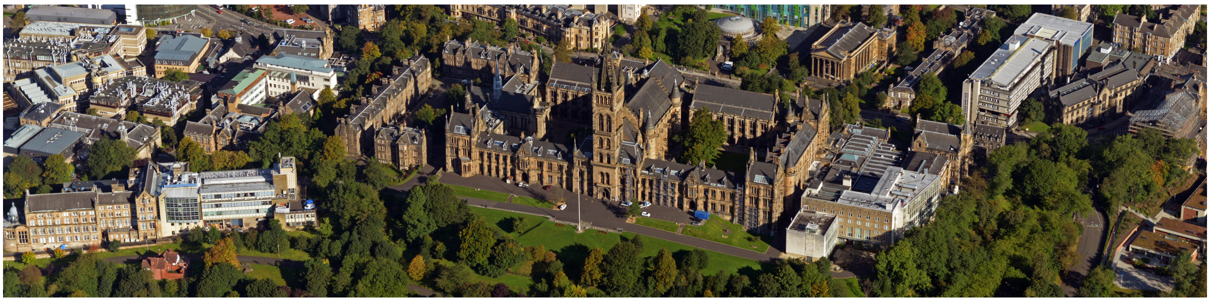
MATCH-UP 2015:

The 3rd International Workshop On Matching Under Preferences



16-18 April 2015
University of Glasgow, Scotland

http://www.optimalmatching.com/MATCHUP2015/

University of Glasgow

COST
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

sicsa* The Scottish Informatics & Computer Science Alliance

COST is supported by the EU Framework Programme Horizon 2020

MATCH-UP 2015:
The 3rd International Workshop on Matching Under Preferences

Preface

This is the third instalment in the series of interdisciplinary MATCH-UP workshops, with the first taking place in Reykjavík in 2008 (as a satellite workshop of ICALP 2008) and the second being held in Budapest in 2012 (co-located with the SING conference).

Five months after the 2012 workshop, the Nobel Prize in Economic Sciences was announced, with Al Roth and Lloyd Shapley receiving the award jointly "for the theory of stable allocations and the practice of market design". This was a tremendously exciting development for the research area and has no doubt contributed to an upsurge in interest in matching under preferences.

Since that time, computer scientists, economists and others have contributed to the growing body of literature in this domain, and it became clear that we should plan for another MATCH-UP. Key to the decision as to "when" and "where" was the idea of Ulle Endriss to co-locate the workshop with a meeting of COST Action IC1205 on Computational Social Choice, which has a working group on Matching.

The COST IC1205 meeting on Fair Division and Matching directly precedes MATCH-UP, and indeed the two events overlap on the morning of 16 April. This has allowed many MATCH-UP delegates to benefit from the travel funding available from the COST IC1205 budget, which in turn has no doubt helped to increase the overall numbers attending MATCH-UP. We are grateful for the valuable support from COST Action IC1205.

As in 2012, our call for papers invited two types of submission. Format A papers were required to be original and at most 12-pages long, whilst Format B papers had no restriction on length or originality. Accepted Format A papers appear in these proceedings in full, whilst only one-page abstracts of accepted Format B papers are included in what follows.

We received 56 submissions (13 Format A and 43 Format B), which were reasonably well-balanced in terms of representing the computing science and economics communities. Due to time constraints and our desire to avoid parallel sessions, we accepted 38 submissions (8 Format A and 30 Format B). To accommodate this number of papers, the workshop was extended to three days from the originally-planned duration of two days.

To give more authors an opportunity to present their work, we additionally accepted posters for presentation at a poster session. One-page abstracts of these posters (received in time for inclusion) also feature in these proceedings.

We feel that these papers and posters represent an excellent snapshot of the current state of the art regarding research in the area of matching problems with preferences.

We would like to thank the Programme Committee (and additional reviewers), the invited speakers and the authors of all submitted papers and posters for their important contributions to the scientific aspects of this workshop. Moreover we would like to thank the members of the Organising Committee, and Baharak Rastegari in particular, for all of their efforts. Additionally we thank several colleagues at the University of Glasgow for their assistance, namely Tania Galabova, May Gallagher, Lucinda Hay, Steven Kendrick, Jean Lindsay, Lydia Marshall, Aileen Orr and Sheena Phillips.

Last but not least, we acknowledge with gratitude the sponsorship that we received from SICSA (The Scottish Informatics and Computer Science Alliance) which enabled us to fund several assisted places for PhD students in Computer Science at SICSA universities.

Péter Biró
David Manlove
MATCH-UP 2015 PC co-chairs

Programme Committee

Itai Ashlagi	Massachusetts Institute of Technology
Eduardo Azevedo	Wharton School of the University of Pennsylvania
Péter Biró	Hungarian Academy of Sciences (co-chair)
Estelle Cantillon	Université Libre de Bruxelles
Brian Dean	Clemson University
Federico Echenique	California Institute of Technology
Lars Ehlers	Université de Montréal
Tamás Fleiner	Budapest University of Technology and Economics
Chien-Chung Huang	Chalmers University of Technology
Nicole Immorlica	Microsoft Research
Yuichiro Kamada	University of California, Berkeley
Flip Klijn	Universitat Autònoma de Barcelona
David Manlove	University of Glasgow (co-chair)
Antonio Miralles	Universitat Autònoma de Barcelona
Shuichi Miyazaki	Kyoto University
Baharak Rastegari	University of Glasgow
Jay Sethuraman	Columbia University
Akihasa Tamura	Keio University
Utku Ünver	Boston College
Alexander Westkamp	Maastricht University

Organising Committee

(All University of Glasgow unless stated otherwise)

Péter Biró	Hungarian Academy of Sciences
Frances Cooper	
Augustine Kwanashie	
David Manlove	(co-chair)
Iain McBride	
Ciaran McCreesh	
Hervé Moulin	
Josue Ortega	
Baharak Rastegari	(co-chair)

Contents

Invited Speaker Abstracts

Format A Papers

Format B Paper Abstracts

Poster Abstracts

Invited Speaker Abstracts

Assignment of teachers to schools—a new variation on an old theme

Katarína Cechlárová
Pavol Jozef Šafárik University in Košice, Slovakia

Several countries more or less successfully use centralized matching schemes for assigning teachers to vacant positions at schools. We explore combinatorial and computational aspects of a possible similar scheme motivated by the situation characteristic for Slovak and Czech education system where each teacher specializes in two subjects. We present a model that takes into consideration that schools may have different capacities for each subject and show that its combinatorial structure leads to intractable problems even under several strong restrictions concerning the total number of subjects, partial capacities of schools and the number of acceptable schools each teacher is allowed to list. We propose several approximation algorithms. Finally, we present integer programming models and their application to real data.

Fair Stable Matchings

Christine Cheng
University of Wisconsin-Milwaukee, USA

It has long been known that Gale and Shapley's deferred acceptance algorithm outputs stable matchings that are highly biased towards one side of the matching. This has motivated the study of fair stable matchings.

In the first part of the talk, we will discuss the class of globally fair stable matchings. Their fairness is derived from the fact that they are good representatives of the set of stable matchings. We will also describe how our results extend to other objects that form a distributive lattice.

In the second part of the talk, we will consider the Random Order Mechanism (ROM), an iterative version of the deferred acceptance algorithm. When the ordering of the agents is chosen uniformly at random, one can argue that the process ROM uses for arriving at a stable matching is procedurally fair. We shall present various computational results with regards to the stable matchings that ROM can output.

One dimensional mechanism design

Hervé Moulin
University of Glasgow, UK

When agents' allocations are one-dimensional and preferences are convex, the three perennial goals of mechanism design, efficiency, prior-free incentive compatibility and fairness (horizontal equity) are compatible. This has been known for decades in the cases of voting and of division of a non disposable commodity. We show that it is in fact true when the range of allocation profiles is an arbitrary convex and compact set.

Examples include: load balancing with arbitrary flow graph constraints; coordinating joint work inside a team or across teams, when individual contributions are substitutable or complementary; and any joint venture with a convex technology where each agent provides a single input or consumes a single output.

The set of efficient, incentive compatible and fair mechanisms is very rich, and additional requirements such as consistency are needed to identify reasonable candidates.

Format A Papers

The (Non)-Existence of Stable Mechanisms in Incomplete Information Environments

Nick Arnosti*, Nicole Immorlica†, Brendan Lucier†

February 16, 2015

Abstract

We consider two-sided matching markets, and study the incentives of agents to circumvent a centralized clearing house by signing binding contracts with one another. It is well-known that if the clearing house implements a stable match and preferences are known, then no group of agents can profitably deviate in this manner.

We ask whether this property holds even when agents have *incomplete information* about their own preferences or the preferences of others. We find that it does not. In particular, when agents are uncertain about the preferences of others, *every* mechanism is susceptible to deviations by groups of agents. When, in addition, agents are uncertain about their *own* preferences, every mechanism is susceptible to deviations in which a single pair of agents agrees in advance to match to each other.

1 Introduction

In entry-level labor markets, a large number of workers, having just completed their training, simultaneously seek jobs at firms. These markets are especially prone to certain failures, including unraveling, in which workers receive job offers well before they finish their training, and exploding offers, in which job offers have incredibly short expiration dates. In the medical intern market, for instance, prior to the introduction of the centralized clearing house (the *National Residency Matching Program*, or NRMP), medical students received offers for residency programs at US hospitals two years in advance of their employment date (Roth and Xing, 1994). In the market for law clerks, law students have reported receiving exploding offers in which they were asked to accept or reject the position on the spot (Roth and Xing, 1994).

In many cases, including the medical intern market in the United States and United Kingdom and the hiring of law students in Canada, governing agencies try to circumvent these market failures by introducing a centralized clearing house which solicits the preferences of all participants and uses these to recommend a matching (Roth, 1991). One main challenge of this approach is that of incentivizing participation. Should a worker and firm suspect they each prefer the other to their assignment by the clearing house, then they would likely match with each other and not participate in the centralized mechanism. Perhaps for this reason, clearing houses that fail to select a stable match have often had difficulty attracting participants and been discontinued (Roth, 1991).

Empirically, however, even clearing houses which produce stable matches may fail to prevent early contracting. Examples include the market for Canadian law students (Roth and Xing, 1994) and the American gastroenterology match (Niederle and Roth, 2004; McKinney et al., 2005). This is perhaps puzzling, as selecting a stable match ensures that no group of participants can profitably circumvent the clearing house ex-post.

*Dept. of Management Science & Engineering, Stanford University. Work conducted at Microsoft Research.
†Microsoft Research.

Our work offers one possible explanation for this phenomenon. While stable clearing houses ensure that for *fixed, known* preferences, no coalition can profitably deviate, in most natural settings, participants contemplating deviation do so without complete knowledge of others' preferences (and sometimes even their own preferences). Our main finding is that in the presence of such uncertainty, *no mechanism* can prevent agents from signing mutually beneficial side contracts.

We model uncertainty in preferences by assuming that agents have a common prior over the set of possible preference profiles, and may in addition know their own preferences. We consider two cases. In one, agents have no private information when contracting, and their decision of whether to sign a side contract depends only on the prior (and the mechanism used by the clearing house). In the second case, agents know their own preferences, but not those of others. When deciding whether to sign a side contract, agents consider their own preferences, along with the information revealed by the willingness (or unwillingness) of fellow agents to sign the proposed contract.

Note that with incomplete preference information, agents perceive the partner that they are assigned by a given mechanism to be a random variable. In order to study incentives for agents to deviate from the centralized clearing house, we must specify a way for agents to compare lotteries over match partners. One seemingly natural model is that each agent gets, from each potential partner, a utility from being matched to that partner. When deciding between two uncertain outcomes, agents simply compare their corresponding expected utilities. Much of the previous literature has taken this approach, and indeed, it is straightforward to discover circumstances under which agents would rationally contract early (see Appendix A). Such cases are perhaps unsurprising; after all, the central clearing houses that we study solicit only ordinal preference lists, while the competing mechanisms may be designed with agents' cardinal utilities in mind.

For this reason, we consider a purely ordinal notion of what it means for an agent to prefer one allocation to another. In our model, an agent debating between two uncertain outcomes chooses to sign a side contract only if the rank that they assign their partner under the proposed contract strictly first-order stochastically dominates the rank that they anticipate if all agents participate in the clearing house. This is a strong requirement, by which we mean that it is easy for a mechanism to be stable under this definition, relative to a definition relying on expected utility. For instance, this definition rules out examples of beneficial deviations, such as that given in Appendix A, where agents match to an acceptable, if sub-optimal, partner in order to avoid the possibility of a "bad" outcome.

Despite the strong requirements we impose on beneficial deviations, we show that every mechanism is vulnerable to side contracts when agents are initially uncertain about their preferences or the preferences of others. On the other hand, when agents are certain about their own preferences but not about the preferences of others, then there do exist mechanisms that resist the formation of side contracts, when those contracts are limited to involving only a pair of agents (i.e., one from each side of the market).

2 Related Work

Roth (1989) and Roth and Rothblum (1999) are among the first papers to model incomplete information in matching markets. These papers focus on the strategic implications of preference uncertainty, meaning that they study the question of whether agents should truthfully report to the clearinghouse. Our work, while it uses a similar preference model, assumes that the clearing house can observe agent preferences. While this assumption may be realistic in some settings, we adopt it primarily in order to separate the strategic manipulation of matching mechanisms (as studied in the above papers) from the topic of early contracting that is the focus of this work.

Since the seminal work of Roth and Xing (1994), the relationship between stability and unraveling has been studied using observational studies, laboratory experiments, and a range of theoretical models. Although some work concluded that stability played an important role in encouraging participation (Roth, 1991; Kagel and Roth, 2000), other papers note that uncertainty may cause unraveling to occur even if a stable matching mechanism is used.

A common theme in these papers is that unraveling is driven by the motive of "insurance." For example, the closely related models of Li and Rosen (1998); Suen (2000); Li and Suen (2000, 2004) study two-sided assignment models with transfers in which binding contracts may be signed in one of two periods (before or after revelation of pertinent information). In each of these papers, unraveling occurs (despite the stability of the second-round matching) because of agents' risk-aversion: when agents are risk-neutral, no early matches form.

Even in models in which transfers are not possible (and so the notion of risk aversion has no obvious definition), the motive of insurance often drives early matching. The models presented by Roth and Xing (1994), Halaburda (2010), and Du and Livne (2014) assume that agents have underlying cardinal utilities for each match, and compare lotteries over matchings by computing expected utilities. They demonstrate that unraveling may occur if, for example, workers are willing to accept an offer from their second-ranked firm (foregoing a chance to be matched to their top choice) in order to ensure that they do not match to a less-preferred option.[1]

While insurance may play a role in the early contracting observed by Roth and Xing (1994), one contribution of our work is to show that it is not necessary to obtain such behavior. In this work, we show that even if agents are unwilling to forego top choices in order to avoid lower-ranked ones, they might rationally contract early with one another. Put another way, we demonstrate that some opportunities for early contracting may be identified on the basis of ordinal information alone (without making assumptions about agents' unobservable cardinal utilities).

The works of Manjunath (2013) and Gudmundsson (2014) consider the stochastic dominance notion used in this paper; however they treat only the case (referred to in this paper as "ex-post") where the preferences of agents are fixed, and the only randomness comes from the assignment mechanism. One contribution of our work is to define a stochastic dominance notion of stability under asymmetric information. This can be somewhat challenging, as agents' actions signal information about their type, which in turn might influence the actions of others.[2]

Perhaps the paper that is closest in spirit to ours is that of Peivandi and Vohra (2013), which considers the operation of a centralized exchange in a two-sided setting with transferrable utility. One of their main findings is that every trading mechanism can be blocked by an alternative; our results have a similar flavor, although they are established in a setting with non-transferrable utility.

3 Model and Notation

In this section, we introduce our notation, and define what it means for a matching to be *ex-post*, *interim*, or *ex-ante* stable.

There is a (finite, non-empty) set M of men and a (finite, non-empty) set W of women.

Definition 1.
Given M and W, a **matching** *is a function $\mu : M \cup W \to M \cup W$ satisfying:*

[1] In many-to-one settings, Sonmez (1999) demonstrates that even in full-information environments, it may be possible for agents to profitably pre-arrange matches (a follow-up by Afacan (2013) studies the welfare effects of such pre-arrangements). In order for all parties involved to strictly benefit, it must be the case that the firm hires (at least) one inferior worker in order to boost competition for their remaining spots (and thereby receive a worker who they would be otherwise unable to hire). Thus, the profitability of such an arrangement again relies on assumptions about the firm's underlying cardinal utility function.

[2] The work of Liu et al. (2014) has recently grappled with this inference procedure, and defined a notion of stable matching under uncertainty. Their model differs substantially from the one considered here: it takes a matching μ as given, and assumes that agents know the quality of their current match, but must make inferences about potential partners to whom they are not currently matched.

1. *For each* $m \in M$*,* $\mu(m) \in W \cup \{m\}$

2. *For each* $w \in W$*,* $\mu(w) \in M \cup \{w\}$

3. *For each* $m \in M$ *and* $w \in W$*,* $\mu(m) = w$ *if and only if* $\mu(w) = m$*.*

We let $\mathcal{M}(M, W)$ be the set of matchings on M, W.

Given a set S, define $\mathcal{R}(S)$ to be the set of one-to-one functions mapping S onto $\{1, 2, \ldots, |S|\}$. Given $m \in M$, let $P_m \in \mathcal{R}(W \cup \{m\})$ be m's ordinal preference relation over women (and the option of remaining unmatched). Similarly, for $w \in W$, let $P_w \in \mathcal{R}(M \cup \{w\})$ be w's ordinal preference relation over the men. We think of $P_m(w)$ as giving the *rank* that m assigns to w; that is, $P_m(w) = 1$ implies that matching to w is m's most-preferred outcome.

Given sets M and W, we let $\mathcal{P}(M, W) = \prod_{m \in M} \mathcal{R}(W \cup \{m\}) \times \prod_{w \in W} \mathcal{R}(M \cup \{w\})$ be the set of possible preference profiles. We use P to denote an arbitrary element of $\mathcal{P}(M, W)$, and use ψ to denote a probability distribution over $\mathcal{P}(M, W)$. We use P_A to refer to the preferences of agents in the set A under profile P, and use P_a (rather than the more cumbersome $P_{\{a\}}$) to refer to the preferences of agent a.

Definition 2. *Given* M *and* W*, and* $P \in \mathcal{P}(M, W)$*, we say that matching* μ *is* **stable at preference profile** P *if and only if the following conditions hold.*

1. *For each* $a \in M \cup W$*,* $P_a(\mu(a)) \leq P_a(a)$*.*

2. *For each* $m \in M$ *and* $w \in W$ *such that* $P_m(\mu(m)) > P_m(w)$*, we have* $P_w(\mu(w)) < P_w(m)$*.*

This is the standard notion of stability; the first condition states that agents may only be matched to partners whom they prefer to going unmatched, and the second states that whenever m prefers w to his partner under μ, it must be that w prefers her partner under μ to m.

In what follows, we fix M and W, and omit the dependence of $\mathcal{M}$ and $\mathcal{P}$ on the sets M and W. We define a *mechanism* to be a (possibly random) mapping $\phi : \mathcal{P} \to \mathcal{M}$. We use A' to denote a subset of $M \cup W$.

We now define what it means for a coalition of agents to *block* the mechanism ϕ, and what it means for a *mechanism* (rather than a matching) to be stable. Because we wish to consider randomized mechanisms, we must have a way for agents to compare lotteries over outcomes. As mentioned in the introduction, our notion of blocking relates to stochastic dominance. Given random variables $X, Y \in \mathbb{N}$, say that X *first-order stochastically dominates* Y (denoted $X \succ Y$) if for all $n \in \mathbb{N}$, $\Pr(X \leq n) \geq \Pr(Y \leq n)$, with strict inequality for at least one value of n.

An astute reader will note that this definition reverses the usual inequalities; that is, $X \succ Y$ implies that X is "smaller" than Y. We adopt this convention because below, X and Y will represent the ranks assigned by each agent to their partner (where the most preferred option has a rank of one), and thus by our convention, $X \succ Y$ means that X is preferred to Y.

Definition 3 (Ex-Post Stability)**.** *Given* M, W *and a profile* $P \in \mathcal{P}(M, W)$*, coalition* A' **blocks mechanism** ϕ **ex-post** *at* P *if there exists a mechanism* ϕ' *such that for each* $a \in A'$*,*

1. $\Pr(\phi'(P)(a) \in A') = 1$*, and*

2. $P_a(\phi'(P)(a)) \succ P_a(\phi(P)(a))$*.*

Mechanism ϕ *is* **ex-post stable at profile** P *if no coalition of agents blocks* ϕ *ex-post at* P*.*
Mechanism ϕ *is* **ex-post stable** *if it is ex-post stable at* P *for all* $P \in \mathcal{P}(M, W)$*.*
Mechanism ϕ *is* **ex-post pairwise stable** *if for all* P*, no coalition consisting of at most one man and at most one woman blocks* ϕ *ex post at* P*.*

Note that in the above setting, because P is fixed, the mechanism ϕ' is really just a random matching. The first condition in the definition requires that the deviating agents can implement this alternative (random) matching without the cooperation of the other agents; the second condition requires that for each agent, the random variable denoting the rank of his partner under the alternative ϕ' stochastically dominates the rank of his partner under the original mechanism.

Note that if the mechanism ϕ is deterministic, then it is ex-post pairwise stable if and only if the matching it produces is stable in the sense of Definition 2.

The above notions of blocking and stability are concerned only with cases where the preference profile P is fixed. In this paper, we assume that at the time of choosing between mechanisms ϕ and ϕ', agents have incomplete information about the profile P that will eventually be realized (and used to implement a matching). We model this incomplete information by assuming that it is common knowledge that P is drawn from a prior ψ over $\mathcal{P}$. Given a mechanism ϕ, each agent may use ψ to determine the ex-ante distribution of the rank of the partner that they will be assigned by ϕ. This allows us to define what it means for a coalition to block ϕ ex-ante, and for a mechanism ϕ to be ex-ante stable.

Definition 4 (Ex-Ante Stability)**.** *Given M, W and a prior ψ over $\mathcal{P}(M, W)$, coalition A'* **blocks mechanism ϕ ex-ante** *at ψ if there exists a mechanism ϕ' such that if P is drawn from the prior ψ, then for each $a \in A'$,*

1. $\Pr(\phi'(P)(a) \in A') = 1$*, and*
2. $P_a(\phi'(P)(a)) \succ P_a(\phi(P)(a))$*.*

Mechanism ϕ is **ex-ante stable at prior ψ** *if no coalition of agents blocks ϕ ex-ante at ψ.*
Mechanism ϕ is **ex-ante stable** *if it is ex-ante stable at ψ for all priors ψ.*
Mechanism ϕ is **ex-ante pairwise stable** *if, for all priors ψ, no coalition consisting of at most one man and at most one woman blocks ϕ ex-ante at ψ.*

Note that the only difference between ex-ante and ex-post stability is that the randomness in Definition 4 is over both the realized profile P and the matching produced by ϕ, whereas in Definition 3, the profile P is deterministic. Put another way, the mechanism ϕ is ex-post stable if and only if it is ex-ante stable at all deterministic distributions ψ.

The notions of ex-ante and ex-post stability defined above are fairly straightforward because the information available to each agent is identical. In order to study the case where each agent knows his or her own preferences but not the preferences of others, we must define an appropriate notion of a blocking coalition. In particular, if man m decides to enter into a contract with woman w, m knows not only his own preferences, but also learns about those of w from the fact that she is willing to sign the contract. Our definition of what it means for a coalition to block ϕ in the interim takes this into account.

In words, given the common prior ψ, we say that a coalition A' *blocks ϕ in the interim* if there exists a preference profile P that occurs with positive probability under ψ such that when preferences are P, all members of A' agree that the outcome of ϕ' stochastically dominates that of ϕ, given their own preferences and the fact that other members of A' also prefer ϕ'. We formally define this concept below, where we use the notation $\psi(\cdot)$ to represent the probability measure assigned by the distribution ψ to the argument.

