) showed that as long as the feasibility constraints satisfy the heredity property, cutoff stability (a natural notion of stability that is stronger than weak stability of Kamada and Kojima (2017b)) is satisfied. The idea behind cutoff stability is as follows. Cutoff stability requires that fairness as well as what is called cutoff non-wastefulness are satisfied. Cutoff non-wastefulness requires that either an agent cannot leave her match and join a more preferred institution h without violating some feasibility constraint or if it can, then there exists another agent that is at higher priority at h that cannot leave her match and join h without violating some feasibility constraint. Aziz, Baychkov, and Biró (2021) presented a simple but widely applicable algorithm that works for any matching problem with feasibility constraints. If the constraints satisfy the heredity property, the algorithm returns a cutoff stable outcome.

Algorithm 2 Adaptive Deferred Acceptance (ADA)

Input: master-list $L = (i_1, i_2, \ldots)$ and upper quotas q_H
Output: matching Y

Initialization: $q_h^1 \leftarrow q_h$ for each $h \in H$, $Y \leftarrow \emptyset$. Proceed to **Stage** 1.

Stage k**:** Proceed to **Round** 1.

Round t**:** Select t agents from the top of L. Let Y' denote the matching obtained by the standard deferred acceptance for the selected agents under $(q_h^k)_{h \in H}$.

(i) If all agents in L are already selected, then $Y \leftarrow Y \cup Y'$, output Y and terminate the mechanism.

(ii) If no new institution is forbidden, then proceed to **Round** $t+1$.

(iii) Otherwise, $Y \leftarrow Y \cup Y'$. Remove t agents from the top of L. For each institution h that is forbidden, set q_h^{k+1} to 0. For each $h \in H$, which is not forbidden, set q_h^{k+1} to $q_h^k - |Y_h'|$. Proceed to **Stage** $k+1$.

Let $d : H \rightarrow [0, 1, \ldots, |N|+1]$ be the cutoff score function, where $d(h)$ is the *cutoff* at institution h. Without loss of generality we assume that each institution h assigns a score to each agent i in accordance with its preference list, that is i has score $|N| - k + 1$ if she is ranked k-th by institution h. Given cutoff scores d, we say that agent i is *admissible* to institution h if her score achieves the cutoff. Cutoff scores d *induce* matching M, if every agent is matched to the best institution of her preference where she is admissible.

Let d^{-h} denote the cutoff scores after decreasing the cutoff of h by one, and keeping the other cutoffs the same, i.e., $d^{-h}(h) = d(h) - 1$, and $d^{-h}(h') = d(h')$ for every $h' \neq h$. We say that cutoffs d are *minimal* if we cannot decrease the cutoff score of any institution without making the induced matching infeasible. More formally for every institution h, either $d(h) = 0$ or the matching induced by d^{-h}, which we call M^{-d}, is infeasible. The formal definition of the algorithm presented by Aziz, Baychkov, and Biró (2021) is given as Algorithm 3.

Algorithm 3 Cutoff algorithm for heredity constraints.

Input: lists $\succ_h$ for all $h \in H$ and $\succ_i$ for all $i \in N$; feasibility function f; institution order $P^* = (h_1, ..., h_k)$
Output: Matching M and corresponding cutoffs d_M

1: Initialize M to empty and $d_M(h) = |N| + 1$ for every institution h.
2: **while** Cutoff d_M are not minimal **do**
3: Locate the first h_j in the list P^* such that M^{-h_j} is feasible.
4: Let $M = M^{-h_j}$ and $d_M = d_M^{-h_j}$.
5: **end while**

Intra-Institution Heredity Constraints

Kamada and Kojima (2022) consider a more restricted version of heredity constraints that are *intra-institution* and show that an outcome that satisfies the following properties exists: fairness, feasibility, agent-optimal among all matchings that satisfy fairness and feasibility.