Definition 5 (Interim Stability)**.** *Given M, W, and a prior ψ over $\mathcal{P}(M, W)$, coalition A'* **blocks mechanism ϕ in the interim** *if there exists a mechanism ϕ', and for each $a \in A'$, a subset of preferences $\mathcal{R}_a$ satisfying the following:*

1. *For each $P \in \mathcal{P}$, $\Pr(\phi'(P)(a) \in A') = 1$.*
2. *For each agent $a \in A'$ and each preference profile $\tilde{P}_a$, $\tilde{P}_a \in \mathcal{R}_a$ if and only if*

(a) $\psi(Y_a(\tilde{P}_a)) > 0$*, where* $Y_a(\tilde{P}_a) = \{P\colon P_a = \tilde{P}_a\} \cap \{P\colon P_{a'} \in \mathcal{R}_{a'}\ \forall a' \in A' \backslash \{a\}\}$

(b) When P *is drawn from the conditional distribution of* ψ *given* $Y_a(\tilde{P}_a)$*, we have* $P_a(\phi'(P)(a)) \succ P_a(\phi(P)(a))$*.*

Mechanism ϕ *is* **interim stable at** ψ *if no coalition of agents blocks* ϕ *in the interim at* ψ*.*
Mechanism ϕ *is* **interim stable** *if it is interim stable at* ψ *for all distributions* ψ*.*
Mechanism ϕ *is* **interim pairwise stable** *if, for all priors* ψ*, no coalition consisting of at most one man and at most one woman blocks* ϕ *in the interim at* ψ*.*

To motivate the above definition of an interim blocking coalition, consider a game in which a moderator approaches a subset A' of agents, and asks each whether they would prefer to be matched according to the mechanism ϕ (proposed by the central clearing house) or the alternative ϕ' (which matches agents in A' to each other). Only if all agents agree that they would prefer ϕ' is this mechanism used. Condition 1 simply states that the mechanism ϕ' generates matchings among the (potentially) deviating coalition A'.

We think of $\mathcal{R}_a$ as being a set of preferences for which agent a agrees to use mechanism ϕ'. The set $Y_a(\tilde{P}_a)$ is the set of profiles which agent a considers possible, conditioned on the events $P_a = \tilde{P}_a$ and the fact that all other agents in A' agree to use mechanism ϕ'. Condition 2 is a consistency condition on the preference subsets $\mathcal{R}_a$: 2a) states that agents in A' should agree to ϕ' only if they believe that there is a chance that the other agents in A' will also agree to ϕ' (that is, if ψ assigns positive mass to Y_a); moreover, 2b) states that in the cases when $P_a \in \mathcal{R}_a$ *and the other agents select* ϕ', it should be the case that a "prefers" the mechanism ϕ' to ϕ (here and in the remainder of the paper, when we write that agent a prefers ϕ' to ϕ, we mean that *given the information available to* a, the rank of a's partner under ϕ' stochastically dominates the rank of a's partner under ϕ).

We now move on to our main results.

4 Results

We begin with the following observation, which states that the three notions of stability discussed above are comparable, in that ex-ante stability is a stronger requirement than interim stability, which is in turn a stronger requirement than ex-post stability.

Lemma 1.
If ϕ *is ex-ante (pairwise) stable, then it is interim (pairwise) stable.*
If ϕ *is interim (pairwise) stable, then it is ex-post (pairwise) stable.*

Proof. We argue the contrapositive in both cases. Suppose that ϕ is not ex-post stable. This implies that there exists a preference profile P, a coalition A', and a mechanism ϕ' that only matches agents in A' to each other, such that all agents in A' prefer ϕ' to ϕ, given P. If we take ψ to place all of its mass on profile P, then (trivially) A' also blocks ϕ in the interim, proving that ϕ is not interim stable.

Suppose now that ϕ is not interim stable. This implies that there exists a distribution ψ over $\mathcal{P}$, a coalition A', a mechanism ϕ' that only matches agents in A' to each other, and preference orderings $\mathcal{R}_a$ satisfying the following conditions: the set of profiles $Y = \{P : \forall a \in A', P_a \in \mathcal{R}_a\}$ has positive mass $\psi(Y) > 0$; and conditioned on the profile being in Y, agents in A' want to switch to ϕ' , i.e., for all $a \in A'$ and *for all* $P_a \in \mathcal{R}_a$ agent a prefers ϕ' to ϕ conditioned on the profile being in Y. Thus, agent a must prefer ϕ' even ex ante (conditioned only on $P \in Y$).

If we take ψ' to be the conditional distribution of ψ given $P \in Y$, it follows that under ψ', all agents $a \in A'$ prefer mechanism ϕ' to mechanism ϕ ex-ante, so ϕ is not ex-ante stable. □

4.1 Ex-post Stability

We now consider each of our three notions of stability in turn, beginning with ex-post stability. By Lemma 1, ex-post stability is the easiest of the three conditions to satisfy. Indeed, we show there not only exist ex-post stable mechanisms, but that any mechanism that commits to always returning a stable matching is ex-post stable.

Theorem 1.
Any mechanism that produces a stable matching with certainty is ex-post stable.

Note that if the mechanism ϕ is deterministic, then (trivially) it is ex-post stable if and only if it always produces a stable matching. Thus, for deterministic mechanisms, our notion of ex-post stability coincides with the "standard" definition of a stable mechanism. Theorem 1 states further that any mechanism that randomizes among stable matchings is also ex-post stable. This fact appears as Proposition 3 in (Manjunath, 2013).[3]

We next show in Example 1 that the converse of Theorem 1 does not hold. That is, there exist randomized mechanisms ϕ which sometimes select unstable matches but are nevertheless ex-post stable. In this and other examples, we use the notation $P_m : w_1, w_2, w_3$ as shorthand indicating that m ranks w_1 first, w_2 second, w_3 third, and considers going unmatched to be the least desirable outcome.

Example 1.

$$\begin{array}{llll} P_{m_1} & : w_1, w_2, w_3 & P_{w_1} & : m_3, m_2, m_1 \\ P_{m_2} & : w_1, w_3, w_2 & P_{w_2} & : m_2, m_1, m_3 \\ P_{m_3} & : w_2, w_1, w_3 & P_{w_3} & : m_3, m_2, m_1 \end{array}$$

There is a unique stable match, given by $\{m_1w_2, m_2w_3, m_3w_1\}$.

Lemma 2. *For the market described in Example 1, no coalition blocks the mechanism that outputs a uniform random matching.*

Proof. Because the random matching gives each agent their first choice with positive probability, if agent a is in a blocking coalition, then it must be that the agent that a most prefers is also in this coalition. Furthermore, any blocking mechanism must always match all participants, and thus any blocking coalition must have an equal number of men and women. Thus, the only possible blocking coalitions are $\{m_2, m_3, w_1, w_2\}$ or all six agents. The first coalition cannot block; if the probability that m_2 and w_2 are matched exceeds 1/3, m_2 will not participate. If the probability that m_3 and w_2 are matched exceeds 1/3, then w_2 will not participate. But at least one of these quantities must be at least 1/2.

Considering a mechanism that all agents participate in, for any set of weights on the six possible matchings, we can explicitly write inequalities saying that each agent must get their first choice with probability at least 1/3, and their last with probability at most 1/3. Solving these inequalities indicates that any random matching μ that (weakly) dominates a uniform random matching must satisfy

$$\Pr(\mu = \{m_1w_1, m_2w_2, m_3w_3\}) = \Pr(\mu = \{m_1w_2, m_2w_3, m_3w_1\}) = \Pr(\mu = \{m_1w_3, m_2w_1, m_3w_2\}),$$

$$\Pr(\mu = \{m_1w_1, m_2w_3, m_3w_2\}) = \Pr(\mu = \{m_1w_2, m_2w_1, m_3w_3\}) = \Pr(\mu = \{m_1w_3, m_2w_2, m_3w_1\}).$$

But any such mechanism gives each agent their first, second and third choices with equal probability, and thus does not strictly dominate the uniform random matching. □

Finally, the following lemma establishes a simple necessary condition for ex-post incentive compatibility. This condition will be useful for establishing non-existence of stable outcomes under other notions of stability.

Lemma 3.
If mechanism ϕ is ex-post pairwise stable, then if man m and woman w rank each other first under P, it follows that $\Pr(\phi(P)(m) = w) = 1$.

Proof. This follows immediately: if $\phi(P)$ matches m and w with probability less than one, then m and w can deviate and match to each other, and both strictly benefit from doing so. □

[3]We thank an anonymous reviewer for the reference.

4.2 Interim Stability

The fact that a mechanism which (on fixed input) outputs a uniform random matching is ex-post stable suggests that our notion of a blocking coalition, which relies on ordinal stochastic dominance, is very strict, and that many mechanisms may in fact be stable under this definition even with incomplete information. We show in this section that this intuition is incorrect: despite the strictness of our definition of a blocking coalition, it turns out that *no* mechanism is interim stable.

Theorem 2.
No mechanism is interim stable.

Proof. In the proof, we refer to *permutations* of a given preference profile P, which informally are preference profiles that are equivalent to P after a relabeling of agents. Formally, given a permutation σ on the set $M \cup W$ which satisfies $\sigma(M) = M$ and $\sigma(W) = W$, we say that P' is the **permutation of P obtained by** σ if for all $a \in M \cup W$ and a' in the domain of P_a, it holds that $P_a(a') = P'_{\sigma(a)}(\sigma(a'))$.

The proof of Theorem 2 uses the following example.

Example 2. *Suppose that each agent's preferences are iid uniform over the other side, and consider the following preference profile, which we denote P:*

$$\begin{array}{llcll} P_{m_1}: & w_1, w_2, w_3 & \quad & P_{w_1}: & m_1, m_2, m_3 \\ P_{m_2}: & w_1, w_3, w_2 & & P_{w_2}: & m_1, m_3, m_2 \\ P_{m_3}: & w_3, w_1, w_2 & & P_{w_3}: & m_3, m_1, m_2 \end{array}$$

Note that under profile P, m_1 and w_1 rank each other first, as do m_3 and w_3. By Lemma 1, if ϕ is interim stable, it must be ex-post stable. By Lemma 3, given this P, any ex-post stable mechanism must produce the match $\{m_1w_1, m_2w_2, m_3w_3\}$ with certainty. Furthermore, if preference profile P' is a permutation of P, then the matching $\phi(P')$ must simply permute $\{m_1w_1, m_2w_2, m_3w_3\}$ accordingly. Thus, on any permutation of P, ϕ gives four agents their first choices, and two agents their third choices.

Define the mechanism ϕ' as follows:

- If P' is the permutation of P obtained by σ, then
$$\phi'(P') = \{\sigma(m_1)\sigma(w_2), \sigma(m_2)\sigma(w_1), \sigma(m_3)\sigma(w_3)\}.$$
- On any profile that is not a permutation of P, ϕ' mimics ϕ.

Note that on profile P, ϕ' gives four agents their first choices, and two agents their second choices. If each agent's preferences are iid uniform over the other side, then each agent considers his or herself equally likely to play each role in the profile P (by symmetry, this is true even after agents observe their own preferences, as they know nothing about the preferences of others). Thus, conditioned on the preference profile being a permutation of P, all agents' interim expected allocation under ϕ offers a $2/3$ chance of getting their first choice and a $1/3$ chance of getting their third choice, while their interim allocation under ϕ' offers a $2/3$ chance of getting their first choice and a $1/3$ chance of getting their second choice. Because ϕ' and ϕ are identical on profiles which are not permutations of P, it follows that all agents strictly prefer ϕ' to ϕ ex-ante. □

The intuition behind the above example is as follows. Stable matchings may be "inefficient", meaning that it might be possible to separate a stable partnership (m_1, w_1) at little cost to m_1 and w_1, while providing large gains to their new partners (say m_2 and w_2). When agents lack the information necessary to determine whether they are likely to play the role of m_1 or m_2, they will gladly go along with the more efficient (though ex-post unstable) mechanism.

Note that in addition to proving that no mechanism is interim stable *for all priors*, Example 2 demonstrates that when the priori ψ is (canonically) taken to be uniform on $\mathcal{P}$, there exists no mechanism which is interim stable *at the prior* ψ (this follows because if ϕ sometimes fails to match pairs who rank each other first, then such pairs have a strict incentive to deviate; if ϕ always matches mutual first choices, then all agents prefer to deviate to the mechanism ϕ' described above).

Although Theorem 2 establishes that it is impossible to design a mechanism ϕ that eliminates profitable deviations, note that the deviating coalition in Example 2 involves six agents, and the contract ϕ' is fairly complex. In many settings, such coordinated action may seem implausible, and one might ask whether there exist mechanisms that are at least immune to deviations by *pairs* of agents. The following theorem shows that the complexity of Example 2 is necessary: any mechanism that always produces a stable match is indeed interim pairwise stable.[4]

Lemma 4.
Any mechanism which produces a stable match with certainty is interim pairwise stable.

Proof. Seeking a contradiction, suppose that ϕ always produces a stable match. Fix a man m, and a woman w with whom he might block ϕ in the interim. Note that m must prefer w to going unmatched; otherwise, no deviation with w can strictly benefit him. Thus, the best outcome (for m) from a contract with w is that they are matched with certainty. According to the definition of an interim blocking pair, m must believe that receiving w with certainty stochastically dominates the outcome of ϕ; that is to say, m must be certain that ϕ will give him nobody better than w. Because ϕ produces a stable match, it follows that in cases where m chooses to contract with w, ϕ always assigns to w a partner that she (weakly) prefers to m, and thus she will not participate. □

4.3 Ex-ante Stability

In some settings, it is natural to model agents as being uncertain not only about the rankings of others, but also about their own preferences. One might hope that the result of Theorem 4 extends to this setting; that is, that if ϕ produces a stable match with certainty, it remains immune to pairwise deviations ex-ante. Theorem 3 states that this is not the case: ex-ante, no mechanism is even pairwise stable.

Theorem 3.
No mechanism is ex-ante pairwise stable.

Proof. The proof of Theorem 3 uses the following example.

Example 3. *Suppose that there are three men and three women, and fix $p \in (0, 1/4)$. The prior ψ is that preferences are drawn independently as follows:*

$$\begin{array}{llll}
P_{m_1} = \begin{cases} w_1, w_3, w_2 & w.p.\ 1-2p \\ w_2, w_1, w_3 & w.p.\ p \\ w_3, w_2, w_1 & w.p.\ p \end{cases} & & P_{w_1} = \begin{cases} m_1, m_3, m_2 & w.p.\ 1-2p \\ m_2, m_1, m_3 & w.p.\ p \\ m_3, m_2, m_1 & w.p.\ p \end{cases} & \\
P_{m_2} = \quad w_1, w_2 & & P_{w_2} = \quad m_1, m_2 & \\
P_{m_3} = \quad w_3 & & P_{w_3} = \quad m_3 &
\end{array}$$

[4]This result relies crucially on the fact that we're using the notion of stochastic dominance to determine blocking pairs. If agents instead evaluate lotteries over matches by computing expected utilities, it is easy to construct examples where two agents rank each other second, and both prefer matching with certainty to the risk of getting a lower-ranked alternative from ϕ (see Appendix A).

Because m_3 and w_3 always rank each other first, we know by Lemmas 1 and 3 that if mechanism ϕ is ex-ante pairwise stable, it matches m_3 and w_3 with certainty. Applying Lemma 3 to the submarket $(\{m_1, m_2\}, \{w_1, w_2\})$, we conclude that

1. Whenever m_1 prefers w_2 to w_1, ϕ must match m_1 with w_2 (and m_2 with w_1) with certainty.
2. Whenever w_1 prefers m_2 to m_1, ϕ must match w_1 with m_2 (and m_1 with w_2) with certainty.
3. Whenever m_1 prefers w_1 to w_2 and w_1 prefers m_1 to m_2, ϕ must match m_1 with w_1.

After doing the relevant algebra, we see that w_1 and m_1 each get their first choice with probability $1-3p+4p^2$, their second choice with probability p, and their third choice with probability $2p-4p^2$. If w_1 and m_1 were to match to each other, they would get their first choice with probability $1-2p$, their second with probability p, and their third with probability p; an outcome that they both prefer. It follows that ϕ is not ex-ante pairwise stable, completing the proof. □

The basic intuition for Example 3 is similar to that of Example 2. When m_1 ranks w_1 first and w_1 does not return the favor, it is unstable for them to match and m_1 will receive his third choice. In this case, it would (informally) be more "efficient" (considering only the welfare of m_1 and w_1) to match m_1 with w_1; doing so improves the ranking that m_1 assigns his partner by two positions, while only lowering the ranking that w_1 assigns her partner by one. Because men and women play symmetric roles in the above example, ex-ante, both m_1 and w_1 prefer the more efficient solution in which they always match to each other.

5 Discussion

In this paper, we extended the notion of stability to settings in which agents are uncertain about their own preferences and/or the preferences of others. We observed that when agents can sign contracts before preferences are fully known, every matching mechanism is susceptible to unraveling. While past work has reached conclusions which sound similar, we argue that our results are stronger in several ways.

First, previous results have assumed that agents are expected utility maximizers, and relied on particular assumptions about the utilities that agents get from each potential partner. Our work uses the stronger notion of stochastic dominance to determine blocking coalitions, and notes that there may exist opportunities for profitable circumvention of a central matching mechanism even when agents are unwilling to sacrifice the chance of a terrific match in order to avoid a poor one.

Second, not only can every mechanism be blocked under *some* prior, but also, for some priors, it is impossible to design a mechanism that is interim stable *at that prior*. This striking conclusion is similar to that of Peivandi and Vohra (2013), who find (in a bilateral transferable utility setting) that for some priors over agent types, every potential mechanism of trade can be blocked.

In light of the above findings, one might naturally ask how it is that many centralized clearing houses have managed to persist. One possible explanation is that the problematic priors are in some way "unnatural" and unlikely to arise in practice. We argue that this is not the case: Example 2 shows that blocking coalitions exist when agent preferences are independent and maximally uncertain, Example 3 shows that they may exist even when the preferences of most agents are known, and in Appendix B we show that they may exist even when one side has perfectly correlated (i.e. ex-post identical) preferences.

A more plausible explanation for the persistence of centralized clearing houses is that although mutually profitable early contracting opportunities may exist, agents lack the ability to identify and/or act on them. To take one example, even when profitable early contracting opportunities can be identified, agents may lack the ability to write binding contracts with one another (whereas our work assumes that they possess such commitment power). We leave a more complete discussion of the reasons that stable matching mechanisms might persist in some cases and fail in others to future work.

References

Afacan, M. O. (2013). The welfare effects of pre-arrangements in matching markets. *Economic Theory*, 53(1):139–151.

Damiano, E., Li, H., and Suen, W. (2005). Unravelling of dynamic sorting. *The Review of Economic Studies*, 72(4):1057–1076.

Du, S. and Livne, Y. (2014). Rigidity of transfers and unraveling in matching markets. *Available at SSRN*.

Echenique, F. and Pereyra, J. S. (2013). Strategic uncertainty and unraveling in matching markets.

Gudmundsson, J. (2014). Sequences in Pairing Problems: A new approach to reconcile stability with strategy-proofness for elementary matching problems. 2014 Papers pgu351, Job Market Papers.

Halaburda, H. (2010). Unravelling in two-sided matching markets and similarity of preferences. *Games and Economic Behavior*, 69(2):365 – 393.

Kagel, J. H. and Roth, A. E. (2000). The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *The Quarterly Journal of Economics*, 115(1):pp. 201–235.

Li, H. and Rosen, S. (1998). Unraveling in matching markets. *The American Economic Review*, 88(3):pp. 371–387.

Li, H. and Suen, W. (2000). Risk sharing, sorting, and early contracting. *Journal of Political Economy*, 108(5):pp. 1058–1091.

Li, H. and Suen, W. (2004). Self-fulfilling early-contracting rush. *International Economic Review*, 45(1):301–324.

Liu, Q., Mailath, G. J., Postlewaite, A., and Samuelson, L. (2014). Stable matching with incomplete information. *Econometrica*, 82(2):541–587.

Manjunath, V. (2013). Stability and the core of probabilistic marriage problems. Technical report, Working paper.

McKinney, C. N., Niederle, M., and Roth, A. E. (2005). The collapse of a medical labor clearinghouse (and why such failures are rare). *The American Economic Review*, 95(3):pp. 878–889.

Niederle, M. and Roth, A. E. (2004). The gastroenterology fellowship match: How it failed and why it could succeed once again. *Gastroenterology*, 127(2):658 – 666.

Peivandi, A. and Vohra, R. (2013). On fragmented markets.

Roth, A. E. (1989). Two-sided matching with incomplete information about others' preferences. *Games and Economic Behavior*, 1(2):191 – 209.

Roth, A. E. (1991). A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the united kingdom. *The American Economic Review*, 81(3):pp. 415–440.

Roth, A. E. and Rothblum, U. G. (1999). Truncation strategies in matching markets in search of advice for participants. *Econometrica*, 67(1):21–43.

Roth, A. E. and Xing, X. (1994). Jumping the gun: Imperfections and institutions related to the timing of market transactions. *The American Economic Review*, 84(4):pp. 992–1044.

Sonmez, T. (1999). Can pre-arranged matches be avoided in two-sided matching markets? *Journal of Economic Theory*, 86(1):148 – 156.

Suen, W. (2000). A competitive theory of equilibrium and disequilibrium unravelling in two-sided matching. *The RAND Journal of Economics*, 31(1):pp. 101–120.

A Interim Pairwise (In)Stability

The following example shows that Theorem 4 depends on our stochastic dominance notion of a blocking pair; if agents compare lotteries by computing expected utilities, then pairs of agents might benefit from circumventing a mechanism that always produces a stable match.

Example 4. *There are three agents on each side. Men m_2 and m_3 are known to rank women in the order w_1, w_2, w_3; m_1 has this preference with probability $1 - p$, and with probability p ranks w_3 first. Symmetrically, women w_2 and w_3 are known to rank men in the order m_1, m_2, m_3; w_1 has this preference with probability $1 - p$, and with probability p ranks m_3 first.*

For any realization, there is a unique stable match; note that when m_1 ranks w_3 first and w_1 ranks m_1 first and m_3 last, this match gives w_2 her least-preferred partner, m_3. Under a stable matching mechanism, both m_2 and w_2 get their first choice with probability $p(1 - p)$, their second choice with probability $(1 - p)^2 + p^2$, and their third choice with probability $p(1 - p)$. So long as their utility from their second choice is above their average utility from a lottery over their first and third choices, m_2 and w_2 prefer matching with one another to the outcome of the stable matching.

B Perfectly Correlated Preferences

Theorem 3 demonstrates that a stable matching mechanism may be blocked ex-ante by a coalition when preferences are drawn independently and uniformly at random.

The following example considers an opposite extreme extreme, where one side has identical preferences ex-post. It demonstrates that even in this case, it may be possible for a coalition to profitably deviate ex-ante from a mechanism that always selects the unique stable matching.

In this appendix, we use the language of "schools" and "students," and assume that schools all rank students according to a common test.

Example 5. *Each student has one of four possible preference profiles, drawn independently:*

$$\begin{array}{lll} A, B, C & w.p. & (1-\delta)/2 \\ A, C, B & w.p. & \delta/2 \\ B, A, C & w.p. & (1-\delta)/2 \\ B, C, A & w.p. & \delta/2 \end{array}$$

Schools have aligned preferences ex-post. The possibilities are the following:

$$\begin{array}{lll} 1, 2, 3 & w.p. & (1-\epsilon)/2 \\ 1, 3, 2 & w.p. & \epsilon/2 \\ 2, 1, 3 & w.p. & (1-\epsilon)/2 \\ 2, 3, 1 & w.p. & \epsilon/2 \end{array}$$

If all agents participate in an assortative match, schools A and B get their first, second, and third choices with probabilities $(\frac{1}{2}, \frac{1-\delta}{2}, \frac{\delta}{2})$ respectively. Students 1 and 2 get their first, second, and third choices with probabilities $\left(\frac{3}{4}, \frac{1}{4}, 0\right) - \frac{\epsilon}{8}\left(2 - \delta, 2 - 5\delta + 3\delta^2, -4 + 6\delta - 3\delta^2\right)$.

If only $(A, B, 1, 2)$ participate in an assortative match, then the associated match probabilities for schools A and B are $(\frac{1}{2}, \frac{1-\epsilon}{2}, \frac{\epsilon}{2})$, and for students 1 and 2 are $\left(\frac{3}{4}, \frac{1}{4}, 0\right) - \frac{\delta}{4}\left(0, 1, -1\right)$.

All four of $A, B, 1, 2$ prefer the latter option if $\epsilon < \delta < 2\epsilon(1 - \frac{3}{2}\delta + \frac{3}{4}\delta^2)$.

Possible and Necessary Allocations via Sequential Mechanisms

Haris Aziz
NICTA and UNSW,
Sydney 2033, Australia

Toby Walsh
NICTA and UNSW,
Sydney 2033, Australia

Lirong Xia
RPI
NY 12180, USA

Abstract

A simple mechanism for allocating indivisible resources is sequential allocation in which agents take turns to pick items. We focus on *possible* and *necessary allocation* problems, checking whether allocations of a given form occur in *some* or *all* mechanisms for several commonly used classes of sequential allocation mechanisms. In particular, we consider whether a given agent receives a given item, a set of items, or a subset of items for natural classes of sequential allocation mechanisms: balanced, recursively balanced, balanced alternation, and strict alternation. We present characterizations of the allocations that result respectively from the classes, which extend the well-known characterization by Brams and King [2005] for policies without restrictions. In addition, we examine the computational complexity of possible and necessary allocation problems for these classes.

1 Introduction

Efficient and fair allocation of resources is a pressing problem within society today. One important and challenging case is the fair allocation of indivisible items [Chevaleyre et al., 2006, Bouveret and Lang, 2008, Bouveret et al., 2010, Aziz et al., 2014b, Aziz, 2014]. This covers a wide range of problems including the allocation of classes to students, landing slots to airlines, players to teams, and houses to people. A simple but popular mechanism to allocate indivisible items is *sequential allocation* [Bouveret and Lang, 2011, Brams and Taylor, 1996, Kohler and Chandrasekaran, 1971, Levine and Stange, 2012]. In sequential allocation, agents simply take turns to pick the most preferred item that has not yet been taken. Besides its simplicity, it has a number of advantages including the fact that the mechanism can be implemented in a distributed manner and that agents do not need to submit cardinal utilities. Well-known mechanisms like serial dictatorship [Svensson, 1999] fall under the umbrella of sequential mechanisms.

The sequential allocation mechanism leaves open the particular order of turns (the so called "policy") [Kalinowski et al., 2013a, Bouveret and Lang, 2014]. Should it be a *balanced* policy i.e., each agent gets the same total number of turns? Or should it be *recursively balanced* so that turns occur in rounds, and each agent gets one turn per round? Or perhaps it would be fairer to alternate but reverse the order of the agents in successive rounds: $a_1 \rhd a_2 \rhd a_3 \rhd a_3 \rhd a_2 \rhd a_1 \dots$ so that agent a_1 takes the first and sixth turn? This particular type of policy is used, for example, by the Harvard Business School to allocate courses to students [Budish and Cantillion, 2012] and is referred to as a *balanced alternation* policy. Another class of policies is *strict alternation* in which the same ordering is used in each round, such as $a_1 \rhd a_2 \rhd a_3 \rhd a_1 \rhd a_2 \rhd a_3 \dots$. The sets of balanced alternation and strict alternation policies are subsets of the set of recursively balanced policies which itself is a subset of the set of balanced policies.

We consider here the situation where a policy is chosen from a family of such policies. For example, at the Harvard Business School, a policy is chosen at random from the space of all balanced alternation policies. As a second example, the policy might be left to the discretion of the chair but, for fairness, it is restricted to one of the recursively balanced policies. Despite uncertainty in the policy, we might be interested in the possible or necessary outcomes. For example, can I get my three most preferred courses? Do I necessarily get my two most preferred courses? We examine the complexity of checking such questions. There are several high-stake applications for these results. For example, sequential allocation is used in professional sports 'drafts' [Brams and Straffin, 1979]. The precise policy chosen from among the set of admissible policies can critically affect which teams (read agents) get which players (read items).

The problems of checking whether an agent can get some item or set of items in a policy or in all policies is closely related to the problem of 'control' of the central organizer. For example, if an agent gets an item in all feasible policies, then it means that the chair cannot ensure that the agent does not get the item. Apart from strategic motivation, the problems we consider also have a design motivation. The central designer may want to consider all feasible policies uniformly at random (as is the case in random serial dictatorship [Aziz et al., 2013, Saban and Sethuraman, 2013]) and use them to find the probability that a certain item or set of item is given to an agent. The probability can be a suggestion of time sharing of an item. The problem of checking whether an agent gets a certain item or set of items in some policy is equiva-

Problems	Sequential Policy Class				
	Any	Balanced	Recursively Balanced	Strict Alternation	Balanced Alternation
PossibleItem	in P	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)
NecessaryItem	in P	coNPC (Thm. 9); in P for const. k (Thm. 10)	coNPC for all $k \geq 2$ (Thm. 13)	coNPC for all $k \geq 2$ (Thm. 19)	coNPC for all $k \geq 2$ (Thm. 21)
PossibleSet	in P	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)
NecessarySet	in P	in P (Thm. 11)	coNPC for all $k \geq 2$ (Thm. 13)	coNPC for all $k \geq 2$ (Thm. 19)	coNPC for all $k \geq 2$ (Thm. 22)
Top-k PossibleSet	in P	in P (trivial)	NPC for all $k \geq 3$ (Thm. 15); in P for $k = 2$ (Thm. 14)	NPC for all $k \geq 3$ (Thm. 18); in P for $k = 2$ (Thm. 17)	NPC for all $k \geq 2$ (Thm. 21)
Top-k NecessarySet	in P	in P (Thm. 11)	coNPC for all $k \geq 2$ (Thm. 13)	coNPC for all $k \geq 2$ (Thm. 19)	coNPC for all $k \geq 2$ (Thm. 22)
PossibleSubset	in P	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)	NPC (Thm. 4)
NecessarySubset	in P	coNPC (Thm. 9); in P for const. k (Thm. 8)	coNPC for all $k \geq 2$ (Thm. 13)	coNPC for all $k \geq 2$ (Thm. 19)	coNPC for all $k \geq 2$ (Thm. 21)
PossibleAssignment	in P	in P (Coro. 1)	in P (Coro. 2)	in P (Coro. 3)	in P (Coro. 4)
NecessaryAssignment	in P	in P (Thm. 7)	in P (Thm. 12)	in P (Thm. 16)	in P (Thm. 20)

Table 1: Complexity of possible and necessary allocation for sequential allocation. All possible allocation problems are NPC for $k = 1$. All necessary problems are in P for $k = 1$.

lent to checking whether an agent gets a certain item or set of items with non-zero probability. Similarly, the problem of checking whether an agent gets a certain item or set of items in all policy is equivalent to checking whether an agent gets a certain item or set of items with probability one.

We let $A = \{a_1, \dots, a_n\}$ denote a set of n agents, and I denote the set of $m = kn$ items[1]. $P = (P_1, \dots, P_n)$ is the profile of agents' preferences where each P_j is a linear order over I. Let M denote an assignment of all items to agents, that is, $M : I \rightarrow A$. We will denote a class of policies by $\mathcal{C}$. Any policy π specifies the $|I|$ turns of the agents. When an agent takes her turn, she picks her most preferred item that has not yet been allocated.

Example 1. *Consider the setting in which $A = \{a_1, a_2\}$, $I = \{b, c, d, e\}$, the preferences of agent a_1 are $b \succ c \succ d \succ e$ and of agent a_2 are $b \succ d \succ c \succ e$. Then for the policy $a_1 \rhd a_2 \rhd a_2 \rhd a_1$, agent a_1 gets $\{b, e\}$ whilst a_2 gets $\{c, d\}$.*

We consider the following natural computational problems.

(i) PossibleAssignment: Given (A, I, P, M) and policy class $\mathcal{C}$, does there exist a policy in $\mathcal{C}$ which results in M?; (ii) necessaryassignment : Given (A, I, P, M), and policy class $\mathcal{C}$, is M the result of all policies in $\mathcal{C}$? ; (iii) PossibleItem: Given (A, I, P, a_j, o) where $a_j \in A$ and $o \in I$, and policy class $\mathcal{C}$, does there exist a policy in $\mathcal{C}$ such that agent a_j gets item o? ; (iv) NecessaryItem: Given (A, I, P, a_j, o) where $a_j \in A$ and $o \in I$, and policy class $\mathcal{C}$, does agent a_j get item o for all policies in $\mathcal{C}$? ; (v) PossibleSet: Given (A, I, P, a_j, I') where $a_j \in A$ and $I' \subseteq I$, and policy class $\mathcal{C}$, does there exist a policy in $\mathcal{C}$ such that agent a_j gets exactly I'? ; (vi) NecessarySet: Given (A, I, P, a_j, I') where $a_j \in A$ and $I' \subseteq I$, and policy class $\mathcal{C}$, does agent a_j get exactly I' for all policies in $\mathcal{C}$? ; (vii) PossibleSubset: Given (A, I, P, a_j, I') where $a_j \in A$ and $I' \subseteq I$, and policy class $\mathcal{C}$, does there exist a policy in $\mathcal{C}$ such that agent a_j gets I'? ; $(viii)$ NecessarySubset: Given (A, I, P, a_j, I') where $a_j \in A$ and $I' \subseteq I$, and policy class $\mathcal{C}$ does agent a_j get I' for all policies in $\mathcal{C}$?

[1] This is without loss of generality since we can add dummy items of no utility to any agent.

We will consider problems top-k PossibleSet and top-k NecessarySet that are restrictions of PossibleSet and NecessarySet in which the set of items I' is the set of top k items of the distinguished agent. When policies are chosen at random, the possible and necessary allocation problems we consider are also fundamental to understand more complex problems of computing the probability of certain allocations.

Contributions: Our contributions are two fold. First, we provide necessary and sufficient conditions for an allocation to be the outcome of balanced, recursively balanced, balanced alternation, and strict alternation policies respectively. Previously Brams and King [2005] characterized the outcomes of arbitrary policies. In a similar vein, we provide sufficient and necessary conditions for more interesting classes of policies such as recursively balanced and balanced alternation. Second, we provide a detailed analysis of the computational complexity of possible and necessary allocations under sequential policies. Table 1 summarizes our complexity results. Our NP/coNP-completeness results also imply that there exists no polynomial-time algorithm that can approximate within any factor the number of admissible policies which do or do not satisfy the target goals.

Related Work. Sequential allocation has been considered in the operations research and fair division literature (e.g. [Kohler and Chandrasekaran, 1971, Brams and Taylor, 1996]). It was popularized within the AI literature as a simple yet effective distributed mechanism [Bouveret and Lang, 2011] and has been studied in more detail subsequently [Kalinowski et al., 2013a,b, Bouveret and Lang, 2014, 2011, 2014].

The problems considered in the paper are similar in spirit to a class of control problems studied in voting theory: if it is possible to select a voting rule from the set of voting rules, can one be selected to obtain a certain outcome [Erdélyi and Elkind, 2012]. They are also related to a class of control problems in knockout tournaments: does there exist a draw of a tournament for which a given player wins the tournament [Vu et al., 2009, Aziz et al., 2014a]. Possible and necessary winners have also been considered in voting theory [Konczak and Lang, 2005, Xia and Conitzer, 2011, Aziz et al., 2012].

When $n = m$, serial dictatorship is a well-known mechanism in which there is an ordering of agents and with respect to that ordering agents pick the most preferred unallocated item in their turns [Svensson, 1999]. We note that serial dictatorship for $n = m$ is a balanced, recursively balanced and balanced alternation policy.

2 Characterizations of Outcomes of Sequential Allocation

In this section we provide necessary and sufficient conditions for a given allocation to be the outcome of a balanced policy, recursively balanced policy, or balanced alternation policy. We first define conditions on an allocation M. An allocation is *Pareto optimal* if there is no other allocation in which each item of each agent is replaced by at least as preferred an item and at least one item of some agent is replaced by a more preferred item.

Condition 1. *M is Pareto Optimal.*

Condition 2. *M is balanced.*

It is well-known that Condition 1 characterizes outcomes of all sequential allocation mechanisms (without constraints). Brams and King [2005] proved that an assignment is achievable via sequential allocation iff it satisfies Condition 1. The theorem of Brams and King [2005] generalized the characterization of Abdulkadiroğlu and Sönmez [1998] of Pareto optimal assignments as outcomes of serial dictatorships when $m = n$. We first observe the following simple adaptation of the characterization of Brams and King [2005] to characterize possible outcomes of balanced policies:

Remark 1. *Given a profile P, an allocation M is the outcome of a balanced policy if and only if M satisfies Conditions 1 and 2.*

Given a balanced allocation M, for each agent $a_j \in A$ and each $i \leq k$, let p^i_j denote the item that is ranked at the i-th position by agent a_j among all items allocated to agent a_j by M. The third condition requires that for all $1 \leq t < s \leq k$, no agent prefers the s-th ranked item allocated to any other agent to the t-th ranked item allocated to her.

Condition 3. *For all $1 \leq t < s \leq k$ and all pairs of agent $a_j, a_{j'}$, agent a_j prefers p^t_j to $p^s_{j'}$.*

The next theorem states that Conditions 1 through 3 characterize outcomes of recursively balanced policies.

Theorem 1. *Given a profile P, an allocation M is the outcome of a recursively balanced policy if and only if it satisfies Conditions 1, 2, and 3.*

Proof. To prove the "only if" direction, clearly if M is the outcome of a recursively balanced policy then Condition 1 and 2 are satisfied. If Condition 3 is not satisfied, then there exists $1 \leq t < s \leq k$ and a pair of agents $a_j, a_{j'}$ such that agent a_j prefers $p^s_{j'}$ to p^t_j. We note that in the round when agent a_j is about to choose p^t_j according to M, $p^s_{j'}$ is still available, because it is allocated by M in a later round. However, in this case agent a_j will not choose p^t_j because it is not her top-ranked available item, which is a contradiction.

To prove the "if" direction, for any allocation M that satisfies the three conditions we will construct a recursively balanced policy π. For each $i \leq k = m/n$, we let **phase** i denote the $((i-1)n+1)$-th round through in-th round. It follows that for all $i \leq k$, $\{p^i_j : j \leq n\}$ are allocated in phase i. Because of Condition 3, $\{p^i_j : j \leq n\}$ is a Pareto optimal allocation when all items in $\{p^{i'}_j : i' < i, j \leq n\}$ are removed. Therefore there exists an order π_i over A that gives this allocation. Let $\pi = \pi_1 \rhd \pi_2 \rhd \cdots \rhd \pi_k$. It is not hard to verify that π is recursively balanced and M is the outcome of π. $\square$

Given a profile P and an allocation M that is the outcome of a recursively balanced policy, that is, it satisfies the three conditions as proved in Theorem 1, we construct a directed graph $G_M = (A, E)$, where the vertices are the agents, and we add the edges in the following way. For each odd $i \leq k$, we add a directed edge $a_{j'} \to a_j$ if and only if agent a_j prefers $p^i_{j'}$ to p^i_j and the edge is not already in G_M; for each even $i \leq k$, we add a directed edge $a_j \to a_{j'}$ if and only if agent a_j prefers $p^i_{j'}$ to p^i_j and the edge is not already in G_M.

Condition 4. *Suppose M is the outcome of a recursively balanced policy. There is no cycle in G_M.*

Theorem 2. *An allocation M is achievable by a balanced alternation policy if and only if satisfies Conditions 1, 2, 3, and 4.*

Proof. The "only if" direction: Suppose M is achievable by a balanced alternation policy π. Let π' denote the suborder of π from round 1 to round n. Let $G_{\pi'} = (A, E')$ denote the directed graph where the vertices are the agents and there is an edge $a_{j'} \to a_j$ if and only if $a_{j'} \rhd_{\pi'} a_j$. It is easy to see that $G_{\pi'}$ is acyclic and complete. We claim that G_M is a subgraph of $G_{\pi'}$. For the sake of contradiction suppose there is an edge $a_j \to a_{j'}$ in G_M but not in $G_{\pi'}$. If $a_j \to a_{j'}$ is added to G_M in an odd round i, then it means that agent j' prefers p^i_j to $p^i_{j'}$. Because $a_j \to a_{j'}$ is not in $G_{\pi'}$, $a_{j'} \rhd_{\pi'} a_j$. This means that right before $a_{j'}$ choosing $p^i_{j'}$ in M, p^i_j is still available, which contradicts the assumption that $a_{j'}$ chooses $p^i_{j'}$ in M. If $a_j \to a_{j'}$ is added to G_M in an even round, then following a similar argument we can also derive a contradiction. Therefore, G_M is a subgraph of $G_{\pi'}$, which means that G_M is acyclic.

The "if" direction: Suppose the four conditions are satisfied. Because G_M has no cycle, we can find a linear order π' over A such that G_M is a subgraph of $G_{\pi'}$. We next prove that M is achievable by the balanced alternation policy π whose first n rounds are π'. For the sake of contradiction suppose this is not true and let t denote the earliest round that the allocation in π differs the allocation in M. Let a_j denote the agent at the t-th round of π, let $p^{i'}_{j'}$ denote the item she gets at round t in π, and let p^i_j denote the item that she is supposed to get according to M. Due to Condition 3, $i' \leq i$. If $i' < i$ then agent $a_{j'}$ didn't get item $p^{i'}_{j'}$ in a previous round, which contradicts the selection of t. Therefore $i' = i$. If i is odd, then there is an edge $a_{j'} \to a_j$ in G_M, which means that $a_{j'} \rhd_{\pi'} a_j$. This means that $a_{j'}$ would have chosen $p^i_{j'}$ in a previous round, which is a contradiction. If i is even, then a similar contradiction can be derived. Therefore M is achievable by π. ☐

Given a profile P and an allocation M that is the outcome of a recursively balanced policy, that is, it satisfies the three conditions as proved in Theorem 1, we construct a directed graph $H_M = (A, E)$, where the vertices are the agents, and we add the edges in the following way. For each $j \leq n$ and $i \leq k$, we let p_j^i denote the item that is ranked at the i-th position among all items allocated to agent j. For each $i \leq k$, if we add a directed edge $a_{j'} \rightarrow a_j$ if j prefers $p_{j'}^i$ to p_j^i if the edge is not already there.

Condition 5. *Suppose M is the outcome of a recursively balanced policy. There is no cycle in H_M.*

Theorem 3. *An allocation M is achievable by a strict alternation policy if and only if satisfies Condition 1, 2, 3, and 5.*

Proof. The "only if" direction: If M is an outcome of a recursively balanced policy but does not satisfy 5, then this means that there is a cycle in H_M. Let agents a_i and a_j be in the cycle. This means that a_i is before a_j in one round and a_j is before a_i in some other round.

The "if" direction: Now assume that M is an outcome of a recursively balanced policy but is not alternating. This means that there exist at least two agents a_i and a_j such that a_i comes before a_j in one round and a_j comes before a_i in some other round. But this means that there is cycle $a_i \rightarrow a_j \rightarrow a_i$ in graph H_M. ☐

3 General Complexity Results

Before we delve into the complexity results, we observe the following reductions between various problems.

Lemma 1. *Fixing the policy class to be one of $\{$all, balanced policies, recursively balanced policies, balanced alternation policies$\}$, there exist polynomial-time many-one reductions between the following problems:* POSSIBLESET *to* POSSIBLESUBSET*;* POSSIBLEITEM *to* POSSIBLESUBSET*; Top-k* POSSIBLESET *to* POSSIBLESET*;* NECESSARYSET *to* NECESSARYSUBSET*;* NECESSARYITEM *to* NECESSARYSUBSET*; and Top-k* NECESSARYSET *to* NECESSARYSET.

A polynomial-time many-one reduction from problem Q to problem Q' means that if Q is NP(coNP)-hard then Q' is also NP(coNP)-hard, and if Q' is in P then Q is also in P. We also note the following. For $n = 2$, POSSIBLEASSIGNMENT and POSSIBLESET are equivalent for any type of policies. Since $n = 2$, the allocation of one agent completely determines the overall assignment.

For $m = n$, checking whether there is a serial dictatorship under which each agent gets exactly one item and a designated agent a_j gets item o is NP-complete [Theorem 2, Saban and Sethuraman, 2013]. They also proved that for $m = n$, checking if for all serial dictatorships, agent a_j gets item o is polynomial-time solvable. Hence, we get the following statements.

Theorem 4. POSSIBLEITEM *and* POSSIBLESET *is NP-complete for balanced, recursively balanced as well as balanced alternation policies.*

Theorem 4 does not necessarily hold if we consider the top element or the top k elements. Therefore, we will especially consider top-k POSSIBLESET. Fimilarly, we get that for $m = n$, NECESSARYITEM and NECESSARYSET is polynomial-time solvable for balanced, recursively balanced, and balanced alternation policies.

For arbitrary policies, we first observe that POSSIBLEITEM, NECESSARYITEM and NECESSARYSET are trivial: POSSIBLEITEM always has a yes answer (just give all the turns to that agent) and NECESSARYITEM and NECESSARYSET always have a no answer (just don't give the agent any turn). Similarly, NECESSARYASSIGNMENT always has a no answer.

Theorem 5. POSSIBLEASSIGNMENT *is polynomial-time solvable for arbitrary policies.*

Proof. By the characterization of Brams and King [2005], all we need to do is to check whether the assignment is Pareto optimal. It can be checked in polynomial time $O(|I|^2)$ whether a given assignment is Pareto optimal via an extension of a result Abraham et al. [2005]. ☐

There is also a polynomial-time algorithm for POSSIBLESET for arbitrary policies via a greedy approach.

Theorem 6. POSSIBLESET *is polynomial-time solvable for arbitrary policies.*

4 Balanced Policies

In contrast to arbitrary policies, POSSIBLEITEM, NECESSARYITEM, NECESSARYSET, and NECESSARYASSIGNMENT are more interesting for balanced policies since we may be restricted in allocating items to a given agent to ensure balance. Before we consider them, we get the following corollary of Remark 1.

Corollary 1. POSSIBLEASSIGNMENT *for balanced assignments is in P.*

Note that an assignment is achieved via all balanced policies iff the assignment is the unique balanced assignment that is Pareto optimal. This is only possible if each agent gets his top k items. Hence, we obtain the following.

Theorem 7. NECESSARYASSIGNMENT *for balanced assignments is in P.*

Compared to NECESSARYASSIGNMENT, the other 'necessary' problems are intractable.

Theorem 8. *For any constant k,* NECESSARYSET *and* NECESSARYSUBSET *for balanced policies are in P.*

Proof. W.l.o.g. given a NECESSARYSET instance (A, I, P, a_1, I'), if I' is not the top-ranked k items of agent a_1 then it is a "No" instance because we can simply let agent a_1 choose items in the first k rounds. When I' is top-ranked k items of agent a_1, (A, I, P, a_1, I') is a "No" instance if and only if (A, I, P, a_1, o) is a "No" instance for some $o \in I'$, which can be checked in polynomial time by Theorem 10. A similar algorithm works for NecessarySubset. □

Theorem 9. NecessaryItem *and* NecessarySubset *for balanced policies where k is not fixed is coNP-complete.*

Proof. Membership in coNP is obvious. By Lemma 1 it suffices to prove that NecessaryItem is coNP-hard, which we will prove by a reduction from PossibleItem for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. Let (A, I, P, a_1, o) denote an instance of the possible allocation problem for $k = 1$, where $A = \{a_1, \ldots, a_n\}$, $I = \{o_1, \ldots, o_n\}$, $o \in I$, $P = (P_1, \ldots, P_n)$ is the preference profile of the n agents, and we are asked whether it is possible for agent a_1 to get item o in some sequential allocation. Given (A, I, P, a_1, o), we construct the following NecessaryItem instance.

Agents: $A' = A \cup \{a_{n+1}\}$.

Items: $I' = I \cup D \cup F_1 \cup \cdots \cup F_n$, where $|D| = n-1$ and for each $a_j \in A$, $|F_j| = n-2$. We have $|I'| = (n+1)(n-1)$ and $k = n-1$.

Preferences:

- The preferences of a_1 is $[F_1 \succ P_1 \succ \text{others}]$.
- For any $j \leq n$, the preferences of a_j are obtained from $[F_j \succ P_j]$ by replacing o by D, and then add o to the bottom position.
- The preferences for a_{n+1} is $[o \succ \text{others}]$.

We are asked whether agent a_{n+1} always gets item o.

If (A, I, P, a_1, o) has a solution π, we show that the NecessaryItem instance is a "No" instance by considering $\underbrace{\pi \rhd \cdots \rhd \pi}_{n-1} \rhd \underbrace{a_{n+1} \rhd \cdots \rhd a_{n+1}}_{n-1}$. In the first $(n-2)n$ rounds all F_j's are allocated to agent a_j's. In the following n rounds o is allocated to a_1, which means that a_{n+1} does not get o.

Suppose the NecessaryItem instance is a "No" instance and agent $n+1$ does not get o in a balanced policy π'. Because agent a_2 through a_n rank o in their bottom position, o must be allocated to agent a_1. Clearly in the first $n-2$ times when agent a_1 through a_n choose items, they will choose F_1 through F_n respectively. Let π denote the order over which agents a_1 through a_n choose items for the last time. We obtain another order π^* over A from π by moving all agents who choose an item in D after agent a_1 while keeping other orders unchanged. It is not hard to see that the outcomes of running π and π^* are the same from the first round until agent a_1 gets o. This means that π^* is a solution to (A, I, P, a_1, o). □

The problems becomes easier when k is constant or we are concerned about top k items.

Theorem 10. *For any constant k,* NecessaryItem *for balanced policies is in P.*

Proof. Given a NecessaryItem instance (A, I, P, a_1, o), if o is ranked below the k-th position by agent a_1 then we can return "No", because by letting agent a_1 choose in the first k rounds she does not get item o. Suppose o is ranked at the k'-th position by agent a_1 with $k' \leq k$, the next claim provides an equivalent condition to check whether the NecessaryItem instance is a "No" instance.