10 Conclusions

We surveyed recent work on two-sided matching with constraints. We focused on the work that takes stability and fairness as central concerns. In recent years, work on other objectives such as popularity has also been extended to more complex feasibility constraints (see, e.g., Kamiyama (2020b)). We restricted our presentation to many-to-one matching problems in which each agent is matched to at most one institution. There are many results in the literature that pertain to many-to-many matchings, a model motivated by several relevant applications such as the resident allocation problem with couples (Biró and Klijn 2013), the assignment of papers to reviewers (Garg et al. 2010), and course allocation (Budish et al. 2017). Another stream of work is on exchange problems (Abdulkadiroglu and Sönmez 1999; Guillen and Kesten 2012; Suzuki, Tamura, and Yokoo 2018) in which individual rationality and Pareto optimality are major concerns.

Acknowledgments

The authors thank Jeremy Vollen and the anonymous reviewers of AAAI 2022 for helpful comments. Aziz is supported by DSTG under the project "Auctioning for distributed multi vehicle planning" (MYIP 9953) and by AFOSR under the project "Efficient and fair decentralized task allocation algorithms for autonomous vehicles" (FA2386-20-1-4063). Biró gratefully acknowledges financial support from the Hungarian Scientific Research Fund, OTKA, Grant No. K128611, and the Hungarian Academy of Sciences, Momentum Grant No. LP2021-2. Yokoo is partially supported by JSPS KAKENHI Grant Numbers JP20H00609 and JP21H04979.

References

Abdulkadiroglu, A.; and Sönmez, T. 1999. House Allocation with Existing Tenants. *Journal of Economic Theory*, 88: 233–260.

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

Abraham, D. J.; Irving, R. W.; and Manlove, D. F. 2007. Two Algorithms for the Student-Project Allocation Problem. *Journal of Discrete Algorithms*, 5(1): 73–90.

Andersson, T.; and Ehlers, L. 2016. Assigning refugees to landlords in Sweden: Stable maximum matchings. Technical Report 2016-08, Université de Montréal Papyrus Institutional Repository.

Aygün, O.; and Bó, I. 2020. College admission with multidimensional privileges: The Brazilian affirmative action case. *American Economic Journal: Microeconomics*.

Aygün, O.; and Sönmez, T. 2013. Matching with contracts: Comment. *American Economic Review*, 103(5): 2050–51.

Aygün, O.; and Turhan, B. 2020. Dynamic reserves in matching markets: Theory and applications. *Journal of Economic Theory*, 188.

Aziz, H.; Baychkov, A.; and Biró, P. 2020. Summer Internship Matching with Funding Constraints. In *Proc. of 19th AAMAS Conference*.

Aziz, H.; Baychkov, A.; and Biró, P. 2021. Cutoff stability under distributional constraints with an application to summer internship matching. Technical report.

Aziz, H.; and Brandl, F. 2021. Efficient, Fair, and Incentive-Compatible Healthcare Rationing. In *EC '21: The 22nd ACM Conference on Economics and Computation, Budapest, Hungary, July 18-23, 2021*, 103–104. ACM.

Aziz, H.; Chen, J.; Gaspers, S.; and Sun, Z. 2018. Stability and Pareto Optimality in Refugee Allocation Matchings. In *Proc. of 17th AAMAS Conference*, 964–972.

Aziz, H.; Gaspers, S.; Sun, Z.; and Walsh, T. 2019. From Matching with Diversity Constraints to Matching with Regional Quotas. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, 377–385. International Foundation for Autonomous Agents and Multiagent Systems.

Biró, P.; Fleiner, T.; Irving, R. W.; and Manlove, D. F. 2010. The College Admissions problem with lower and common quotas. *Theoretical Computer Science*, 411(34): 3136 – 3153.

Biró, P.; Hassidim, A.; Romm, A.; Shorrer, R. I.; and Sóvágó, S. 2020. Need versus Merit: The Large Core of College Admissions Markets. *arXiv preprint arXiv:2010.08631*.

Biró, P.; and Kiselgof, S. 2015. College admissions with stable score-limits. *Central European Journal of Operations Research*, 23(4): 727–741.

Biró, P.; and Klijn, F. 2013. Matching with couples: a multidisciplinary survey. *International Game Theory Review*, 15(02): 1340008.