Claim 1. *Suppose o is ranked at the k'-th position by agent a_1 with $k' \leq k$, the* NecessaryItem *instance (A, I, P, a_1, o) is a "No" instance if and only if there exists a balanced policy π such that (i) agent a_1 picks items in the first $k'-1$ rounds and the last $k-k'+1$ rounds, and (ii) agent a_1 does not get o.*

Let I^* denote agent a_1's top $k'-1$ items. In light of the claim above, to check whether the (A, I, P, a_1, o) is a "No" instance, it suffices to check for every set of $k-k'+1$ items ranked below the k'-th position by agent a_1, denoted by I', whether it is possible for agent a_1 to get I^* and I' by a balanced policy where agent a_1 picks items in the first $k'-1$ rounds and the last $k-k'+1$ rounds. To this end, for each $I' \subseteq I - I^* - \{o\}$ with $|I'| = k-k'+1$, we construct the following maximum flow problem $F_{I'}$, which can be solved in polynomial-time by e.g. the Ford-Fulkerson algorithm.

- **Vertices:** s, t, $A - \{a_1\}$, $I - I' - I^*$.
- **Edges and weights:** For each $a \in A - \{a_1\}$, there is an edge $s \to a$ with weight k; for each $a \in A - \{a_1\}$ and $c \in I - I' - I^*$ such that agent a ranks c above all items in I', there is an edge $a \to c$ with weight 1; for each $c \in I - I' - I^*$, there is an edge $c \to t$ with weight 1.
- **We are asked** whether the maximum amount of flow from s to t is $k(n-1)$ (the maximum possible flow from s to t).

Claim 2. *(A, I, P, a_1, o) is a "No" instance if and only if there exists $I' \subseteq I - I^* - \{o\}$ with $|I'| = k-k'+1$ such that $F_{I'}$ has a solution.*

Because k is a constant, the number of I' we will check is a constant. Algorithm 1 is a polynomial algorithm for NECESSARYITEM with balanced policies. □

Algorithm 1: NecessaryItem for balanced policies.

```
Input: A NecessaryItem instance (A, I, P, a_j, o).
if o is ranked below the k-th position by agent a_j then
    return "No".
end
Let I* denote agent a_j's top k' − 1 items.
for I' ⊆ I − I* − {o} with |I'| = k − k' + 1 do
    if F_{|I'|} has a solution then
        return "No"
    end
end
return "Yes".
```

Theorem 11. NecessarySet *and top-k* NecessarySet *for balanced policies are in P even when k is not fixed.*

Proof. Given an instance of NECESSARYSET, if the target set is not top-k then the answer is "No" because we can simply let the agent choose k items in the first k rounds. It remains to show that top-k NECESSARYSET for balanced policies is in P. That is, given (A, I, P, a_1), we can check in polynomial time whether there is a balanced policy π for which agent a_1 does not get exactly her top k items.

For NECESSARYSET, suppose agent a_1 does not get her top-k items under π. Let π' denote the order obtained from π by moving all agent a_1's turns to the end while keeping the other orders unchanged. It is easy to see that agent a_1 does not get her top-k items under π' either. Therefore, NECESSARYSET is equivalent to checking whether there exists an order π where agent a_1 picks item in the last k rounds so that agent a_1 does not get at least one of her top-k items.

We consider an equivalent, reduced allocation instance where the agents are $\{a_1, a_2, \ldots, a_n\}$, and there are $k(n-1)+1$ items $I' = (I - I^*) \cup \{c\}$, where I^* is agent a_1's top-k items. Agent a_j's preferences over I' are obtained from P_j by replacing the first occurrence of items in I^* by c, and then removing all items in I^* while keeping the order of other items the same. We are asked whether there exists an order π where agent a_1 is the last to pick and a_1 picks a single item, and each other agents picks k times, so that agent a_1 does not get item c. This problem can be solved by a polynomial-time algorithm based on maximum flows that is similar to the algorithm for NECESSARYITEM for balanced policies in Theorem 10. □

5 Recursively Balanced Policies

From Theorem 1, we get the following corollary.

Corollary 2. POSSIBLEASSIGNMENT *for recursively balanced policies is in P.*

We also report computational results for problems other than POSSIBLEASSIGNMENT. The following algorithm works via a greedy approach.

Theorem 12. NECESSARYASSIGNMENT *for recursively balanced policies is in P.*

The other 'necessary problems' turn out to be computationally intractable.

Theorem 13. *For $k \geq 2$,* NECESSARYITEM*,* NECESSARYSET*, top-k* NECESSARYSET*, and* NECESSARYSUBSET *for recursively balanced policies are coNP-complete.*

On the other hand, Top-2 POSSIBLESET is easy via a reduction to maximum matching.

Theorem 14. *Top-k* POSSIBLESET *for recursively balanced policies is in P for $k = 2$.*

Finally, top-k-POSSIBLESET is NP-complete iff $k \geq 3$.

Theorem 15. *For all $k \geq 3$, top-k* POSSIBLESET *for balanced policies is NP-complete.*

6 Strict Alternation Policies

Since there are $n!$ possible strict alternation policies, so if n is constant, then all problems can be solved in polynomial time by brute force search. Note that such an argument does not apply to recursively balanced policies. As a result of our characterization of strict alternation outcomes (Theorem 3), we get the following.

Corollary 3. POSSIBLEASSIGNMENT *for strict alternation polices is in P.*

We also present other computational results.

Theorem 16. NECESSARYASSIGNMENT *for strict alternation polices is in P.*

Theorem 17. *Top-k* POSSIBLESET *for strict alternation policies is in P for $k = 2$.*

For Theorem 17, the polynomial-time algorithm is similar to the algorithm for Theorem 14. The next theorems state that the remaining problems are hard to compute. Both theorems are proved by reductions from POSSIBLEITEM.

Theorem 18. *For all $k \geq 3$, top-k* POSSIBLESET *is NP-complete for strict alternation policies.*

Theorem 19. *For all $k \geq 2$,* NECESSARYITEM*,* NECESSARYSET*, top-k* NECESSARYSET*, and* NECESSARYSUBSET *are coNP-complete for strict alternation policies.*

7 Balanced Alternation Policies

If n is constant, then all problems can be solved in polynomial time by brute force search. As a result of our characterization of balanced alternation outcomes (Theorem 2), we get the following.

Corollary 4. POSSIBLEASSIGNMENT *for balanced alternation polices is in P.*

NECESSARYASSIGNMENT can be solved efficiently as well.

Theorem 20. NECESSARYASSIGNMENT *for balanced alternation polices is in P.*

We already know that for $k = m/n = 1$, top-k possible and necessary problems can be solved in polynomial time. The next theorems state that for any other k, they are NP-complete for balanced alternation policies. Theorem 21 is proved by a reduction from the EXACT 3-COVER problem and Theorem 22 is proved by a reduction from the POSSIBLEITEM problem.

Theorem 21. *For all $k \geq 2$, top-k* POSSIBLESET *is NP-complete,* NECESSARYITEM *is coNP-complete, and* NECESSARYSUBSET *is coNP-complete for balanced alternation policies.*

Theorem 22. *For all $k \geq 2$, top-k* NECESSARYSET *for balanced alternation policies is coNP-complete.*

8 Conclusions

We have studied sequential allocation mechanisms where the policy has not been fixed or has been fixed but not announced. We have characterized the allocations achievable with common classes of policies.We have also identified the computational complexity of identifying the possible or necessary items, set or subset of items to be allocated to an agent when using one of the policy classes. There are interesting future directions including considering other common classes of policies, as well as other properties of the outcome like the possible or necessary welfare.

References

A. Abdulkadiroğlu and T. Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–702, 1998.

D. J. Abraham, K. Cechlárová, D. Manlove, and K. Mehlhorn. Pareto optimality in house allocation problems. In *Proc. of the 16th International Symposium on Algorithms and Computation (ISAAC)*, volume 3341 of *LNCS*, pages 1163–1175, 2005.

H. Aziz. A note on the undercut procedure. In *Proc. of the 13th AAMAS Conference*, pages 1361–1362, 2014.

H. Aziz, M. Brill, F. Fischer, P. Harrenstein, J. Lang, and H. G. Seedig. Possible and necessary winners of partial tournaments. In *Proc. of the 11th AAMAS Conference*, pages 585–592. IFAAMAS, 2012.

H. Aziz, F. Brandt, and M. Brill. The computational complexity of random serial dictatorship. *Economics Letters*, 121(3):341–345, 2013.

H. Aziz, S. Gaspers, S. Mackenzie, N. Mattei, P. Stursberg, and T. Walsh. Fixing a balanced knockout tournament. In *Proc. of the 28th AAAI Conference*, pages 552–558, 2014a.

H. Aziz, S. Gaspers, S. Mackenzie, and T. Walsh. Fair assignment of indivisible objects under ordinal preferences. In *Proc. of the 13th AAMAS Conference*, pages 1305–1312, 2014b.

Y. Bachrach, N. Betzler, and P. Faliszewski. Probabilistic possible winner determination. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 697–702, 2010.

D. Baumeister and J. Rothe. Taking the final step to a full dichotomy of the possible winner problem in pure scoring rules. In *Proceedings of The 19th European Conference on Artificial Intelligence (ECAI)*, 2010.

N. Betzler and B. Dorn. Towards a dichotomy for the possible winner problem in elections based on scoring rules. *Journal of Computer and System Sciences*, 76(8):812–836, 2010.

S. Bouveret and J. Lang. Efficiency and envy-freeness in fair division of indivisible goods: logical representation and complexity. *Journal of Artificial Intelligence Research*, 32 (1):525–564, 2008.

S. Bouveret and J. Lang. A general elicitation-free protocol for allocating indivisible goods. In *Proc. of the 22 IJCAI*, pages 73–78, 2011.

S. Bouveret and J. Lang. Manipulating picking sequences. In *In Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 141–146, 2014.

S. Bouveret, U. Endriss, and J. Lang. Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *Proc. of the 19th European Conference on Artificial Intelligence (ECAI)*, pages 387–392, 2010.

S. J. Brams and D. L. King. Efficient fair division: Help the worst off or avoid envy? *Rationality and Society*, 17(4): 387–421, 2005.

S. J. Brams and P. D. Straffin. Prisoners' dilemma and professional sports drafts. *The American Mathematical Monthly*, 86(2):80–88, 1979.

S. J. Brams and A. D. Taylor. *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge University Press, 1996.

E. Budish and E. Cantillion. The multi-unit assignment problem: Theory and evidence from course allocation at Harvard. *American Economic Review*, 102(5):2237–2271, 2012.

Y. Chevaleyre, P. E. Dunne, U. Endriss, J. Lang, M. Lemaître, N. Maudet, J. Padget, S. Phelps, J. A. Rodríguez-Aguilar, and P. Sousa. Issues in multiagent resource allocation. *Informatica*, 30:3–31, 2006.

G. Erdélyi and E. Elkind. Manipulation under voting rule uncertainty. In *Proc. of the 11th AAMAS Conference*, pages 627–634, 2012.

T. Kalinowski, N. Narodytska, and T. Walsh. A social welfare optimal sequential allocation procedure. In *Proc. of the 22nd IJCAI*, pages 227–233, 2013a.

T. Kalinowski, N. Narodytska, T. Walsh, and L. Xia. Strategic behavior when allocating indivisible goods sequentially. In *Proc. of the 27th AAAI Conference*, pages 452–458, 2013b.

D. A. Kohler and R. Chandrasekaran. A class of sequential games. *Operations Research*, 19(2):270–277, 1971.

K. Konczak and J. Lang. Voting procedures with incomplete preferences. In *Multidisciplinary Workshop on Advances in Preference Handling*, 2005.

L. Levine and K. E. Stange. How to make the most of a shared meal: Plan the last bite first. *The American Mathematical Monthly*, 119(7):550–565, 2012.

D. Saban and J. Sethuraman. The complexity of computing the random priority allocation matrix. In Y. Chen and N. Immorlica, editors, *Proc. of the 9th WINE*, LNCS, 2013.

L-G Svensson. Strategy-proof allocation of indivisible goods. *Social Choice and Welfare*, 16(4):557–567, 1999.

T. Vu, A. Altman, and Y. Shoham. On the complexity of schedule control problems for knockout tournaments. In *Proc. of the 8th AAMAS Conference*, pages 225–232, 2009.

L. Xia and V. Conitzer. Determining possible and necessary winners under common voting rules given partial orders. *JAIR*, 41(2):25–67, 2011.

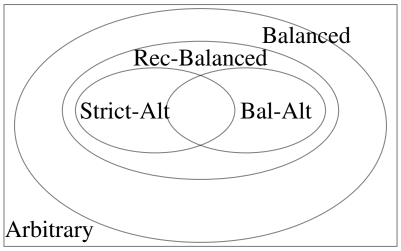


Figure 1: Inclusion relationships between sets of policies. We use abbreviations Rec-Balanced (recursively balanced); Strict-Alt (strict alternation), and Bal-Alt (balanced alternation).

Relation between policy classes

In Figure 1.

Proof of Theorem 6

Proof. Let the target allocation of agent a_i be S. If there is any agent $a_j \in A \setminus \{a_i\}$ who wants to pick an item $o' \in I \setminus S$, let him pick it. If no agent in $A \setminus \{i\}$ wants to pick such an item $o' \in I \setminus S$, and i does not want to pick an item from S return no. If no agent in $A \setminus \{a_i\}$ wants to pick such an item $o' \in I \setminus S$, and a_i wants to pick an item $o \in S$, let a_i pick o. If some agents in $A \setminus \{a_i\}$ wants to pick such an item $o \in S$, and also i wants to pick $o \in S$, then we let a_i pick o. Repeat the process until all the items are allocated or we return no at some point.

We now argue for the correctness of the algorithm. Observe the order in which agent a_1 picks items in S is exactly according to his preferences.

Claim 3. *Let us consider the first pick in the algorithm. If agent a_1 picks an item $o = \max_{\succsim_i}(S)$, then if there exists a policy π in which agent a_i gets S, then there also exists a policy π' in which agent a_1 first picks o and agent i gets S overall.*

Proof. In π, by the time agent a_i picks his second most preferred item from S, all items more preferred have already been allocated. In π, if $a_i \neq \pi(1)$, then we can obtain π' by bringing a_i to the first place and having all the other turns in the same order. Note that in π', for any agent's turn the set of available items are either the same or o is the extra item missing. However since o was not even chosen by the latter agents, the picking outcomes of π and π' are identical. □

Claim 4. *Let us consider the first pick in the algorithm. If some agent a_j picks an item $o' \in A \setminus S$ in the algorithm, then if there exists a policy in which agent a_i gets S, then there also exists a policy in which agent a_j first picks o' and agent a_i gets S overall.*

Proof. In π, if $a_j \neq \pi(1)$, then we can obtain π' by bringing a_j to the first place and having all the other turns in the same order. If j does not get o' in π, then when we construct π' we simply delete the turn of the agent who got o'. Note that in π', for any agent's turn the set of available items are either the same or o' is the extra item missing. However since o' was not even chosen by the latter agents, the picking outcomes of π and π' are identical. □

By inductively applying Claims 3 and 4, we know that as long as a policy exists in which i gets allocation S, our algorithm can construct a policy in which i gets allocation S. □

Proof of Theorem 10

Proof. In a NECESSARYITEM instance we can assume the distinguished agent is a_1. Given (A, I, P, a_1, o), if o is ranked below the k-th position by agent a_1 then it we can return "No", because by letting agent a_1 choose in the first k rounds she does not get item o.

Suppose o is ranked at the k'-th position by agent a_1 with $k' \leq k$, the next claim provides an equivalence condition to check whether the NECESSARYITEM instance is a "No" instance.

Claim 5. *Suppose o is ranked at the k'-th position by agent a_1 with $k' \leq k$, the NECESSARYITEM instance (A, I, P, a_1, o) is a "No" instance if and only if there exists a balanced policy π such that (i) agent a_1 picks items in the first $k' - 1$ rounds and the last $k - k' + 1$ rounds, and (ii) agent a_1 does not get o.*

Proof. Suppose there exists a balanced policy π' such that agent a_1 does not get item o, then we obtain π^* from π' by moving the first $k' - 1$ occurrences of agent a_1 to the beginning of the sequence while keeping other positions unchanged. When preforming π^*, in the first $k' - 1$ rounds agent a_1 gets her top $k' - 1$ items.

By the next time agent a_1 picks an item in π^*, o must have been chosen by another agent. To see why this is true, for each agent from the k'-th round until agent a_1's next turn in π^*, we compare side by side the items allocated before this agent's turn by π^* and by π'. It is not hard to see by induction that the item allocated by π^* before agent a_1's next turn is a superset of the item allocated by π' before agent a_1's k'-th turn. Because the latter contains o, agent a_1 does not get o in π^*.

Then, we obtain π from π^* by moving the k'-th through the k-th occurrence of agent a_1 to the end of the sequence while keeping other positions unchanged. It is easy to see that agent a_1 does not get o in π. This completes the proof. □

Let I^* denote agent a_1's top $k' - 1$ items. In light of the claim above, to check whether the (A, I, P, a_1, o) is a "No" instance, it suffices to check for every set of $k - k' + 1$ items ranked below the k'-th position by agent a_1, denoted by I', whether it is possible for agent a_1 to get I^* and I' by a balanced policy where agent a_1 picks items in the first $k' - 1$ rounds and the last $k - k' + 1$ rounds. To this end, for each $I' \subseteq I - I^* - \{o\}$ with $|I'| = k - k' + 1$, we construct the following maximum flow problem $F_{I'}$, which can be solved in polynomial-time by e.g. the Ford-Fulkerson algorithm.

- **Vertices:** $s, t, A - \{a_1\}, I - I' - I^*$.
- **Edges and weights:** For each $a \in A - \{a_1\}$, there is an edge $s \to a$ with weight k; for each $a \in A - \{a_1\}$ and $c \in I - I' - I^*$ such that agent a ranks c above all items in I', there is an edge $a \to c$ with weight 1; for each $c \in I - I' - I^*$, there is an edge $c \to t$ with weight 1.
- **We are asked** whether the maximum amount of flow s to t is $k(n-1)$ (the maximum possible flow from s to t).

Claim 6. *(A, I, P, o) is a "No" instance if and only if there exists $I' \subseteq I - I^* - \{o\}$ with $|I'| = k - k' + 1$ such that $F_{I'}$ has a solution.*

Proof. If (A, I, P, o) is a "No" instance, then by Claim 1 there exists π such that agent a_1 picks items in the first $k' - 1$ rounds and the last $k - k' + 1$ rounds, and agent a_1 gets $I^* \cup I'$ for some $I' \subseteq I - I^* - \{o\}$. For each agent a_j with $j \neq 2$, let there be a flow of amount k from s to a_j and a flow of amount 1 from a_j to all items that are allocated to her in π. Moreover, let there be a flow of amount 1 from any $c \in I - I^* - \{o\}$ to t. It is easy to check that the amount of flow is $k(n-1)$.

If $F_{I'}$ has a solution, then there exists an integer solution because all weights are integers. This means that there exists an assignment of all items in $I - I' - I^*$ to agent 2 through n such that no agent gets an item that is ranked below any item in I^*. Starting from this allocation, after implementing all trading cycles we obtain a Pareto optimal allocation where $I - I' - I^*$ are allocated to agent 2 through n, and still no agent gets an item that is ranked below any item in I^*. By Proposition 1 in Brams and King, there exists a balanced policy π^* that gives this allocation. It follows that agent a_1 does not get o under the balanced policy $\pi = \underbrace{a_1 \rhd \ldots \rhd a_1}_{k'-1} \rhd \pi^* \rhd \underbrace{a_1 \rhd \ldots \rhd a_1}_{k-k'+1}$.

□

Because k is a constant, the number of I' we will check is a constant. The polynomial algorithm for NECESSARYITEM for balanced policies is presented as Algorithm 1. □

Proof of Theorem 12

Proof. In the allocation p, let p_i^j be the j-th most preferred item for agent i among his set of k allocated items.

Claim 7. *If there exists a recursively balanced policy achieving the target allocation. Then, in any such recursively balanced policy, we know that in each t-th round, each agent gets item p_i^t.*

We initialize t to 1 i.e., focus on the first round. We check if there is an agent whose turn has not come in the round whose most preferred unallocated item is not p_i^t. In this case return "no". Otherwise, we complete the round in any arbitrary order. If all the items are allocated, we return "yes". If $t \neq k$, we increment t by one and repeat the process.

We now argue for correctness. If the algorithm returns no, then we know that there is a recursively balanced policy that does not achieve the allocation. This policy was partially built during the algorithm and can be completed in an arbitrary way to get an allocation that is not the same as the target allocation. Now assume for contradiction that there is a policy which does not achieve the allocation but the algorithm incorrectly returns yes. Consider the first round where the algorithm makes a mistake. But in each round, each agent had a unique and mutually exclusive most preferred unallocated item. Hence no matter which policy we implement in the round, the allocation and the set of unallocated items after the round stays the same. Hence a contradiction. □

Proof of Theorem 13

Proof Sketch. Membership in coNP is obvious. By Lemma 1 it suffices to show coNP-hardness for NECESSARYITEM and top-k NECESSARYSET. We will prove the co-NP-hardness for them for $k = 2$ by the same reduction from POSSIBLEITEM for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. The proof for other $k \geq 2$ can be done similarly by constructing preferences so that the distinguished agent always get her top $k - 2$ items. Let (A, I, P, a_1, o) denote an instance of POSSIBLEITEM for $k = 1$, where $A = \{a_1, \ldots, a_n\}$, $I = \{o_1, \ldots, o_n\}$, $o \in I$, $P = (P_1, \ldots, P_n)$ is the preference profile of the n agents, and we are asked wether it is possible for agent a_1 to get item o in some sequential allocation. Given (A, I, P, a_1, o), we construct the following necessary allocation instance.

Agents: $A' = A \cup \{a_{n+1}\}$.
Items: $I' = I \cup \{c, d\} \cup D$, where $|D| = n + 1$.
Preferences:

- The preferences of a_1 is obtained from P_1 by inserting d right before o, and append the other items such that the bottom item is c.
- For any $2 \leq j \leq n$, the preferences of a_j is obtained from P_j by replacing o by D and then appending the remaining items such that the bottom items are $c \succ d \succ o$.
- The preferences for a_{n+1} is $[c \succ o \succ \text{others} \succ d]$.

For NECESSARYITEM, we are asked whether agent a_{n+1} always get item o; for top-k NECESSARYSET, we are asked whether agent a_{n+1} always get $\{c, o\}$, which are her top-2 items.

Suppose the (A, I, P, a_1, o) has a solution, denoted by π. We claim that $\pi' = a_{n+1} \rhd \pi \rhd a_1 \rhd (A' - \{a_1\})$ is a "No" answer to the NECESSARYITEM and top-k NECESSARYSET instance. Following π', in the first round a_{n+1} gets c. In the next n rounds a_1 gets d. Then in the $(n+2)$-th round agent a_1 gets item o, which means that a_{n+1} does not get item o after all items are allocated.

We note that a_{n+1} always get item c for any recursively balanced policy. We next show that if NECESSARYITEM or top-k NECESSARYSET instance is a "No" instance, then the POSSIBLEITEM instance is a "Yes" instance. Suppose π' is a recursively balanced policy such that a_{n+1} does not get o. We let **phase 1** denote the first $n + 1$ rounds, and let **phase 2** denote the $(n + 2)$-th through $2(n + 1)$-th round.

Because o is the least preferred item for all agents except a_1 and a_{n+1}, if a_{n+1} does not get o in the second phase, then o must be allocated to a_1. This is because for the sake of contradiction suppose o is allocated to agent a_j with $j \neq 1, n$, then a_j must be the last agent in π' and o is not chosen in any previous round. However, when it is a_n's turn in the second phase, o is still available, which means that a_n would have chosen o and contradicts the assumption that a_j gets o.

Claim 8. *If a_1 gets o under π', then a_1 gets d in the first phase.*

Proof. For the sake of contradiction, suppose in the first phase a_1 does not get d, then either she gets an item before d, or she gets o, because it is impossible for a_1 to get an item after o otherwise another agent must get o in the first phase, which is impossible as we just argued above.

- If a_1 gets an item before d in the first phase, then in order for a_1 to get o in the second phase, d must be chosen by another agent. Clearly d cannot be chosen by a_{n+1} before a_1 gets o, because d is the bottom item by a_{n+1}, which means that the only possibility for a_{n+1} to get d is that a_{n+1} is the last agent in π'. If d is chosen by a_j with $j \leq n$, then because d, o are the bottom two items by a_j, the last two agents in π' must be $a_j \rhd a_1$. Therefore, when a_{n+1} chooses an item in the second phase, o is still available, which means that a_{n+1} gets o in π', a contradiction to the assumption that a_{n+1} does not get her top-2 items.
- If a_1 gets o in the first phase, then it means that another agent must get d in the first phase, which is impossible because all other agents rank d within their bottom two positions, which means that the earliest round that any of them can get d is $2n+1$.

□

Let π denote the order over A that is obtained from the first phase of π' by removing a_{n+1}, and them moving all agents who get an item in D after a_1. We claim that π is a solution to (A, I, P, a_1, o), because when it is a_1's round all items before o must be chosen and o has not been chosen (if another agent gets o before a_1 in π then the same agent must get an item in D in the first phase of π', which contradicts the construction of π). This proves the co-NP-completeness of the allocation problems mentioned in the theorem. □

Proof of Theorem 14

Proof. We give agent a_1 the first turns in each round. He is guaranteed to get s_1. We now construct a bipartite graph $G = ((A \setminus \{a_1\}) \cup (I \setminus \{s_1\}), E)$ in which each $\{i, o\} \in E$ iff o is strictly more preferred for i than s_2. We check whether G admits a perfect matching. If G does not admit a perfect matching, we return no. Otherwise, there exists a recursively balanced policy for which agent a_1 gets s_1 and s_2.

Claim 9. *G admits a perfect matching if and only if there a recursively balanced policy for which a_1 gets $\{s_1, s_2\}$.*

Proof. If G admits a perfect matching, then each agent in $A \setminus \{a_1\}$ can get a more preferred item than s_2 in the first round. If this particular allocation is not Pareto optimal for agents in $A \setminus \{a_1\}$ for items among $I \setminus \{s_1\}$, we can easily compute a Pareto optimal Pareto improvement over this allocation by implementing trading cycles as in setting of house allocation with existing tenants. This takes at most $O(n^3)$. Hence, we can compute an allocation in which each agent in $A \setminus \{a_1\}$ gets a strictly more preferred item than s_2 and this allocation for agents in $A \setminus \{a_1\}$ is Pareto optimal. Since the allocation is Pareto optimal, we can easily build up a policy which achieves this Pareto optimal allocation via the characterization of Brams. In the second round, a_1 gets s_2 and then subsequently we don't care who gets what because agent a_1 has already got s_1 and s_2.

If G does not admit a perfect matching, then there is no allocation in which each agent in $A \setminus \{a_1\}$ get a strictly better item than s_2 in $I \setminus \{s_1\}$. Hence in each policy in the first round, some agent in $A \setminus \{a_1\}$ will get s_2. □

□

Proof of Theorem 15

Proof. Membership in NP is obvious. We prove that top-k PossibleSet for $k = 3$ is NP-hard by a reduction from PossibleItem for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. Hardness for other k's can be proved similarly by constructing preferences so that the distinguished agent always get her top $k - 2$ items. Let (A, I, P, a_1, o) denote an instance of PossibleItem for $k = 1$, where $A = \{a_1, \dots, a_n\}$, $I = \{o_1, \dots, o_n\}$, $o \in I$, $P = (P_1, \dots, P_n)$ is the preference profile of the n agents, and we are asked wether it is possible for agent a_1 to get item o in some sequential allocation. Given (A, I, P, a_1, o), we construct the following PossibleSet instance.

Agents: $A' = A \cup \{a_{n+1}\} \cup \{d_1, \dots, d_{n-1}\}$.

Items: $I' = I \cup \{c_1, c_2, c_3\} \cup D \cup E \cup F$, where $|D| = |E| = n - 1$ and $|F| = 3n - 1$. We have $|I'| = 6n$.

Preferences:

- The preferences of a_1 is $[P_1 \succ \text{others} \succ c_1 \succ c_2 \succ c_3]$.
- For any $2 \leq j \leq n$, the preferences of a_j is obtained from $[P_j \succ \text{others} \succ c_1 \succ c_2 \succ c_3 \succ E]$ by switching o and E.
- The preferences for a_{n+1} is $[c_1 \succ c_2 \succ c_3 \succ \text{others}]$.
- For all $j \leq n-1$, the preferences for d_j is $[D \succ ((I - \{o\}) \cup E) \succ c_3 \succ c_2 \succ c_1 \succ \text{others}]$.

We are asked whether agent a_{n+1} can get items $\{c_1, c_2, c_3\}$, which are her top-3 items.

If (A, I, P, a_1, o) has a solution π, we show that the top-3 PossibleSet instance is a "Yes" instance by considering $\pi' = \underbrace{a_{n+1} \rhd d_1 \rhd \cdots \rhd d_{n-1} \rhd \pi}_{\text{Phase 1}} \rhd \underbrace{a_{n+1} \rhd d_1 \rhd \cdots \rhd d_{n-1} \rhd A}_{\text{Phase 2}} \rhd \underbrace{a_{n+1} \rhd \text{others}}_{\text{Phase 3}}$. In the first phase a_{n+1} gets c_1; d_j's get D a_1 gets o and other agents in A get $n - 1$ items in $(I - \{o\}) \cup E$. In the second phase a_{n+1} gets c_2; d_j's get the remaining $n - 1$ items in $(I - \{o\}) \cup E$; agents in A get n items in F. In the third phase a_{n+1} gets c_3.

Suppose the top-3 PossibleSet instance is a "Yes" instance and agent a_{n+1} gets $\{c_1, c_2, c_3\}$ in a recursively balanced policy π'. Let π denote the order over which agents a_1 through n choose items in the first phase of π'. We obtain another order π^* over A from π by moving all agents who choose an item in D after agent a_1 without changing the order of other agents. We claim that π^* is a solution to (A, I, P, a_1, o). For the sake of contradiction suppose π^* is not a solution to (A, I, P, a_1, o). It follows that in the first phase of π' agent a_1 gets an item she ranks higher than o, because no other agents can get o. This means that in the first phase n items in $(I - \{o\}) \cup E$ are chosen by A. We note that in the first phase d_j's must chose items in D. Then in the second phase at least one d_j will choose $\{c_3\}$, because there are $n-1$ of them and only $2(n-1) - n = n - 2$ items available before $\{c_3\}$. This contradicts the assumption that a_{n+1} gets c_3. □

Proof of Theorem 16

Proof. We prove that an assignment M is the outcome of all strict alternating policies iff in each round, each agent has a unique most preferred item from among the unallocated items from the previous round. If in each round, each agent gets the most preferred item from among the unallocated items from the previous round, the order does not matter in any round. Hence all alternating policies result in M.

Now assume that it is not the case that in each round, each agent gets the most preferred item from among the unallocated items from the previous round. Then, there exist at least two agent who have the same most preferred item from among the remaining items. Therefore, a different relative order among such agents results in different allocations which means that M is not the unique outcome of all strict alternating policies. □

Proof of Theorem 18

Proof. Membership in NP is obvious. We prove that top-k PossibleSet for $k = 3$ is NP-hard by a reduction from PossibleItem for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. The reduction is similar to the proof of Theorem 15. Hardness for other k's can be proved similarly by constructing preferences so that the distinguished agent always get her top $k - 2$ items. Let (A, I, P, a_1, o) denote an instance of PossibleItem for $k = 1$, where $A = \{a_1, \dots, a_n\}$, $I = \{o_1, \dots, o_n\}$, $o \in I$, $P = (P_1, \dots, P_n)$ is the preference profile of the n agents, and we are asked wether it is possible for agent a_1 to get item o in some sequential allocation. Given (A, I, P, a_1, o), we construct the following PossibleSet instance.

Agents: $A' = A \cup \{a_{n+1}\} \cup \{d_1, \dots, d_{n-1}\}$.

Items: $I' = I \cup \{c_1, c_2, c_3\} \cup D \cup E \cup F$, where $|D| = |E| = n - 1$ and $|F| = 3n - 1$. We have $|I'| = 6n$.

Preferences:

- The preferences of a_1 is $[P_1 \succ \text{others} \succ c_1 \succ c_2 \succ c_3]$.
- For any $2 \leq j \leq n$, the preferences of a_j is obtained from $[P_j \succ \text{others} \succ c_1 \succ c_2 \succ c_3 \succ E]$ by switching o and E.
- The preferences for a_{n+1} is $[c_1 \succ c_2 \succ c_3 \succ \text{others}]$.
- For all $j \leq n-1$, the preferences for d_j is $[D \succ ((I - \{o\}) \cup E) \succ c_3 \succ c_2 \succ c_1 \succ \text{others}]$.

We are asked whether agent a_{n+1} can get items $\{c_1, c_2, c_3\}$, which are her top-3 items.

If (A, I, P, a_1, o) has a solution π, we show that the top-3 PossibleSet instance is a "Yes" instance by considering $\pi' =$

$\underbrace{a_{n+1} \rhd d_1 \rhd \cdots \rhd d_{n-1} \rhd \pi}_{\text{Phase 1}} \rhd \underbrace{a_{n+1} \rhd d_1 \rhd \cdots \rhd d_{n-1} \rhd \pi}_{\text{Phase 2}} \rhd$ $\underbrace{a_{n+1} \rhd d_1 \rhd \cdots \rhd d_{n-1} \rhd \pi}_{\text{Phase 3}}$. In the first phase a_{n+1} gets c_1, a_1 gets o; other agents in A get $n-1$ items in $(I - \{o\}) \cup E$; d_j's get D. In the second phase a_{n+1} gets c_2; d_j's get the remaining $n-1$ items in $(I - \{o\}) \cup E$; agents in A get n items in F. In the third phase a_{n+1} gets c_3.

Suppose the top-3 PossibleSet instance is a "Yes" instance and agent a_{n+1} gets $\{c_1, c_2, c_3\}$ in a strict alternation policy π'. Let π denote the order over which agents a_1 through n choose items in the first phase of π'. We obtain another order π^* over A from π by moving all agents who choose an item in D after agent a_1 without changing the order of other agents. We claim that π^* is a solution to (A, I, P, a_1, o). For the sake of contradiction suppose π^* is not a solution to (A, I, P, a_1, o). It follows that in the first phase of π' agent a_1 gets an item she ranks higher than o, because no other agents can get o. This means that in the first phase n items in $(I - \{o\}) \cup E$ are chosen by A. We note that in the first phase d_j's must chose items in D. Then in the second phase at least one d_j will choose $\{c_3\}$, because there are $n-1$ of them and only $2(n-1)-n = n-2$ items available before $\{c_3\}$. This contradicts the assumption that a_{n+1} gets c_3. □

Proof of Theorem 19

Proof Sketch. The proof is similar to the proof of Theorem 13. Membership in coNP is obvious. By Lemma 1 it suffices to show coNP-hardness for NecessaryItem and top-k NecessarySet. We will prove the co-NP-hardness for them for $k = 2$ by the same reduction from PossibleItem for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. The proof for other $k \geq 2$ can be done similarly by constructing preferences so that the distinguished agent always get her top $k-2$ items. Let (A, I, P, a_1, o) denote an instance of PossibleItem for $k = 1$, where $A = \{a_1, \ldots, a_n\}$, $I = \{o_1, \ldots, o_n\}$, $o \in I$, $P = (P_1, \ldots, P_n)$ is the preference profile of the n agents, and we are asked wether it is possible for agent a_1 to get item o by some strict alternation policy. Given (A, I, P, a_1, o), we construct the following necessary allocation instance.

Agents: $A' = A \cup \{a_{n+1}\}$.
Items: $I' = I \cup \{c, d\} \cup D$, where $|D| = n+1$.
Preferences:

- The preferences of a_1 is obtained from P_1 by inserting d right before o, and append the other items such that the bottom item is c.
- For any $2 \leq j \leq n$, the preferences of a_j is obtained from P_j by replacing o by D and then appending the remaining items such that the bottom items are $c \succ d \succ o$.
- The preferences for a_{n+1} is $[c \succ o \succ \text{others} \succ d]$.

For NecessaryItem, we are asked whether agent a_{n+1} always get item o; for top-k NecessarySet, we are asked whether agent a_{n+1} always get $\{c, o\}$, which are her top-2 items.

Suppose the (A, I, P, a_1, o) has a solution, denoted by π. We claim that $\pi' = \underbrace{\pi \rhd a_{n+1}}_{\text{Phase 1}} \rhd \underbrace{\pi \rhd a_{n+1}}_{\text{Phase 2}}$ is a "No" answer to the NecessaryItem and top-k NecessarySet instance. Following π', in phase a_1 gets d gets d and a_{n+1} gets c. In phase 2 a_1 gets o, which means that a_{n+1} does not get item o after all items are allocated.

We next show that if NecessaryItem or top-k NecessarySet instance is a "No" instance, then the PossibleItem instance is a "Yes" instance. We note that a_{n+1} always get item c in the first phase of any strict alternation policy. Let π' denote a strict alternation policy where a_{n+1} does not get o. If a_1 does not get d in the first phase, then following a similar argument in the proof of Theorem 13, we have that a_{n+1} gets o in the second phase, which is a contradiction. Therefore, a_1 must get d in the first phase.

Let π denote the order over A that is obtained from the first phase of π' by removing a_{n+1}, and them moving all agents who get an item in D after a_1. We claim that π is a solution to (A, I, P, a_1, o), because when it is a_1's round all items before o must be chosen and o has not been chosen (if another agent gets o before a_1 in π then the same agent must get an item in D in the first phase of π', which contradicts the construction of π). This proves the co-NP-completeness of the allocation problems mentioned in the theorem. □

Proof of Theorem 21

Proof. Membership in NP and coNP are obvious. By Lemma 1, if NecessaryItem is coNP-hard then NecessarySubset is coNP-hard. We show the NP-hardness of top-k PossibleSet and coNP-hardness of NecessaryItem by the same reduction from exact 3-cover (X3C) for $k = 2$. Hardness for other k can be proved similarly by constructing preferences so that the distinguished agent always get her top $k-2$ items. In an X3C instance $(\mathcal{S}, X)$, we are given $\mathcal{S} = \{S_1, \ldots, S_t\}$ and $X = \{x_1, \ldots, x_q\}$, such that q is a multiple of 3 and for all $j \leq t$, $|S_j| = 3$ and $S_j \subseteq X$; we are asked whether there exists a subset of $q/3$ elements of $\mathcal{S}$ whose union is exactly X.

Given an X3C instance $(\mathcal{S}, X)$, we construct the following agents, items, and preferences.

Agents: $A = \{a\} \cup \bigcup_{j \leq t} \mathcal{S}_j \cup X \cup C$, where $C = \{c_1, \ldots, c_{q/3}\}$ and $\mathcal{S}_j = \{S_j, S_j^{j_1}, S_j^{j_2}, S_j^{j_1}\}$ such that $j \leq t$, j_1, j_2, j_3 are the indices of elements S_j. That is, $S_j = \{x_{j_1}, x_{j_2}, x_{j_3}\}$. We note that $|A| = 4t + 4q/3 + 1$.

Items: $8t + 8q/3 + 2$ items are defined as follows. Let $I = \{a, b, c\} \cup \bigcup_{j \leq t} \mathcal{S}_j \cup D \cup E \cup F$, where $|D| = 8q/3$, $E = q/3$, and $F = 4t - q/3 - 1$. We note that $|I| = 2|A|$. For each $i \leq q$, we let K_i denote the sets in $\mathcal{S}$ that cover x_i. That is, $K_i = \{S \in \mathcal{S} : x_i \in S\}$.

Preferences are illustrated in Table 2.

Agent	Preferences
a:	$a \succ b \succ c \succ$ others
$\forall j, S_j$:	$S_j \succ a \succ D \succ b \succ$ others $\succ c$
$\forall j, s = 1,2,3, S_j^{j_s}$:	$S_j \succ S_j^{j_s} \succ a \succ D \succ b \succ$ others $\succ c$
$\forall i, x_i$:	$K_i \succ b \succ$ others $\succ c$
$\forall k \leq q/3, c_k$:	$a \succ S_1 \succ \ldots \succ S_t \succ E \succ$ others $\succ c$

Table 2: Agents' preferences, where $K_i = \{S \in \mathcal{S} : x_i \in S\}$.

For top-2 PossibleSet, we are asked whether agent a can get $\{a, b\}$. For NecessaryItem, we are asked whether agent a always get item c.

If the X3C instance has a solution, w.l.o.g. $\{S_1, \ldots, S_{q/3}\}$, we show that there exists a solution to the constructive control problem and destructive control problem described above. For each $j \leq t$, we let $L_j = S_j \rhd S_j^{j_1} \rhd S_j^{j_2} \rhd S_j^{j_3}$. Let the order π over agents be the following.

$$\pi = L_{q/3+1} \rhd L_{q/3+2} \rhd \cdots \rhd L_t \rhd X \rhd a \rhd C \rhd L_1 \rhd \cdots \rhd L_{q/3}$$

The balanced alternation policy is thus $\pi \rhd \text{inv}(\pi)$, where $\text{inv}(\pi)$ is the inverse order of π. It is not hard to verify that in the first round the allocation w.r.t. π is as follows:

- for each $j \geq q/3+1$, agent S_j gets item S_j and agent $S_j^{j_s}$ gets item $S_j^{j_s}$;
- for each $i \leq q$, agent x_i get S_j^i for the (only) $j \leq q/3$ such that $x_i \in S_j$;
- agent a gets item a;
- for each $k \leq q/3$, agent c_k gets item S_k;
- for each $j \leq q/3$ and $s = 1, 2, 3$, agent S_j gets an item in D and agent $S_j^{j_s}$ gets an item in D.

In the second round, the allocation w.r.t. $\text{inv}(\pi)$ is as follows:

- for each $j \leq q/3$ and $s = 1, 2, 3$, agent S_j gets an item in D and agent $S_j^{j_s}$ gets an item in D; all items in D ($|D| = 8q/3$) are allocated;
- for each $k \leq q/3$, agent c_k gets an item in E; all items in E are allocated ($|E| = q/3$).
- agent a gets item b;
- other agents get the remaining items.

Specifically, agent a gets $\{a, b\}$.

Now suppose the constructive control has a solution, namely there exists an order π over A such that in the sequential allocation w.r.t. $\pi \rhd \text{inv}(\pi)$ agent a gets $\{a, b\}$. We next show that the X3C instance has a solution. For convenience, we divide the sequential allocation of $\pi \rhd \text{inv}(\pi)$ into three stages:

- **Stage 1:** turns before agent a's first turn, where each agent ranked before agent a in π chooses an item;
- **Stage 2:** turns between agent a's first turn and agent a's second turn, where each agent ranked after agent a in π chooses two items;
- **Stage 3:** turns after agent a's second turn, where each agent ranked before agent a in π chooses an item.

Claim 10. *Agents in C must be after agent a in π, and they get at least $q/3$ items in $\mathcal{S}$.*

Proof. Because any agent in C ranks item a at their top, all of them must be after agent a in π. We note that $|C| = q/3$, $|E| = q/3$, and each agent in C will choose two items before agent a's second turn. Therefore, agents in C must get at least $q/3$ items in $\mathcal{S}$, otherwise one of them will choose b, which contradicts the assumption that agent a gets b. ☐

W.l.o.g. let $\{S_1, \ldots, S_{q'}\}$ (for some $q' \geq q/3$) be the items in $\mathcal{S}$ that are chosen by agents in C.

Claim 11. *$q' = q/3$. For all $j \leq q/3$, agents in $\mathcal{S}_j$ are ranked after agent a in π, and for all $j \geq q/3+1$, agents in $\mathcal{S}_j$ are ranked before agent a in π.*

Proof. Let $K = \bigcup_{j \leq t} \mathcal{S}_j \cup D$ denote the set of $4t + 8q/3$ items. The crucial observation is that for any agent $s \in \bigcup_{j \leq t} \mathcal{S}_j$, if s is ranked before a in π, then in the sequential allocation she will get at least one item in K, because she picks an item in K in Stage 1, and maybe another item in K in Stage 3; and if s is ranked after a in π, then in the sequential allocation she will get exactly two items in K in Stage 2. Moreover, each agent in X must get at least one item in K and agents in C must get at least $q/3$ items in K. Therefore, agents in $\bigcup_{j \leq t} \mathcal{S}_j$ get no more than $4t + 4q/3$ items in K. Because $|\bigcup_{j \leq t} \mathcal{S}_j| = 4t$, at most $4q/3$ of these agents are ranked after a in π.

On the other hand, for all $j \leq q'$, agents in $\mathcal{S}_j$ must be ranked after all agents in C in π, otherwise some item S_j would have been allocated to an agent in $\mathcal{S}_j$ (because all of them rank item S_j at the top). By Claim 10 all agents in C must be ranked after agent a in π, which means that for all $j \leq q'$, all agents in $\mathcal{S}_j$ are ranked after agent a in π. Because $q' \geq q/3$, we must have that $q' = q/3$ and for all $j \leq q/3$, agents in $\mathcal{S}_j$ are ranked after agent a in π, and for all $j \geq q/3+1$, agents in $\mathcal{S}_j$ are ranked before agent a in π. ☐

Finally, we are ready to show that $\{S_1, \ldots, S_{q/3}\}$ is an exact cover of X. For the sake of contradiction suppose x_i is not covered. Let S_j^i (with $j > q/3$) denote an item that agent x_i gets in the sequential allocation. Because agents in $\mathcal{S}_j$ are before a in π, it follows that agent S_j^i must get item S_j (because her top-ranked items are S_j, S_j^i, a). However, in this case agent S_j must be allocated item a, which contradicts the assumption that agent a gets item a. Therefore, $\{S_1, \ldots, S_{q/3}\}$ is an exact cover of X. This proves the top-2 POSSIBLESET is NP-complete.

We note that item c is the most undesirable item for all agents except agent a, which means that agent a gets item c if and only if she does not get item a and b. This proves that the NECESSARYITEM is coNP-complete. ☐

Proof of Theorem 22

Proof. Membership in coNP is obvious. We prove that top-k NECESSARYSET for $k = 2$ is coNP-hard by a reduction from POSSIBLEITEM for $k = 1$, which is NP-complete [Saban and Sethuraman, 2013]. Hardness for other k's can be proved similarly by constructing preferences so that the distinguished agent always get her top $k-2$ items. Let (A, I, P, a_1, o) denote an instance of possible allocation problem for $k = 1$, where $A = \{a_1, \ldots, a_n\}$, $I = \{o_1, \ldots, o_n\}$, $o \in I$, $P = (P_1, \ldots, P_n)$, and we are asked wether it is possible for agent a_1 to get item o in some sequential allocation. Given (A, I, P, a_1, o), we construct the following top-2 NECESSARYSET instance.

Agents: $A' = A \cup \{a_{n+1}\}$.

Items: $I' = I \cup \{c_1, c_2\} \cup D$, where $|D| = n$. We have $|I'| = 2n+2$.

Preferences:

- The preferences of a_1 is obtained from P_1 by inserting c_2 right after o, and then append $D \succ c_1$.
- For any $j \leq n$, the preferences of a_j is obtained from $[P_j \succ D \succ c_2 \succ c_1]$ by switching o and D.
- The preferences for a_{n+1} is $[c_1 \succ c_2 \succ \text{others} \succ o]$.

We are asked whether agent a_{n+1} always gets items $\{c_1, c_2\}$, which are her top-2 items.

If (A, I, P, a_1, o) has a solution π, we show that the top-2 NECESSARYSET instance is a "No" instance by considering $\pi' = a_{n+1} \rhd \pi \rhd \pi \rhd a_{n+1}$. In the first phase of π', a_{n+1} gets c_1 and a_1 gets o. In the third phase a_1 gets c_2.

Suppose the top-2 NECESSARYSET instance is a "No" instance and agent a_{n+1} does not get $\{c_1, c_2\}$ in an balanced alternation policy π'. It is easy to see that a_{n+1} must get c_1 in the first phase. Suppose a_1 does not get o in the first phase, then in the beginning of the second phase both o and c_2 are still available. In this case a_{n+1} must get c_2, because clearly none of a_2 through a_n can get c_2, which means that a_1 must get c_2 in the second phase. However, this means that o must be chosen by another agent before, which is impossible since it is ranked in the bottom position after c_1 and c_2 are removed by all other agents. Let π^* denote a linear order over A obtained from the restriction of the first phase of π' on A by moving all agents who choose an item in D after agent a_1 without changing other orders. It is not hard to see that π^* is a solution to (A, I, P, a_1, o). ☐

Improved Algorithmic Results for Unsplittable Stable Allocation Problems

Ágnes Cseh*
Institute for Mathematics, TU Berlin
cseh@math.tu-berlin.de

Brian C. Dean†
School of Computing, Clemson University
bcdean@clemson.edu

February 3, 2015

Abstract

The stable allocation problem is a many-to-many generalization of the well-known stable marriage problem, where we seek a bipartite assignment between, say, jobs (of varying sizes) and machines (of varying capacities) that is "stable" based on a set of underlying preference lists submitted by the jobs and machines. Building on the initial work of [5], we study a natural "unsplittable" variant of this problem, where each assigned job must be fully assigned to a single machine. Such unsplittable bipartite assignment problems generally tend to be NP-hard, including previously-proposed variants of the unsplittable stable allocation problem [12]. Our main result is to show that under an alternative model of stability, the unsplittable stable allocation problem becomes solvable in polynomial time; although this model is less likely to admit feasible solutions than the model proposed in [12], we show that in the event there is no feasible solution, our approach computes a solution of minimal total congestion (overfilling of all machines collectively beyond their capacities). We also describe a technique for rounding the solution of a stable allocation problem to produce "relaxed" unsplit solutions that are only mildly infeasible, where each machine is overcongested by at most a single job.