Boehmer, N.; and Heeger, K. 2020. A fine-grained view on stable many-to-one matching problems with lower and upper quotas. In *International Conference on Web and Internet Economics*, 31–44. Springer.

Budish, E.; Cachon, G. P.; Kessler, J. B.; and Othman, A. 2017. Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65(2): 314–336.

Chen, J.; Ganian, R.; and Hamm, T. 2020. Stable Matchings with Diversity Constraints: Affirmative Action is beyond NP. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, 146–152.

Correa, J. R.; Epstein, R.; Escobar, J.; Rios, I.; Bahamondes, B.; Bonet, C.; Epstein, N.; Aramayo, N.; Castillo, M.; Cristi, A.; and Epstein, B. 2019. School Choice in Chile. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019*, 325–343.

Cseh, Á.; and Dean, B. C. 2016. Improved algorithmic results for unsplittable stable allocation problems. *Journal of Combinatorial Optimization*, 32(3): 657–671.

Dean, B. C.; Goemans, M. X.; and Immorlica, N. 2006. The unsplittable stable marriage problem. In *Fourth IFIP International Conference on Theoretical Computer Science-TCS 2006*, 65–75. Springer.

Delacrétaz, D.; Kominers, S. D.; and Teytelboym, A. 2019. Matching Mechanisms for Refugee Resettlement. Working Papers 2019-078, Human Capital and Economic Opportunity Working Group.

Echenique, F.; and Yenmez, M. B. 2015. How to Control Controlled School Choice. *American Economic Review*, 105(8): 2679–94.

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

Ehlers, L.; Hafalir, I. E.; Yenmez, M. B.; and Yildirim, M. A. 2014b. School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds. *Journal of Economic Theory*, 153: 648–683.

Fleiner, T. 2001. A Matroid Generalization of the Stable Matching Polytope. In Gerards, B.; and Aardal, K., eds., *Integer Programming and Combinatorial Optimization: 8th International IPCO Conference, LNCS 2081*, 105–114. Springer-Verlag.

Fleiner, T. 2003. A Fixed-Point Approach to Stable Matchings and Some Applications. *Mathematics of Operations Research*, 28: 103–126.

Fleiner, T.; and Jankó, Z. 2014. Choice function-based two-sided markets: stability, lattice property, path independence and algorithms. *Algorithms*, 7(1): 32–59.

Fleiner, T.; and Kamiyama, N. 2016. A Matroid Approach to Stable Matchings with Lower Quotas. *Mathematics of Operations Research*, 41(2): 734–744.

Fragiadakis, D.; Iwasaki, A.; Troyan, P.; Ueda, S.; and Yokoo, M. 2016. Strategyproof Matching with Minimum Quotas. *ACM Transactions on Economics and Computation*, 4(1): 6:1–6:40.

Gale, D.; and Shapley, L. S. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1): 9–15.

Garg, N.; Kavitha, T.; Kumar, A.; Mehlhorn, K.; and Mestre, J. 2010. Assigning papers to referees. *Algorithmica*, 58(1): 119–136.

Gonczarowski, Y. A.; Nisan, N.; Kovalio, L.; and Romm, A. 2019. Matching for the Israeli "Mechinot" Gap Year: Handling Rich Diversity Requirements. In *Proceedings of the 20th ACM Conference on Economics and Computation*, 321–321.

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

Goto, M.; Kojima, F.; Kurata, R.; Tamura, A.; and Yokoo, M. 2017. Designing matching mechanisms under general distributional constraints. *American Economic Journal: Microeconomics*, 9(2): 226–62.

Guillen, P.; and Kesten, O. 2012. Matching markets with mixed ownership: the case for a real-life assignment mechanism. *International Economic Review*, 53: 1027–1046.

Hatfield, J. W.; and Kominers, S. D. 2016. Hidden substitutes. In *Proceedings of the 16th ACM Conference on Electronic Commerce*.

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

Huang, C.-C. 2010. Circular Stable Matching and 3-way Kidney Transplant. *Algorithmica*, 58: 137–150.

Ismaili, A.; Hamada, N.; Zhang, Y.; Suzuki, T.; and Yokoo, M. 2019. Weighted Matching Markets with Budget Constraints. *Journal of Artificial Intelligence Research*, 65: 393–421.