1 Introduction

Consider a bipartite assignment problem over a graph $G = (V = J \cup M, E)$ involving the assignment of a set of jobs J to a set of machines M. Each job $j \in J$ has a processing time $q(j)$, each machine $m \in M$ has a capacity $q(m)$, and there is a capacity $c(jm)$ for each edge $jm \in E$ governing the maximum amount of job j that can be assigned to machine m. A feasible allocation of jobs to machines is described by a function $x : E \to \mathbb{R}_{\geq 0}$ such that

1. $0 \leq x(jm) \leq c(jm)$ for all edges $jm \in E$,
2. $x(j) := \sum_{m \in M} x(jm) \leq q(j)$ for all jobs $j \in J$, and
3. $x(m) := \sum_{j \in J} x(jm) \leq q(m)$ for all machines $m \in M$.

If $x(jm) \in \{0, q(j)\}$ for all $jm \in E$, we say the allocation is *unsplit*, since each assigned job is assigned in its entirety to a single machine. We often forgo the use of edge capacities $c(jm)$ when discussing unsplit allocations, since an edge jm can simply be deleted if $c(jm) < q(j)$.

*Supported by the Deutsche Telekom Stiftung and the Deutsche Forschungsgemeinschaft within the research training group 'Methods for Discrete Structures' (GRK 1408).

†Supported by USA National Science Foundation CAREER Award CCF-0845593.

Problems of the form above have been extensively studied in the algorithmic literature, where typical objectives are to find a feasible assignment or one of maximum weight (maximizing a linear objective function $\sum_{jm \in E} w(jm)x(jm)$, with $w(jm)$ being the weight of edge jm). While the fractional (splittable) variants of these problems are easy to solve in polynomial time via network flow techniques, it is NP-hard to find an unsplit allocation of either maximum total size $|x| = \sum_{jm \in E} x(jm)$ or of maximum weight; the former is a variant of the multiple subset sum problem [3], and the latter is known as the multiple knapsack problem [4].

In contrast to problems with explicit edge costs, the *stable allocation problem* is an "ordinal" problem variant where the quality of an allocation is expressed in a more game theoretic setting via ranked preference lists submitted by the jobs and machines, with respect to which we seek an assignment that is *stable* (defined shortly). In this paper, we study the stable allocation problem in the unsplittable setting, which was shown to be NP-hard in [12] using one natural definition for stability. We show here that by contrast, a different and more strict notion of stability, proposed initially in [5], leads to an $O(|E|)$ algorithm for the unsplit problem. The tradeoff is that under this different notion of stability, it is unlikely that feasible allocations will exist. However, we show that by relaxing the problem to allow mildly infeasible allocations, our algorithm computes a "relaxed" unsplit stable allocation (in which each machine is filled beyond its capacity by at most the allocation of a single job) in which the total amount of overcongestion across all machines, $\sum_{m \in M} \max(0, x(m) - q(m))$, is minimized (so in particular, if there is a feasible allocation with no congestion, we will find it).

Through the work of several former authors [7, 16, 15], the "relaxed" model has become relatively popular in the context of unsplittable bipartite assignment and unsplittable flow problems. The standard approximation algorithm framework (finding an approximately-optimal, feasible solution) typically does not fit these problems, since finding any feasible solution is typically NP-hard. Instead, authors tend to focus on pseudo-approximation results with minimal congestion per machine or per edge. Analogous results were previously developed for unsplit stable allocation problems in [5], where an unsplit stable allocation can be found in linear time in which each machine is overcongested by at most a single job. The model of stability proposed in [5] is the one we further develop in this paper, and among all of these prior approaches (including those for standard unsplittable bipartite assignment and flows), it seems to be the only unsplit model studied to date in which minimization of *total* congestion is possible in polynomial time. Hence, there is a substantial algorithmic incentive to consider this model, even though its notion of stability is less natural than in [12].

The classical stable marriage problem, perhaps the simplest relative of our problem in the domain of ordinal matching, is known to satisfy a number of remarkable mathematical properties. For example, one can always find stable solutions that are "left-hand-side optimal" or "right-hand-side optimal", and the exact same subset of left-hand side and right-hand side elements are matched in every stable solution (the so-called "rural hospital" theorem, named after applications involving the assignment of medical residents to hospitals). We show natural generalizations of all of these structural properties in our relaxed unsplit stable allocation setting (further justifying the utility of this model from a mathematical perspective). For example we show how to compute in $O(|E|)$ time a "job-optimal" allocation that maximizes the total size $|x|$ of all assigned jobs, and a "machine-optimal" allocation that minimizes $|x|$. It is this machine-optimal solution that we show also minimizes total congestion. In order to produce potentially other solutions (e.g., that might be more fair to both sides), we show also a technique for "rounding" a solution of the fractional stable allocation problem to obtain a relaxed unsplit solution.

2 Background and Preliminaries

2.1 Stable Matching and Allocation Problems

Stable Marriage. The stable marriage (or stable matching) problem takes place on a bipartite graph with men on one side and women on the other, where each individual submits a strictly-ordered, but possibly incomplete preference list of the members of the opposite sex. The goal is to find a matching that is *stable*, containing no *blocking pair* – an unmatched (man, woman) pair (m, w) where m is either unmatched or prefers w to his current partner, and likewise for w.

In their seminal paper [9], Gale and Shapley describe a simple $O(|E|)$ algorithm to find a stable matching for any instance. The most typical incarnation of their algorithm generates a solution that is "man-optimal" and "woman-pessimal", where each man is matched with the best possible partner he could receive in any stable matching, and each woman is matched with the worst possible partner she could receive in any stable matching. By reversing the roles of the men and women, the algorithm can also generate a solution that is simultaneously woman-optimal and man-pessimal.

Stable Allocation. The stable allocation problem was introduced by Baïou and Balinski [1] as a high-multiplicity variant of the stable matching problem, where we match non-unit elements with non-unit elements – here, we speak of matching jobs of varying size with machines of varying capacity. Just as before, jobs and machines submit strict preferences over their outgoing edges in the bipartite assignment graph. If job $j \in J$ prefers machine $m_1 \in M$ to machine $m_2 \in M$, we write $\text{rank}_j(jm_1) > \text{rank}_j(jm_2)$. A stable allocation in this setting is a feasible allocation (as defined in the introduction) where for every edge $jm \in E$ with $x(jm) < c(jm)$, either j is fully assigned to machines at least as good as m, or m is fully assigned to jobs at least as good as j. That is, there can be no *blocking edge* jm where $x(jm) < c(jm)$ and both j and m would prefer to use more of jm. We say that edges with positive x value are in x. If any machine m has $q(m) > \sum_{j \in J} c(jm)$, then $q(m)$ is set to $\sum_{j \in J} c(jm)$. Machines with $x(m) = q(m)$ are *saturated*. Later, when $x(m) > q(m)$ occurs in the relaxed version of the problem, we talk about *over-capacitated* machines. If any job prefers machine m to any of its allocated machines, then m is called *popular*, otherwise m is *unpopular*. Note that all popular machines must be saturated in any stable allocation.

The stable allocation problem can be solved in $O(|E| \log |V|)$ time [6]. There can be many different solutions for the same instance, but they all have the same total allocation $|x|$, and even stronger, the values of $x(j)$ and $x(m)$ for each job and machine remain unchanged across all stable allocations. This holds for both stable marriage [10] and stable allocation [1], moreover, even for stable roommate [11], the non-bipartite version of the problem, and is known as the *rural hospital theorem*. A common application of stable matching in practice is the National Resident Matching Program (NRMP) [14], where medical school graduates in the USA are matched with residency positions at hospitals via a centralized stable matching procedure. A consequence of the rural hospital theorem is that if a less-preferred (typically rural) hospital cannot fill its quota in some stable assignment, then there is no stable assignment in which its quota will be filled.

Like the stable marriage problem, one can always find job-optimal, machine-pessimal and job-pessimal, machine-optimal solutions. To define these notions for the stable allocation problem, Baïou and Balinski [1] define an order on stable solutions based on a *min-min criterion*, where a job j prefers allocation x_1 to allocation x_2 if $x_1(jm) < x_2(jm)$ implies $x_1(jm') = 0$ for every jm' worse than jm for j. A similar relation can be defined for machines as well. Stable matchings and stable allocations both are known to form distributive lattices with an ordering relation based on the min-min criterion.

2.2 Unsplittable Stable Allocation Problems

An unsplit allocation x satisfies $x(jm) \in \{0, q(j)\}$ for all $jm \in E$, so each assigned job is assigned in its entirety to one machine. For simplicity, we introduce a "dummy" machine m_d with high capacity, which acts as the last choice for every job. This lets us assume without loss of generality that an unsplittable allocation always exists in which every job is assigned. In this context, we define the size $|x|$ of an allocation so that jobs assigned to m_d do not count, since they are in reality unassigned. In addition to the application of scheduling jobs in a non-preemptive fashion, a motivating application for the unsplittable stable allocation problem is in assigning personnel with "two-body" constraints. For example, in the NRMP, a married pair of medical school graduates might act as an unsplittable entity of size 2 (this particular application has been studied in substantial detail in the literature; see [2] for further reference).

From an algorithmic standpoint, one of the main results of this paper is that how we define stability in the unsplit case seems quite important. In [12], the following natural definition was proposed: an edge jm is blocking if j prefers m to its current partner, and if m prefers j over $q(j)$ units of its current allocation or unassigned quota. Unfortunately, it was shown in [12] that this definition makes the computation of an unsplit stable allocation NP-hard. We therefore consider an alternative, stricter notion of stability where edge jm is blocking if j prefers m to its current partner, and if m prefers j over *any amount* of its current allocation or has free quota. That is, if j would prefer to be assigned to m over its current partner, than m must be saturated with jobs that m prefers to j. Aside of the integrity constraint, this definition is fully aligned with the classical definition of a stable allocation. As in the splittable case, popular machines must therefore be saturated. Practice shows [13] that if a hospital is willing to hire one person in a couple, but it has no free job opening for the partner, it is most likely amenable to make room for both applicants. Therefore, our definition of a blocking pair serves practical purposes.

Relaxed Unsplit Allocations. The downside of our alternative definition of stability is that it is unlikely to allow feasible unsplit stable allocations to exist in most large instances. Therefore, we consider allowing mildly-infeasible solutions where each machine can be over-capacitated by a single job – a model popularized by previous results in the approximation algorithm literature for standard unsplittable assignment problems [7, 16, 15], and introduced in the context of unsplittable stable allocation by Dean et al. [5]. Specifically, we say that $x : E \to \mathbb{R}_{\geq 0}$ is a *relaxed unsplit allocation* if $x(jm) \in \{0, q(j)\}$ for every edge $jm \in E$, $x(j) \leq q(j)$ for every job $j \in J$, and for each machine m, the removal of the least-preferred job assigned to m would cause $x(m) < q(m)$[1]. Our definition of stability extends readily to the relaxed setting, and we would argue that it is perhaps the most natural mathematical notion of stability to consider in this setting (whereas the form of stability in [12] is probably the most natural for the hard capacity setting). We say a relaxed unsplit allocation x is *stable* if for every edge jm with $x(jm) = 0$, either j is assigned to a machine that j prefers to m, or m's quota is filled or exceeded with jobs that m prefers to j. Otherwise, if edge jm with $x(jm) = 0$ is preferred by j to its allocated machine and m's quota is not filled up with better edges than jm, then *jm blocks x*.

Note that the relaxed unsplit model differs from the non-relaxed unsplit model with capacities inflated by $\max q(j)$, since stability is still defined with respect to original capacities. It may be best to regard "capacities" here as constraints governing start time, rather than completion time of jobs. A machine below its capacity is always willing to launch a new job, irrespective of job size.

[1] The model introduced in [5] allows $x(m) \leq q(m)$, but we believe strict inequality is actually a better choice – mathematically and from a modeling perspective. For example, the old definition applied to a hospital-resident matching scenario with married couples might cause a hospital to accept two more residents than its quota, while the new definition would only require accepting one more resident. The results in [5] hold with either definition.

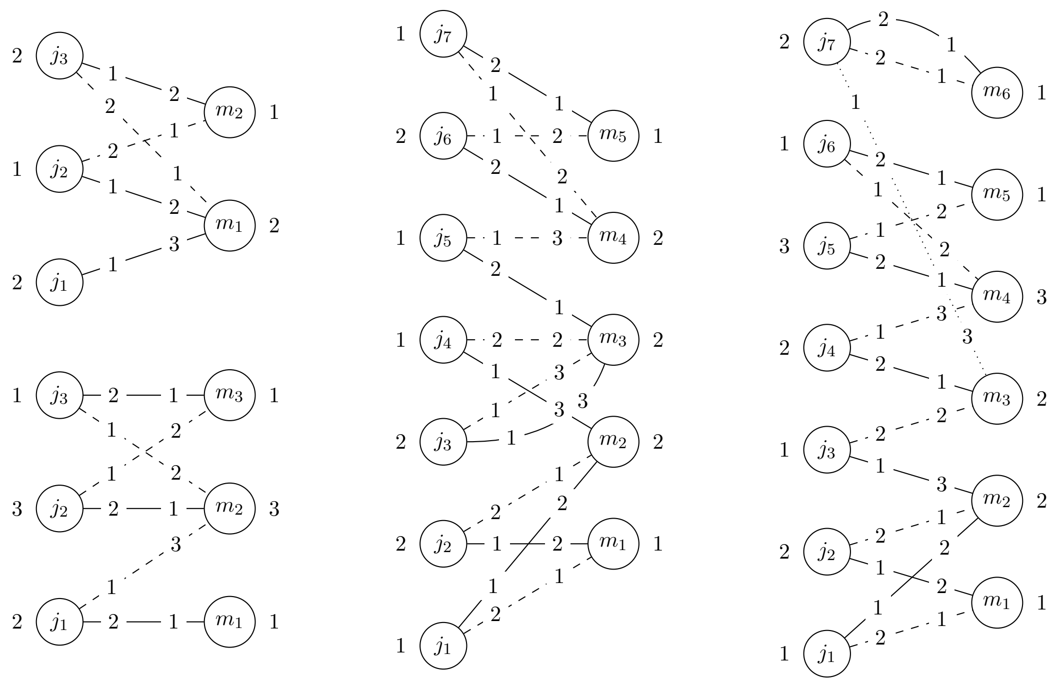


Figure 1: The upper-left instance admits two relaxed unsplit allocations differing in cardinality: The dashed edges form a stable allocation of size 3, while the remaining edges build another stable allocation of size 5. The lower-left example is evidence against an exact rural hospital theorem, where m_1 is empty in one relaxed unsplit stable allocation (given by the dashed edges) but filled beyond its capacity in another (given by the solid edges). The graph in the middle shows two relaxed unsplit allocations that are incomparable from the perspective of m_3. The last instance is a counterexample showing the difficulty of formulating join and meet operations.

3 Machine-Optimal Relaxed Unsplit Allocations

In [5], a version of the Gale-Shapley algorithm is described to find the job-optimal relaxed unsplit stable allocation x_{jopt}. In this context, job-optimal means that there is no relaxed unsplit stable allocation x' such that any job is assigned to a better machine in x' than in x_{jopt}. The implementation described in [5] runs in $O(|E||V|\log|V|)$ time, but $O(|E|)$ is also easy to achieve. In this section, we show how to define and compute a *machine-optimal* relaxed unsplit stable allocation x_{mopt} also in $O(|E|)$ time, and we prove the following:

Theorem 1. *Among all relaxed unsplit stable allocations x, $|x|$ is maximized at $x = x_{jopt}$ and minimized at $x = x_{mopt}$.*

One of the main challenges with computing a machine-optimal allocation is defining machine-optimality. In the stable allocation problem, existence of a machine-optimal allocation follows from the fact that all stable allocations form a distributive lattice under the standard min-min ordering relationship introduced in [1]. However, this ordering seems to depend crucially on the existence of a rural hospital theorem, which no longer holds in the relaxed unsplit case, since relaxed unsplit stable allocations may differ in cardinality (Figure 1). Even an appropriately relaxed version of the rural hospital theorem seems difficult to formulate over relaxed instances: machines can be saturated or even over-capacitated in one relaxed unsplit stable allocation, while being empty in another one (Figure 1). Nonetheless, we can still prove a result in the spirit of the rural hospital theorem, which we discuss further in Section 3.3.

Without an "exact" rural hospital theorem, comparing two allocations using the original min-

$x(jm_d) := q(j)$ for all $j \in J$, $x(jm) := 0$ for every other $jm \in E$
while $\exists m : x(m) < q(m)$ with a non-empty preference list **do**
  m proposes to its best job j with value $q(j)$
  if j prefers m to its current partner **then**
    $x(jm) := q(j)$
    $x(jm') := 0$ for $\forall m' \neq m$
  end if
  delete j from m's preference list
end while

Figure 2: Reversed relaxed unsplit Gale-Shapley algorithm.

min ordering seems problematic, and indeed one can construct instances where two relaxed unsplit stable allocations are incomparable according to this criterion (Figure 1). We therefore adopt a different but nonetheless natural ordering relation: *lexicographical order*. We say that machine m prefers unsplit allocation x_1 to allocation x_2 if the best edge in $x_1 \triangle x_2$ belongs to x_1, where $\triangle$ denotes the symmetric difference operation. The opposite ordering relation is based on the position of jobs, and since jobs are always assigned to machines in an unsplit fashion, the lexicographic and min-min relations are actually the same from the job's perspectives; hence, "job optimal" means the same thing under both. The lexicographical position of the same agent in different allocations can always be compared, and we say a relaxed stable allocation x is *machine-optimal* if it is at least as good for all machines as any other relaxed stable allocation (although we still need to show that such a allocation always exists).

3.1 The Reversed Gale-Shapley Algorithm

For the classical stable marriage problem, the Gale-Shapley algorithm can be reversed easily, with women proposing instead of men, to obtain a woman-optimal solution. We show that this idea can be generalized (carefully accounting for multiple assignment and congestion among machines) to compute a machine-optimal relaxed unsplit stable allocation. Pseudocode for the algorithm appears in Figure 2.

Claim 2. *The algorithm terminates in $O(|E|)$ time.*

Proof. In each step, a job is deleted from a machine's preference list. □

Claim 3. *The algorithm produces an allocation x that is a relaxed unsplit stable allocation.*

Proof. First, we check the three feasibility constraints for x. Since proposals are always made with $q(j)$ and refusals are always full rejections, the quota constraints of the jobs may not be violated. Moreover, each job is assigned to exactly one machine. Machines can be over-capacitated, but deleting the worst job from their preference list results in an allocation under their quota. Otherwise the machine would not have proposed along the last edge. If x is unstable, then there is an empty edge jm blocking x. During the execution, m must have proposed to j. This offer was rejected, because j already had a better partner in the current allocation. Since jobs monotonically improve their position in the allocation, this leads to a contradiction. □

Claim 4. *The output x is the machine-optimal relaxed unsplit stable allocation (i.e., no machine has a better lexicographical position in any other relaxed unsplit stable allocation).*

Proof. Assume that there is a relaxed unsplit stable allocation x', where some machines come better off than in x. To be more precise, in the symmetric difference $x \triangle x'$, the best edge incident to these machines belongs to x'. When running the reversed relaxed unsplit Gale-Shapley algorithm, there is a step when the first such edge jm_1 carries a proposal from m_1 but gets rejected. Otherwise, m_1 filled up or exceeded its quota in x with only better edges than jm_1. Let us consider only this edge first and denote the feasible, but possibly unstable relaxed allocation produced by the algorithm so far by x_0.

When j refused jm_1, it already had a partner m_0 in x_0, better than m_1. Even if there is no guarantee that $jm_0 \in x$, it is sure that $jm_0 \notin x'$ and jm_0 does not block x', though $\text{rank}_j(jm_0) > \text{rank}_j(jm_1)$ for $jm_1 \in x'$. It is only possible if m_0 is saturated or over-capacitated in x' with edges better than jm_0. Since $jm_0 \in x_0$, x_0 may not contain all of these edges, otherwise m_0 is congested in x_0 beyond the level required for a relaxed unsplit allocation. During the execution of the reversed relaxed unsplit Gale-Shapley algorithm, m_0 proposed along all of these edges and got rejected by at least one of them. This edge is never considered again, it may not enter x later. Thus, jm_1 is not the first edge in $x' \setminus x$ that was rejected in the algorithm. □

With this, we completed the constructive proof of the following theorem:

Theorem 5. *The machine-optimal relaxed unsplit stable allocation x_{mopt} can be computed in $O(|E|)$ time.*

3.2 Properties of the Job- and Machine-Optimal Solutions

Theorem 6. *The job-optimal relaxed unsplit stable allocation x_{jopt} is the machine-pessimal relaxed unsplit stable allocation and vice versa, the machine-optimal relaxed unsplit stable allocation x_{mopt} is the job-pessimal relaxed unsplit stable allocation.*

Proof. We start with the first statement. Suppose that there is a relaxed unsplit stable allocation x' that is worse for some machine m than x_{jopt}. This is only possible if m's best edge jm in $x_{\text{jopt}} \triangle x'$ belongs to x_{jopt}. Since x_{jopt} is the job-optimal solution, jm', j's edge in x' is worse than jm. But then, m is saturated or over-capacitated in x' with better edges than jm. We assumed that all edges in x' that are better than jm are also in x_{jopt}. Thus, omitting m's worst job from x_{jopt} leaves m at or over its quota.

The second half of the theorem can be proved similarly, using the reversed Gale-Shapley algorithm. Assume that there is a relaxed unsplit stable allocation x' that assigns some jobs to worse machines than x_{mopt} does. Let us denote the set of edges preferred by any job to its allocated machine in x' by E'. Due to our indirect assumption, E' contains some edges of x_{mopt}. When running the reversed Gale-Shapley algorithm on the instance, there is an edge $jm \in E'$ that is the first edge in E' carrying a proposal. Since j is not yet matched to a better machine, it also accepts this offer. Even if $jm \notin x_{\text{mopt}}$, j's edge in x_{mopt} is at least as good as m, because jobs always improve their position during the course of the reversed Gale-Shapley algorithm. On the other hand, m cannot fulfill its quota in x_{mopt} with better edges than jm, simply because the proposal step along jm took place.

Since $jm \notin x'$, but j prefers jm to its edge in x', m is saturated or over-capacitated with better edges than jm in x'. As observed above, not all of these edges belong to x_{mopt}. Let us denote one of them in $x' \setminus x_{\text{mopt}}$ by $j'm$. Before proposing along jm, m submitted an offer to j' that has been refused. The only reason for such a refusal is that j' has already been matched to a better machine m'. But since $j'm \in x'$, $j'm' \in E'$. This contradicts to our indirect assumption that jm is the first edge in E' that carries a proposal. □

Theorem 1 also follows from the proof above.

We note that although we can compute the job-optimal and machine-optimal relaxed unsplit stable allocations, there in general does not appear to be an obvious underlying lattice structure behind relaxed unsplit solutions. For stable matching or fractional stable allocation, computing the meet or join of two solutions is fairly easy. In order to reach the join (meet) of x_1 and x_2, all machines (jobs) choose the better edge set out of those two allocations [8]. The example in Figure 1 illustrates that this property does not carry over to relaxed unsplit allocations. If all jobs chose the better allocation, m_3 remains empty and j_7m_3 becomes blocking. Similar examples can easily be constructed to show that choosing the worse allocation also can lead to instability.

Our ability to compute x_{mopt} in $O(|E|)$ time now gives us a linear-time method for solving the (non-relaxed) unsplittable stable allocation problem (according to our, stricter notion of stability).

Lemma 7. *If an instance $\mathcal{I}$ admits an unsplit stable assignment x, then the machine-optimal relaxed unsplit stable assignment x_{mopt} on the corresponding relaxed instance $\mathcal{I}'$ is also an unsplit stable assignment on $\mathcal{I}$.*

Proof. Suppose the statement is false, e.g. although there is an unsplit stable assignment x, x_{mopt} is no unsplit stable assignment on $\mathcal{I}$. This can be due to two reasons: either the feasibility or the stability of x_{mopt} is harmed on $\mathcal{I}$. The latter case is easier to handle. An allocation that is feasible on both instances and stable on $\mathcal{I}'$ may not be blocked by any edge on $\mathcal{I}$, since the set of unsaturated edges is identical on both instances. The second case, namely if x_{mopt} violates some feasibility constraint on $\mathcal{I}$, needs more care.

$\mathcal{I}$ and $\mathcal{I}'$ differ only in the constraints on the quota of machines. If x_{mopt} is infeasible on $\mathcal{I}$, then there is a machine m for which $x_{\text{mopt}}(m_1) > q(m_1)$. Regarding the unsplit stable assignment x, the inequality $x(m_1) \leq q(m_1)$ trivially holds. Now we use Theorem 1 for x and x_{mopt} that are both relaxed unsplit stable assignments on $\mathcal{I}'$. This corollary implies that if there is a machine m_1 with $x_{\text{mopt}}(m_1) > x(m_1)$, then another machine m_2 exists for which $x_{\text{mopt}}(m_2) < x(m_2)$ holds.

This machine m_2 plays a crucial role in our proof. It has a lower allocation value in the machine-optimal relaxed solution x_{mopt} than in another relaxed stable solution x on $\mathcal{I}$. Its lexicographical position can only be better in x_{mopt} than in x if the best edge j_2m_2 in $x \triangle x_{\text{mopt}}$ belongs to x_{mopt}. Moreover, $x \triangle x_{\text{mopt}}$ also contains an edge $j_3m_2 \in x$, otherwise $x_{\text{mopt}}(m_2) > x(m_2)$. Naturally, $\text{rank}_m(j_2m_2) < \text{rank}_m(j_3m_2)$. At this point, we use the property that $x_{\text{mopt}}(m_2) < q(m_2)$. Since m_2 has free quota in x_{mopt} and j_3m_2 is not a blocking edge, j_3 must be matched to a machine better than m_2 in x_{mopt}. Thus, there is a job that comes better off in the machine-optimal (and job-pessimal) relaxed solution than in another relaxed stable solution. This contradiction to Theorem 6 finishes our proof. □

Lemma 7 shows that if there is an unsplit solution, it can be found in linear time by computing the machine-optimal relaxed solution. Unfortunately, the existence of such an unsplit assignment is not guaranteed. Our next result applies to the case when no feasible unsplit solution can be found. In terms of congestion, with the machine-optimal solution we come as close as possible to feasibility.

Theorem 8. *Amongst all relaxed unsplit stable solutions, x_{mopt} has the lowest total congestion.*

Proof. Let M_u denote the set of machines that remain under their quota in x_{mopt}. Note that $\sum_{m \notin M_u} x_{\text{mopt}}(m)$, the total allocation value on the remaining machines clearly determines the total congestion of x_{mopt}, given by $\sum_{m \notin M_u} x_{\text{mopt}}(m) - q(m)$. Let x be an arbitrary relaxed solution. Due to Theorem 1, the total allocation value is minimized at x_{mopt}. Therefore, for any relaxed unsplit stable allocation x, the following inequalities hold:

$$\begin{aligned}
\sum_{m\in M} x(m) &\geq \sum_{m\in M} x_{\text{mopt}}(m) \\
\sum_{m\notin M_u} x(m) + \sum_{m\in M_u} x(m) &\geq \sum_{m\notin M_u} x_{\text{mopt}}(m) + \sum_{m\in M_u} x_{\text{mopt}}(m) \\
\sum_{m\notin M_u} x(m) - \sum_{m\notin M_u} x_{\text{mopt}}(m) &\geq \sum_{m\in M_u} x_{\text{mopt}}(m) - \sum_{m\in M_u} x(m) \\
\sum_{m\notin M_u} (x(m) - q(m)) - \sum_{m\notin M_u} (x_{\text{mopt}}(m) - q(m)) &\geq \sum_{m\in M_u} x_{\text{mopt}}(m) - \sum_{m\in M_u} x(m)
\end{aligned}$$

At this point, we investigate the sign of both sides of the last inequality. The core of our proof is to show that for each $m \in M_u$ and relaxed stable solution x, $x_{\text{mopt}}(m) \geq x(m)$. This result, proved below, has two benefits. On one hand, the term on the right hand-side of the last inequality is non-negative. Therefore, the inequality implies that the total congestion on machines in $M \setminus M_u$ is minimized at x_{mopt}. On the other hand, no machine in M_u is over-capacitated in any relaxed solution. Thus, the total congestion is minimized at x_{mopt}. □

Our last observation in this subsection refers to the unsaturated machines.

Lemma 9. *For every $m \in M_u$ and relaxed solution x, the inequality $x_{mopt}(m) \geq x(m)$ holds.*

Proof. Suppose that there is a machine $m \in M_u$ for which $x_{\text{mopt}}(m) < x(m)$ for some relaxed solution x. Since m is unsaturated in x_{mopt}, it is unpopular. On the other hand, there is at least one job j for which $jm \in x \setminus x_{\text{mopt}}$. As m is unpopular in x_{mopt}, j is allocated to a better machine in x_{mopt} than in x. Since x_{mopt} is the job-pessimal solution, we derived a contradiction. □

3.3 A Variant of the "Rural Hospital" Theorem

In the relaxed unsplit case, one can find counterexamples against an exact rural hospital theorem (e.g., where all machines have the same amount of allocation in all relaxed unsplit allocations) or even a weakened theorem stating that all unsaturated / congested machines have the same status in all relaxed unsplit allocations. Lemma 9 above however suggests an alternative variant of "rural hospital" theorem that does hold.

Theorem 10. *A machine m that is not saturated in x_{mopt} will not be saturated in every relaxed unsplit stable solution, and a machine m that is over-capacitated in x_{jopt} must at least be saturated in every relaxed unsplit stable solution.*

Proof. The first part is shown by Lemma 9. For the second part, consider a machine m that is over-capacitated in x_{jopt} but has $x(m) < q(m)$ in some relaxed unsplit allocation x. Consider any job j in $x_{\text{jopt}} \setminus x$, and note that since x_{jopt} is job-optimal, j prefers m to its partner in x. Hence, jm blocks x. □

As of the jobs' side, Theorem 6 already guarantees that if a job is unmatched in x_{jopt}, then it is unmatched in all relaxed stable solutions and similarly, if it is matched in x_{mopt}, then it is matched in all relaxed stable solutions.

4 Rounding Algorithms

We have seen now how to compute x_{jopt} and x_{mopt} in linear time. We now describe how to find potentially other relaxed unsplit solutions by "rounding" solutions to the (fractional) stable allocation problem. For example, this could provide a heuristic for generating relaxed unsplit solutions that are more balanced in terms of fairness between the jobs and machines. Our approach is based on augmentation around *rotations*, alternating cycles that are commonly used in stable matching and allocation problems to move between different stable solutions (see, e.g., [6, 11]).

We begin with a stable allocation x with $x(j) = q(j)$ for every job j, thanks to the existence of a dummy machine. For each job j that is not fully assigned to its first-choice machine, we define its *refusal edge* $r(j)$ to be the worst edge jm incident to j with $x(jm) > 0$. Jobs with refusal edges also have *proposal edges* – namely all their edges ranked better than $r(j)$. Recall that a machine with incoming proposal edges is said to be *popular.* We call a machine *dangerous* if it is over-capacitated and has zero assignment on all its incoming proposal edges.

Claim 11. *Consider a popular machine m in some fractional stable allocation x. Amongst all proposal edges incoming to m, at most one has positive allocation value in x, and this positive proposal edge is ranked lower on m's preference list than any other edge into m with positive allocation.*

Proof. Let $\text{rank}_m(j_1m) > \text{rank}_m(j_2m)$ be proposal edges such that $x(j_1m)$ and $x(j_2m)$ are both positive. Note that j_1m blocks x, since j_1 and m have worse allocated edges in x. A similar argument implies the last part of the claim. □

Our algorithm proceeds by a series of augmentations around rotations, defined as follows. We start from a popular, non-dangerous machine m (if no such machine exists, the algorithm terminates, having reached an unsplit solution). Since m is popular and non-dangerous, it has incoming proposal edges with positive allocation, and due to the preceding claim, it must have exactly one such edge jm. We include jm as well as j's refusal edge jm' in our partial rotation, then continue building the rotation from m' (again finding an incoming proposal edge, etc.). We continue until we close a cycle, visiting some machine m visited earlier (in which case we keep just the cycle as our rotation, not the edges leading up to the cycle), or until we reach a machine m that is unpopular or dangerous, where our rotation ends.

To enact a rotation, we increase the allocation on its proposal edges by ε and decrease along the refusal edges by ε, where ε is chosen to be as large as possible until either (i) a refusal edge along the rotation reaches zero allocation, or (ii) a dangerous machine at the end of the rotation drops down to being exactly saturated from being over-capacitated, and hence ceases to be dangerous. We call case (i) a "regular" augmentation. This concludes the algorithm description.

Claim 12. *The algorithm terminates after $O(|E|)$ augmentations.*

Proof. Jobs remain fully allocated during the whole procedure, and their lexicographical positions never worsen. With every regular augmentation, some edge stops being a refusal edge, and will never again be increased or serve as a proposal or refusal edge. We can therefore have at most $O(|E|)$ regular augmentations. Furthermore, a machine can only become dangerous if one of its incoming refusal pointers reaches zero allocation, so the number of newly-created dangerous machines over the entire algorithm is bounded by $|E|$. Hence, the number of non-regular augmentations is at most $O(|M| + |E|) = O(|E|)$. □

Claim 13. *The final allocation x is a feasible relaxed unsplit assignment.*

Proof. Since we start with a feasible assignment and jobs never lose or gain allocation, the quota condition on jobs cannot be harmed. If there is any edge jm with $0 < x(jm) < q(j)$, then j has at least two positive edges, the better one must be a positive proposal edge. This contradicts the termination condition, and hence x is unsplit.

We now show that deleting the worst job from each machine results in an allocation strictly below the machine's quota. It is clearly true at the beginning, where no machine is over-capacitated (since x starts out as a feasible stable allocation). The only case when $x(m)$ increases is when m is the first machine on a rotation. As such, m has a positive proposal edge jm, which is also its worst allocated edge, due to our earlier claim. If m is not over-capacitated when choosing the rotation, then even if $x(jm)$ rises as high as $q(j)$, this increases $x(m)$ by strictly less than $q(j)$. Thus, deleting jm, the worst allocated edge of m, guarantees that $x(m)$ sinks under $q(m)$. If m is saturated or over-capacitated when choosing the rotation, then jm would have been the best proposal edge of m earlier, when $x(m)$ was not greater than $q(m)$. Thus, assigning j entirely to m does not harm the relaxed quota condition. Let us consider the last step as $x(m)$ exceeded $q(m)$. Again, m was the starting vertex of an augmenting path, having a positive proposal edge. If it was jm, our claim is proved. Otherwise m became over-capacitated while $x(jm)$ was zero, and then increased the allocation on jm. But between those two operations, m had to become dangerous, because it switched its best proposal edge to jm. Dangerous machines never start alternating paths. Thus, we have a contradiction to the fact that we considered the last step when $x(m)$ exceeded $q(m)$. $\square$

Claim 14. *The final allocation x is stable.*

Proof. Suppose some edges block x. Since we started with a stable allocation, there was a step during the execution of the algorithm when the first edge jm became blocking. Before this step, either j or m was saturated or over-capacitated with better edges than jm. The change can be due to two reasons: either j gained allocation on an edge worse than jm, or m gained allocation on an edge worse than jm. As already mentioned, j's lexicographical position never worsens: $\text{rank}_j(p) > \text{rank}_j(r(j))$ always holds. The second event also may not occur, because machines always play their best response strategy. An edge jm that becomes blocking when allocation is increased on an edge worse than it, was already a proposal edge before. Thus, m would have chosen jm, or an edge better than jm to add it to the augmenting path. $\square$

Since each augmentation requires $O(|V|)$ time and there are $O(|E|)$ augmentations, our rounding algorithm runs in $O(|E||V|)$ total time. If desired, dynamic tree data structures can be used (much like in [6]) to augment in $O(\log |V|)$ time, bringing the total time down to just $O(|E| \log |V|)$.

Although jobs improve their lexicographical position in each rotation, the output of the algorithm is not necessarily x_{jopt}. In fact, even x_{mopt} can be reached via this approach. Ideally, this approach can serve as a heuristic to generate many other relaxed unsplit stable allocations, if run from a variety of different initial stable solutions x.

References

[1] M. Baïou and M. Balinski. The stable allocation (or ordinal transportation) problem. *Math. Oper. Res.*, 27(3):485–503, 2002.

[2] P. Biró and F. Klijn. Matching with couples: A multidisciplinary survey. *International Game Theory Review (IGTR)*, 15(02):1340008–1–1, 2013.

[3] A. Caprara, H. Kellerer, and U. Pferschy. A {PTAS} for the multiple subset sum problem with different knapsack capacities. *Information Processing Letters*, 73(3-4):111 – 118, 2000.

[4] C. Chekuri and S. Khanna. A polynomial time approximation scheme for the multiple knapsack problem. *SIAM J. Comput.*, 35(3):713–728, 2005.

[5] B.C. Dean, M.X. Goemans, and N. Immorlica. The unsplittable stable marriage problem. In *IFIP TCS*, volume 209 of *IFIP*, pages 65–75. Springer, 2006.

[6] B.C. Dean and S. Munshi. Faster algorithms for stable allocation problems. *Algorithmica*, 58(1):59–81, 2010.

[7] Y. Dinitz, N. Garg, and M.X. Goemans. On the single-source unsplittable flow problem. *Combinatorica*, 19:17–41, 1999.

[8] T. Fleiner. On stable matchings and flows. In *Proceedings of the 36th International Workshop on Graph-Theoretic Concepts in Computer Science*, WG'10, pages 51–62, 2010.

[9] D. Gale and L. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 1:9–14, 1962.

[10] D. Gale and M. Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11:223–232, 1985.

[11] D. Gusfield and R.W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989.

[12] E. McDermid and D. Manlove. Keeping partners together: algorithmic results for the hospitals/residents problem with couples. *J. Comb. Optim.*, 19(3):279–303, 2010.

[13] A.E. Roth. The national residency matching program as a labor market. *Journal of the American Medical Association*, 275(13):1054–1056, 1996.

[14] A.E. Roth. Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory*, 36(3-4):537–569, 2008.

[15] D.B. Shmoys and É. Tardos. Scheduling unrelated machines with costs. In *Proceedings of SODA*, pages 448–454, 1993.

[16] M. Skutella. Approximating the single source unsplittable min-cost flow problem. In *Proceedings of FOCS*, pages 136–145, 2000.

Factor Revealing LPs and Stable Matching with Ties and Incomplete Lists

Brian C. Dean   Rommel Jalasutram
School of Computing, Clemson University
{bcdean,rjalasu}@clemson.edu

Abstract

Stable matching with ties and incomplete lists (SMTI) is one of the most prominent NP-hard problems in the domain of ordinal matching, where we seek to find a bipartite matching between a set of men and a set of women that is "stable" with respect to their preferences. When ties in preference lists are restricted to one side of the problem, Iwama et al. [9] devised a variant of the famous Gale-Shapley stable matching algorithm that breaks ties using edge weights from a linear programming (LP) relaxation of the problem, leading to an approximation ratio of $\frac{25}{17} \approx 1.4706$. We apply ideas from factor-revealing LPs to show, via computational proof involving the solution of massive LPs, that their analysis can be systematically improved to yield an approximation ratio of $19/13 \approx 1.4615$, improving the best currently-known ratio (obtained via different techniques in [15]) of $\frac{41}{28} \approx 1.4643$.

1 Introduction

Consider an instance I of a bipartite matching problem over a graph $G = (L \cup R, E)$, where L and R represent sets of men and women. Each man $m \in L$ and woman $w \in R$ submits a ranked preference list over potential partners in E. In the context of a matching $M \subseteq E$, edge $(m, w) \in E$ is a *blocking pair* if m is unmatched or prefers w to his partner in M, and similarly w is unmatched or prefers m to her partner in M. Informally, m and w would be unhappy if we imposed the matching M, since they would prefer to be matched with each-other. We say M is *stable* if no blocking pairs exist. Stable matchings and their variants have been a vibrant subfield of algorithmic game theory since the seminal work of Gale and Shapley [1] in the 1960s. In particular, the following results are now well known:

- Gale and Shapley proposed a simple algorithm that computes a stable matching in $O(|E|)$ time (i.e., linear time).
- If preference lists are strictly ordered (no ties) and *complete* ($E = L \times R$), then the Gale-Shapley (GS) algorithm finds a matching M of maximum possible size, with $|M| = \min(|L|, |R|)$. Otherwise, if preference lists are not complete, the resulting matching may include both unmatched men and women, but the unmatched men and women are the same in every stable matching, so all stable matchings have equal cardinality.
- If preference lists are complete but contain ties, the GS algorithm still computes a matching of size $\min(|L|, |R|)$. Here, stability is typically defined such that an edge $(m, w) \notin M$ is blocking if m is unmatched or *strictly* prefers w to his partner in M, and likewise for w.

Stable matching with both ties and incomplete lists (SMTI) is more challenging. Here, stable matchings may differ in cardinality, as shown in Figure 1, and the problem of computing a maximum-cardinality stable matching is NP-hard [7, 12], even with ties restricted to one side of the problem (in preference lists of the women, as is the typical convention). Since one-sided ties and incomplete preference lists are anticipated in many ordinal matching problems in practice (e.g., [6]), the SMTI problem has received substantial attention recently, and is currently one of the most actively-studied problems in the ordinal matching research community. In this paper, we show a polynomial-time algorithm with the strongest approximation ratio for this problem to date, $19/13 \approx 1.4615$. Our approach is somewhat distinct in that we rely on the computational solution of a massive "factor-revealing" LP to systematically generalize and strengthen the analysis of Iwama al. [9]. The application of factor-revealing LPs to online bipartite matching has recently appeared in [2, 3], but to the best of our knowledge, this work is the first involving factor-revealing LPs in the domain of stable matching.

Background. The GS algorithm produces a *maximal* stable matching, since any edge $(m, w) \in E$ between two unmatched individuals m and w would be blocking. It therefore gives a 2-approximate solution, owing to the well-known fact that a maximum matching can be at most twice the cardinality of any maximal matching. Using local search techniques, Iwama et al. [8] improved the factor to $\frac{15}{8}$ for the general case where ties are allowed on both sides. Király [10] introduced the idea of promoting unmatched men to higher levels of priority to break ties, achieving a ratio of $\frac{5}{3}$ for the general case and $\frac{3}{2}$ for the case where ties are restricted to one side. McDermid [13] improved the factor for the general case to $\frac{3}{2}$ by exploiting a classical graph theory result known as the Gallai-Edmonds decomposition, and Paluch [14] and Király [11] gave linear time algorithms obtaining the same ratio. Following this, Iwama, Miyazaki, and Yanagisawa [9] improved the approximation factor for the case with one-sided ties to $\frac{25}{17} \approx 1.4706$, by using the edge weights in a fractional linear programming (LP) relaxation of the problem to break ties in a natural way while running the GS algorithm. More recently, paper, Huang et al. [5] used different techniques to improve the ratio for the one-sided case to $\frac{22}{15} \approx 1.4667$; the analysis was tightened to $\frac{41}{28} \approx 1.4643$ by Radnai [15]. In terms of hardness results, the SMTI problem with two-sided ties cannot be approximated to a ratio better than $\frac{33}{29}$ unless P = NP, or $\frac{4}{3}$ assuming the unique games conjecture [18]. These results even apply when each individual has a single tie of length two. For the one-sided variant, we cannot approximate with a ratio better than $\frac{21}{19}$ unless P = NP [18], or $\frac{5}{4}$ assuming the unique games conjecture [4].

2 Linear Programming Relaxation

The following LP relaxation of the one-sided SMTI problem was initially studied in [16] and [17]. The notation $m \succ_w m'$ indicates that m is preferred to m' by w.

$$
\begin{aligned}
LP(I) = \max \quad & \sum_{(m,w)\in E} x_{mw} & & \\
\text{subject to:} \quad & \sum_{w:(m,w)\in E} x_{mw} \le 1 & \forall m \in L & \quad (1) \\
& \sum_{m:(m,w)\in E} x_{mw} \le 1 & \forall w \in R & \quad (2) \\
& \sum_{w' \succ_m w} x_{mw'} + \sum_{m' \succeq_w m} x_{m'w} \ge 1 & \forall (m,w) \in E & \quad (3) \\
& x_{mw} \ge 0 & \forall (m,w) \in E &
\end{aligned}
$$

For each edge $(m, w) \in E$, the variable $x_{mw} \in [0, 1]$ indicates the extent to which m and w are matched. Constraints (1) and (2) are standard for matching problems, and ensure that each person can be matched to at most one other person. Constraint (3) enforces stability if

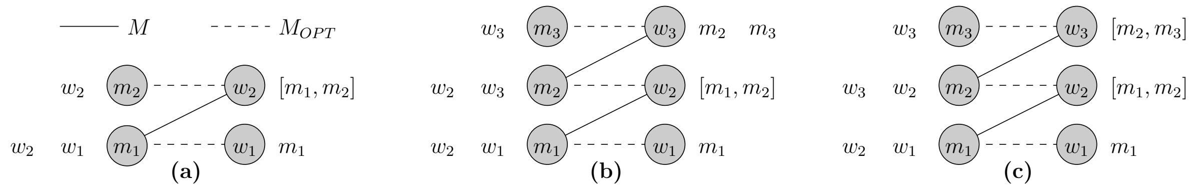


Figure 1: Simple instances admitting two stable matchings M and M_{OPT} of different cardinality: (a) an "augmenting 3-path", and (b)-(c) "augmenting 5-paths". Preference lists are shown alongside each man and woman, with ties in brackets.

x_{mw}'s are restricted to integers, since it states that for every edge $(m, w) \in E$, at least one of m or w should be matched to a partner that is equally or more preferred than along this edge. Since the SMTI problem is exactly modeled by the formulation above with the added constraint that $x_{mw} \in \{0, 1\}$ for all $(m, w) \in E$, we have $LP(I) \geq |M_{OPT}(I)|$ for any instance I, where $M_{OPT}(I)$ is a maximum-cardinality stable matching for I.

Iwama et al. [9] show that the integrality gap of this LP relaxation is at least $1 + \frac{1}{e} \approx 1.3679$. Any analysis based on comparing to the objective of this LP relaxation therefore cannot lead to a stronger approximation guarantee.

3 The GS-LP Algorithm

Our algorithm is essentially the same as used by Iwama et al. [9], which is built from the standard GS algorithm by adding extra machinery to decide how to break ties. Since the GS algorithm is known to generate a stable matching irrespective of how ties are broken, we are therefore still guaranteed to find a stable matching. The standard GS algorithm is quite straightforward to describe: in each iteration, an arbitrary unassigned man proposes to the next woman on his preference list. Each woman tentatively holds on to the best proposal she has received to date, rejecting all other offers. When a man is rejected, he continues issuing proposals down his list. Each edge is proposed along at most once, leading to a linear running time.

3.1 A Prioritization Scheme for Tie-Breaking

As a preprocessing step, we solve the LP relaxation to obtain a solution x^*. Each man m is then assigned integer and fractional priority levels $P_i(m) \in \{0, 1, 2, \ldots\}$ and $P_f(m) \in [0, 1]$, both initially zero. These priorities are used to resolve ties: if w is tentatively matched with m but receives a proposal from a man m' for which $m =_w m'$, she accepts if $P_i(m') > P_i(m)$, or if $P_i(m) = P_i(m')$ but $P_f(m') > P_f(m)$.

Let $w(m)$ denote the farthest woman down m's preference list up to whom he has proposed; initially, $w(m) = \emptyset$. The fractional priority for m is defined as $P_f(m) = \sum_{w \succeq_m w(m)} x^*_{mw}$. Every time m is rejected by $w(m)$, he returns to the start of his preference list and has another chance to propose to all the women up to $w(m)$ in sequence (he may be more successful this time, since his fractional priority now includes the contribution due to $x^*_{m,w(m)}$). If m runs off the end of his preference list (after being rejected by the final woman and then proposing one last time to everyone on the list), he increments the value of his integer priority $P_i(m)$, resets $w(m)$ to $\emptyset$ (so $P_f(m)$ resets to zero as well), and restarts the proposal process from the beginning of his list. Men whose priorities reach a specified threshold T cease to issue any further proposals, and the algorithm terminates when every unmatched man m has $P_i(m) = T$. The algorithm of Iwama et al. uses $T = 3$, but we consider using higher thresholds; it turns out that $T = 4$ is sufficient for our analysis.

The entire algorithm clearly runs in polynomial time. Speedups to the GS-based part above may certainly be possible, although these may offer diminishing returns, since the solution of the LP relaxation will likely dominate the running time.

3.2 Integer Prioritization and Restricted Augmenting Paths

Many prior approximation algorithms for the SMTI problem are structured as above, using the GS algorithm along with some method of prioritization to break ties. The work of Paluch [14] and Kiraly [11] used essentially only the integer part of the prioritization scheme above (for which the LP solution is not needed), and leads to an approximation bound of 3/2 due to the following observation on the resulting matching M. This holds in our case as well since we use the same integer prioritization machinery, and it takes precedence over any tie-breaking decisions made due to fractional priorities.