Kamada, Y.; and Kojima, F. 2015. Efficient matching under distributional constraints: Theory and applications. *The American Economic Review*, 105(1): 67–99.

Kamada, Y.; and Kojima, F. 2017a. Recent Developments in Matching with Constraints. *The American Economic Review*, 107(5): 200–204.

Kamada, Y.; and Kojima, F. 2017b. Stability concepts in matching under distributional constraints. *Journal of Economic Theory*, 168: 107–142.

Kamada, Y.; and Kojima, F. 2018. Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition. *Theoretical Economics*, 13(2): 761–793.

Kamada, Y.; and Kojima, F. 2022. Fair Matching Under Constraints: Theory and Applications. *Review of Economic Studies*.

Kamiyama, N. 2019a. Many-to-Many Stable Matchings with Ties, Master Preference Lists, and Matroid Constraints. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019*, 583–591.

Kamiyama, N. 2019b. Pareto Stable Matchings under One-Sided Matroid Constraints. *SIAM Journal of Discrete Mathematics*, 33(3): 1431–1451.

Kamiyama, N. 2020a. On Stable Matchings with Pairwise Preferences and Matroid Constraints. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '20, Auckland, New Zealand, May 9-13, 2020*, 584–592.

Kamiyama, N. 2020b. Popular matchings with two-sided preference lists and matroid constraints. *Theoretical Computer Science*, 809: 265–276.

Kawase, Y.; and Iwasaki, A. 2017. Near-Feasible Stable Matchings with Budget Constraints. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 242–248.

Kawase, Y.; and Iwasaki, A. 2018. Approximately Stable Matchings With Budget Constraints. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*, 1113–1120.

Kawase, Y.; and Iwasaki, A. 2020. Approximately Stable Matchings with General Constraints. In *Proc. of 19th AAMAS Conference*, 602–610.

Kojima, F. 2019. New Directions of Study in Matching with Constraints. In Laslier, J.-F.; Moulin, H.; Sanver, R.; and Zwicker, W. S., eds., *The Future of Economic Design*. Springer.

Kojima, F.; Tamura, A.; and Yokoo, M. 2018. Designing matching mechanisms under constraints: An approach from discrete convex analysis. *Journal of Economic Theory*, 176: 803–833.

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

Manlove, D. F. 2013. *Algorithmics of Matching Under Preferences*. World Scientific Publishing Company.

McDermid, E. J.; and Manlove, D. F. 2010. Keeping partners together: algorithmic results for the hospitals/residents problem with couples. *Journal of Combinatorial Optimization*, 19(3): 279–303.

Nguyen, T.; and Vohra, R. 2019. Stable Matching with Proportionality Constraints. *Operations Research*, 67(6): 1503–1519.

Pathak, P. A.; Sönmez, T.; Ünver, M. U.; and Yenmez, M. B. 2020. Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing. Boston College Working Papers in Economics 1015, Boston College Department of Economics.

Roth, A. E. 2008. Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory*, 36: 537––569.

Roth, A. E.; and Sotomayor, M. A. O. 1992. Two-Sided Matching. In Aumann, R. J.; and Hart, S., eds., *Handbook of Game Theory with Economic Applications*, volume I. Elsevier.

Sönmez, T.; and Switzer, T. B. 2013. Matching with (Branch-of-Choice) Contracts at the United States Military Academy. *Econometrica*, 81(2): 451–488.

Sun, Z. 2020. *Mechanism Design for Matching with Constraints*. Ph.D. thesis, University of New South Wales, Sydney, Australia.

Suzuki, T.; Tamura, A.; and Yokoo, M. 2018. Efficient Allocation Mechanism with Endowments and Distributional Constraints. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS-2018)*, 50–58.

Yokoi, Y. 2017. A Generalized Polymatroid Approach to Stable Matchings with Lower Quotas. *Mathematics of Operations Research*, 42(1): 238–255.

Yokoi, Y. 2019. Matroidal Choice Functions. *SIAM Journal of Discrete Mathematics*, 33(3): 1712–1724.