Lemma 1. *Integer prioritization, where man m' wins a tie over man m only if $P_i(m') > P_i(m)$, yields a matching M with no augmenting paths of length 3 or shorter.*

Fixing a particular optimal matching M_{OPT}, $M \oplus M_{OPT}$ is the graph formed by taking edges that occur in M or M_{OPT} but not both. An augmenting path is a odd-length path in $M \oplus M_{OPT}$ containing one more edge from M_{OPT} than from M. It is well-known that if all augmenting paths admitted by M have length at least $2k+1$, then $|M_{OPT}| \leq \frac{k+1}{k}|M|$. This is easy to argue by looking at how much $|M|$ expands when transforming M into M_{OPT} by toggling all the edges in $M \oplus M_{OPT}$: cycles, even-length paths, and odd-length non-augmenting paths in $M \oplus M_{OPT}$ can be toggled without increasing $|M|$, and an augmenting path of length $2n + 1 \geq 2k + 1$ has n edges from M and $n + 1$ from M_{OPT}, so its cardinality expands by a factor of $\frac{n+1}{n} \leq \frac{k+1}{k}$.

Proof of Lemma 1. The GS algorithm itself prevents augmenting 1-paths, since it produces a maximal matching. Any augmenting 3-path must have the structure shown in Figure 1(a) in order for its edges in both M and M_{OPT} to be non-blocking. Woman w_1 clearly never saw a proposal, since any woman receiving a proposal will become and henceforth remain matched. Thus, $P_i(m_1) = 0$. However, since m_2 is unassigned, $P_i(m_2) > 0$, and so w_2 would have chosen m_2 over m_1. □

Integer prioritization is not sufficient to prevent all augmenting paths of length 5 or longer, although it can prevent some types of these paths from arising. Iwama et al. show that an augmenting 5-path in the specific form shown in Figure 1(b) can occur, since no amount of prioritization on the single man m_3 can induce w_3 to accept him instead of m_2. Figure 1(c) depicts one of several augmenting 5-paths that cannot occur with $T \geq 3$. To see why, let us number the men and women in any path in $M \oplus M_{OPT}$ sequentially as in Figure 1, so that if the path contains a woman unmatched in M, such a woman is labeled w_1 at the beginning of the path. Edges in M_{OPT} are of the form (m_j, w_j), and edges in M are of the form (m_j, w_{j+1}).

Definition 1 (High and low priority for men)**.** *Man m_j in a path in $M \oplus M_{OPT}$ has high priority if $P_i(m_j) \geq j$ at termination; otherwise m_j has low priority.*

Definition 2 (High and low priority for women)**.** *Any woman w_j in a path in $M \oplus M_{OPT}$ unmatched in M has low priority; otherwise if matched in M, woman w_j has high priority if and only if m_{j-1} (her partner in M) has high priority.*

The following facts are easy to establish:

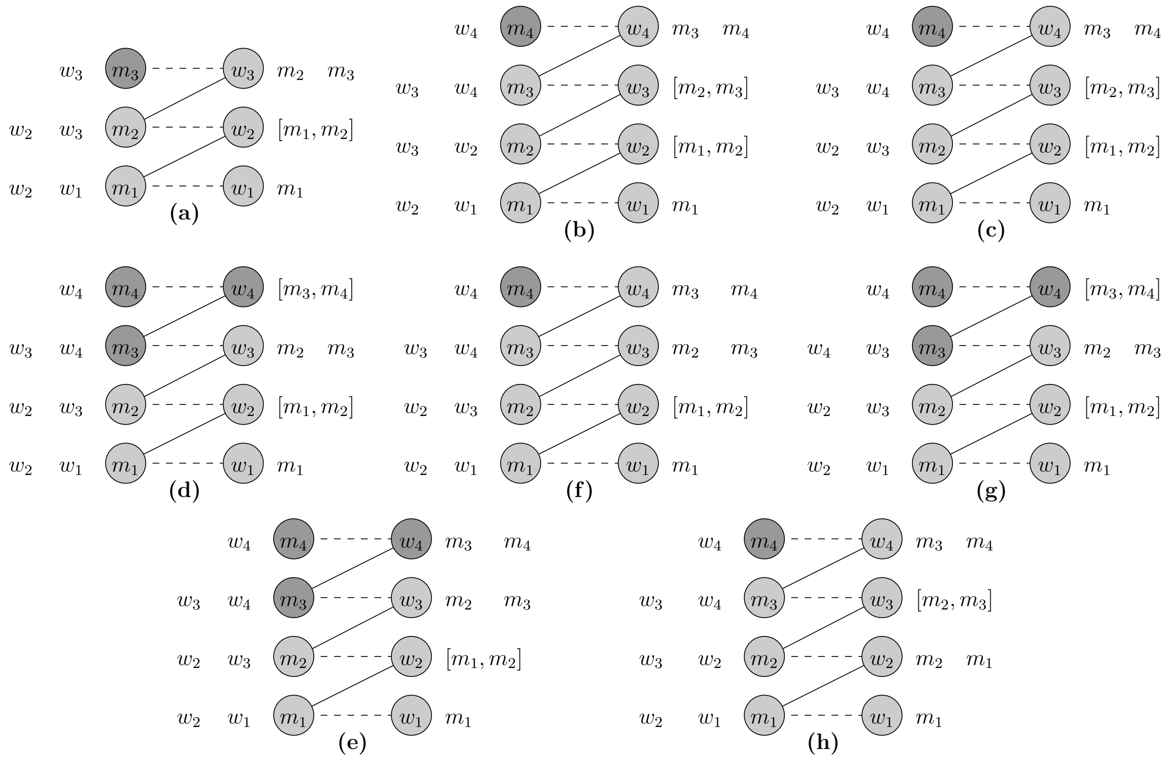


Figure 2: The only possible configurations for (a) an augmenting 5-path (assuming $T \geq 3$), and (b)-(h) augmenting 7-paths (assuming $T \geq 4$). Darkened nodes indicate individuals with high priority.

(i) $P_i(m_j) = 0$ (and hence m_j has low priority) if he is matched in M_{OPT} to a woman (w_j) unmatched in M. Otherwise, he would have proposed to w_j, and she would have accepted. Since matched women stay matched, w_j could not have ended up single.

(ii) Man m_j has $P_i(m_j) = 0$ (and hence has low priority) if $m_j \succ_{w_j} m_{j-1}$, for similar reasons. Otherwise, m_j would have proposed to w_j and she would have accepted, preventing the edge (m_{j-1}, w_j) from becoming part of M.

(iii) If m_j has high priority and $m_j =_{w_j} m_{j-1}$, then m_{j-1} has high priority, since $P_i(m_j) \geq j$ implies that m_j must have been proposed to w_j at least at priority level $j-1$. Hence, $P_i(m_{j-1}) \geq j-1$, or else there is no way w_j could have accepted m_{j-1} in the matching M.

If $T \geq 3$, then the highest man (m_3) in an augmenting 5-path will have high priority, since m_3 ends up unmatched, with priority $P_i(m_3) = T$. For The augmenting 5-path in Figure 1(c) therefore cannot occur because by applying (iii) twice, m_1 would have high priority, contradicting (i).

If $T \geq 4$, the highest man (m_4) in an augmenting 7-path will have high priority, and we can similarly deduce that augmenting 7-paths can only occur in the seven configurations shown in Figure 2(b)-(h). One can programatically enumerate all such possible types of augmenting k-paths for larger values of k by filtering out those that would violate the priority rules above, as well as those where edges in M block M_{OPT} or vice versa. These grow at an exponential rate: there are 37 different augmenting 9-paths, 181 augmenting 11-paths, and so on. As we discover in our analysis, however, we only need to consider up to augmenting 7-paths, so we can set $T = 4$. Raising T higher, while useful for restricting augmenting 9-paths and longer, cannot further help us restrict the set of augmenting 7-paths.

Taking into account the restricted structure of augmenting 5-paths, and by adding fractional prioritization based on the LP relaxation, Iwama et al. were able to show using a fairly complicated "ad hoc" analysis that a solution cannot admit only augmenting 5-paths. Longer paths are also necessary, thereby improving slightly on the 3/2 approximation ratio one would obtain with only augmenting 5-paths. Our analysis follows a similar pattern, but using a more general approach involving a factor-revealing LP that considers augmenting 7-paths and beyond.

4 Towards a Factor-Revealing LP

In order to analyze the approximation factor of the GS-LP algorithm, we look for an instance I maximizing the ratio $A(I) = |M_{OPT}(I)|/|M(I)|$, where $M(I)$ is the smallest possible matching GS-LP can produce when run on I. We search for I systematically by solving a large LP, known as a factor-revealing LP since its output gives us a valid upper bound on the approximation ratio of the GS-LP algorithm. The factor-revealing LP, which we call FLP, has two types of decision variables, "class" variables (α's) and "edge bundle" variables (β's).

4.1 Class variables

Fixing an instance I, consider overlaying the edges of $M = M(I)$ and $M_{OPT} = M_{OPT}(I)$. We find a number of different *classes* of structures, shown in Figure 3, which allow us to characterize the edges in M as well as the individuals who are unmatched in M. We find: (a) edges $(m, w) \in M \cap M_{OPT}$, (b) individuals that are unmatched in both M and M_{OPT}, (c) individuals matched in M_{OPT} but not M (these are the endpoints of paths in $M \oplus M_{OPT}$), (d) edges $(m, w) \in M$ where both m and w are matched, differently, in M_{OPT}, and where (m, w) belongs to a path in $M \oplus M_{OPT}$, (e) edges $(m, w) \in M$ where only w is matched in M_{OPT} (these are found at the ends of paths in $M \oplus M_{OPT}$ that are not augmenting paths), (f) edges $(m, w) \in M$ where only m is matched in M_{OPT} (these are found at the ends of paths in $M \oplus M_{OPT}$ that are not augmenting paths), and (g) edges $(m, w) \in M$ found in cycles in $M \oplus M_{OPT}$.

Every edge $(m, w) \in M$ can be categorized into one of these classes, and similarly every man $m \in L$ and woman $w \in R$ in our instance belongs to exactly one of these classes. Observe that we have further differentiated our classes based on the structure of their preference lists with respect to placement of partners in M and M_{OPT}, and also whether individuals are high or low priority (note that this designation is only defined for individuals on paths in $M \oplus M_{OPT}$). For example, the first class shown in Figure 3(f) represents all edges $(m, w) \in M$ in our instance where m is matched to a better partner (w_o) in M_{OPT} than in M, where w is unmatched in M_{OPT}, and where both m and w have low priority.

Let C denote the set of the classes above (note that soon we will enlarge C to contain even more classes, namely explicit representations of all the valid configurations of augmenting 5-paths, 7-paths, etc.). For each class $c \in C$, let g_c denote the number of edges in M represented by a single instance of c. We have $g_c = 1$ for all classes above except those of types (b) and (c), where $g_c = 0$. Let n_c denote the number of occurrences of class c when we overlay M and M_{OPT}, and let the decision variable α_c represent $n_c/|M|$. Since all edges in M appear in exactly one occurrences of some class,

$$\sum_{c \in C} g_c n_c = |M|,$$

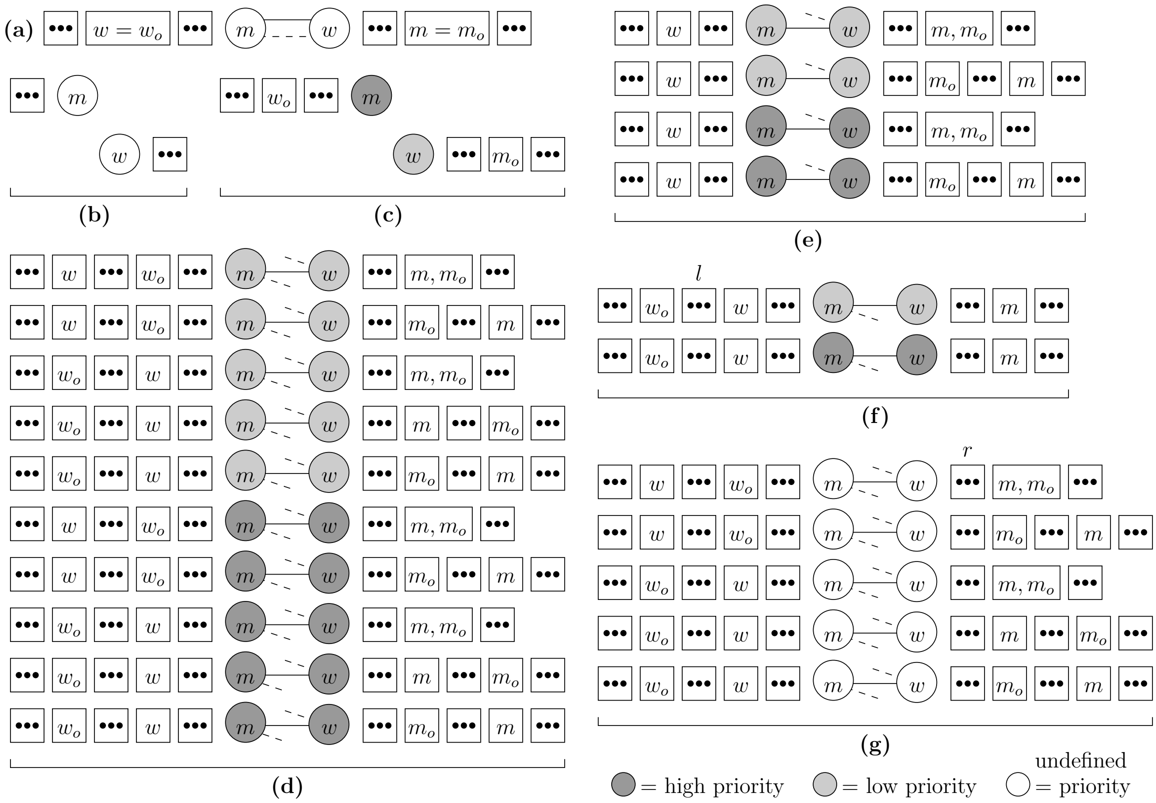


Figure 3: Classes of structures over which the factor-revealing LP optimizes, where m_o and w_o indicate partners in M_{OPT}. Boxes denote distinct "regions" in preference lists, where "$\cdots$" indicates a region with possibly several entries. Since women can have ties in their lists, a right-hand side region with a specific label like "m_2, m_3" represents a block of preference list entries that are all tied with m_2 and m_3. In contrast, a left-hand side region with a specific label like "w_1" represents only a single specific entry in a preference list. The pair (l, r) is one example of a pair of regions between which edges in our instance cannot run, since they would block M. The single man in (c) will have high priority provided that T is sufficiently large.

and by dividing both sides by $|M|$, we obtain

$$\sum_{c \in C} g_c \alpha_c = 1, \tag{4}$$

giving a simplex-like constraint in FLP that governs the relative proportion α_c of each class $c \in C$.

Let f_c denote the "local" approximation factor for class c – that is, an upper bound on the factor by which the edges of M belonging to instances of c would expand if we were to transform M into M_{OPT} by toggling the edges in $M \oplus M_{OPT}$. We set $f_c = 0$ by convention for the special cases of (b) and (c), since these involve no edges from M. We have $f_c = 1$ trivially for classes c of type (a), and we also have $f_c = 1$ for classes c of types (e), (f), and (g) since these can only belong to non-augmenting paths or cycles in $M \oplus M_{OPT}$. For classes c of type (d), we have $f_c = 3/2$, since these represent edges in M that might belong to augmenting paths of length 5 or longer (these could also belong to non-augmenting paths, of course, but in this case the local approximation ratio is at most $1 \leq 3/2$). We let C_R denote the 10 classes of type (d), singling them out since they are the only classes so far with a non-trivial local approximation

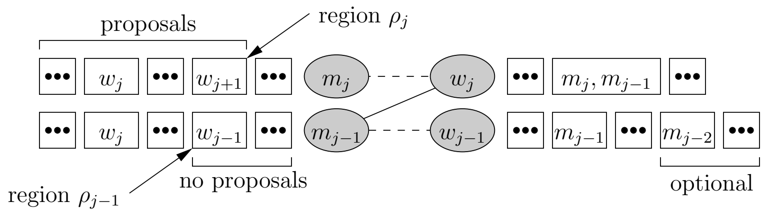


Figure 4: The part of an augmenting path class to which a tie-breaking constraint applies. Woman w_{j-1} could be either unmatched in M or she must prefer her partner m_{j-1} in M_{OPT} to her partner m_{j-2} in M.

ratio. Since every edge in M is represented by an instance of some class $c \in C$, we have

$$|M_{OPT}| \leq \sum_{c \in C} g_c f_c n_c. \tag{5}$$

4.2 Edge Bundle Variables

In addition to the class variables α_c, FLP also contains edge bundle variables β_{lr} that represent the aggregate weight allocated to specific bundles of edges in an optimal solution x^* to $LP(I)$.

In Figure 3, the preference lists in each class $c \in C$ are each partitioned into distinct *regions*. For example, consider again the first of the two classes shown in Figure 3(f), describing an edge $(m, w) \in M$ where m is matched in M_{OPT} (to w_o) but w remains single in M_{OPT}. Here, the preference list of m is divided into five regions: (1) women m prefers more than w_o, (2) the woman w_o, (3) women m prefers less than w_o but more than w, (4) the woman w, and (5) women m prefers less than both w and w_o. The preference list of w is divided into three regions: (1) men she prefers more than m, (2) men she prefers the same amount as m, and (3) men she prefers less than m.

Let Z_L and Z_R respectively denote the sets of all left-hand and right-hand regions across all the classes $c \in C$. For any edge $e = (m, w) \in E$ in our problem instance, we can classify its left endpoint according to its region type $l(e) \in Z_L$ and its right endpoint according to its type $r(e) \in Z_R$. We now define β_{lr} to represent the following quantity:

$$\beta_{lr} = \frac{1}{|M|} \sum_{\substack{e \in E \\ l(e)=l, r(e)=r}} x^*_e. \tag{6}$$

That is, β_{lr} denotes the aggregate weight assigned by $LP(I)$ to the bundle of all edges from region type $l \in Z_L$ to region type $r \in Z_R$. Note that not every pair of regions $(l, r) \in Z_L \times Z_R$ is *valid* in its ability to describe edges from E. The following pairings (l, r) are invalid, since there cannot be any edges $e \in E$ satisfying $l = l(e)$ and $r = r(e)$: (i) region pairs (l, r) corresponding to edges that would block M (for example, the pair (l, r) shown in Figure 3), (ii) region pairs (l, r) corresponding to edges that would block M_{OPT}, (iii) region pairs (l, r) corresponding to edges from a high-priority man m to a woman w, where m is preferred to her partner in M (this is not possible since $P_i(m) > 0$, so m would have proposed to w, and w would have accepted, preventing later acceptance of her partner in M), (iv) region pairs (l, r) corresponding to edges from a high-priority man m to a low-priority woman w, where m is tied with her partner in M (similarly, m would have proposed to w, and w would have accepted, preferring m to her partner in M), and (v) region pairs (l, r) where l is a region representing a single specific individual (e.g., one's partner w_o in M_{OPT}), but r is not a compatible match for that individual (e.g., not labeled with w_o).

Let $B \subset Z_L \times Z_R$ denote the set of all valid region pairs that remain after filtering out the invalid pairs enumerated above. Since every edge $e \in E$ appears in exactly one such pair $(l(e), r(e)) \in B$, we have

$$\sum_{(l,r)\in B} \beta_{lr} = \frac{1}{|M|} \sum_{e\in E} x^*_e. \tag{7}$$

4.3 The Factor-Revealing LP

We can now state FLP:

$$FLP = \max \left\{ \min \left(\sum_{c\in C} g_c f_c \alpha_c, \sum_{(l,r)\in B} \beta_{lr} \right) : (\alpha, \beta) \in P \right\}, \tag{8}$$

where P denotes its feasible region, described with 5 types of constraints:

- *Nonnegativity* constraints: $\alpha, \beta \geq 0$.
- A *simplex* constraint: $\sum g_c \alpha_c = 1$.
- *Tie-breaking* constraints, stating that weights assigned to edge bundles would have resulted in fractional priorities leading to tie-breaking that would have resulted in the formation of augmenting paths. Consider any part of an augmenting path class matching the situation described by Figure 4. Note that $P_i(m_{j-1}) = 0$, since m_{j-1} must never have proposed to w_{j-1}. Therefore, the fact that w_j chose m_{j-1} over m_j implies that $P_i(m_j) = 0$ and so $P'_f(m_j) \leq P'_f(m_{j-1})$ where P'_f denotes P_f at the last point in time at which the algorithm broke the tie. Since m_j is matched with w_{j+1} in M, his fractional priority at that point in time was equal to the weight of all entries in his preference list prior to w_{j+1}:

  $$P'_f(m_j) = \sum_{\substack{e=(m_j,w)\in E \\ l(e)\succ\rho_j}} x^*_e,$$

  where ρ_j is the region of m_j's list containing w_{j+1}, as indicated in the figure. Since m_{j-1} never proposed to w_{j-1} and also never proposed to any woman w matched in M to a man m for which $m_{j-1} \succ_w m$, we have

  $$P'_f(m_{j-1}) \leq \sum_{\substack{e=(m_{j-1},w)\in E \\ l(e)\succ\rho_{j-1}, r(e)\notin Z^M_R}} x^*_e,$$

  where $Z^M_R \subset Z_R$ denotes the set of all right-hand side regions that are more preferred than the partner in M within their same list. Combining these, we have

  $$\sum_{\substack{e=(m_j,w)\in E \\ l(e)\succ\rho_j}} x^*_e \leq \sum_{\substack{e=(m_{j-1},w)\in E \\ l(e)\succ\rho_{j-1}, r(e)\notin Z^M_R}} x^*_e,$$

  which in aggregate across all class instances becomes the following tie-breaking constraint in FLP:

  $$\sum_{\substack{(l,r)\in B \\ l\succ\rho_j}} \beta_{lr} \leq \sum_{\substack{(l,r)\in B \\ l\succ\rho_{j-1}, r\notin Z^M_R}} \beta_{lr}.$$

- *Matching* constraints, stating that the total weight incident to any generic man or woman in class c must not exceed α_c. Consider some generic man m belonging to class c, and let $L(m)$ denote the set of n_c individual men in L represented by the generic man m. Since each individual man $m' \in L(m)$ satisfies constraint (1) in $LP(I)$, we sum these constraints to obtain

$$\sum_{\substack{(m',w)\in E \\ m'\in L(m)}} x^*_{m'w} \leq n_c,$$

  Therefore,

$$\sum_{\substack{(l,r)\in B \\ l\in Z_L(m)}} \beta_{lr} = \frac{1}{|M(I)|} \sum_{\substack{(m',w)\in E \\ m'\in L(m)}} x^*_{m'w} \leq \frac{n_c}{|M(I)|} = \alpha_c,$$

  where $Z_L(m)$ denotes the regions in m's preference list. Similar constraints hold for the women.

- *Stability* constraints, which correspond to aggregations of the stability constraints from the original SMTI LP relaxation across our edge bundle variables. Consider any class c and region pair (l, r) representing an edge explicitly found within c. So far, this includes just the single edge $(m, w) \in M$ represented by class c, but soon we will expand C by including larger augmenting structures that also explicitly include edges from M_{OPT}. We must have

$$\sum_{\substack{(l',r')\in B \\ l'\succ l}} \beta_{l'r'} + \sum_{\substack{(l',r')\in B \\ r'\succeq r}} \beta_{l'r'} \geq \alpha_c, \tag{9}$$

  where the notation $r' \succ r$ indicates that r' is a region more preferred than r belonging to the same preference list. That is, the total weight incident to regions in the preference list of l preceding l, plus the total weight incident to regions in the preference list of r up to and including r, must be at least α_c.

5 Analysis and Refinement of FLP

Lemma 2. *For any instance I, let α and β be defined as above. Then $(\alpha, \beta) \in P$.*

Proof. Nonnegativity of α and β is clear, validity of the simplex constraint is shown in (4), and validity of the tie-breaking and matching constraints are argued above. Validity of the stability constraints are shown exactly the same way, by aggregating constraint (3) from $LP(I)$ over all n_c individual instances of an edge appearing explicitly within some class c. □

We now establish our main result:

Theorem 3. *Let $A(I)$ denote the approximation ratio of GS-LP when run on instance I. Then $FLP \geq A(I)$.*

Proof. Let α, β be defined as above. Lemma 2 shows that $(\alpha, \beta) \in P$. We now have

$$\sum_{c\in C} g_c f_c \alpha_c = \frac{1}{|M(I)|} \sum_{c\in C} g_c f_c n_c \geq \frac{1}{|M(I)|} |M_{OPT}(I)| = A(I),$$

where the first equality follows from the definition of α_c, and the next inequality is a re-statement of (5). From (7), we also have

$$\sum_{(l,r)\in B} \beta_{lr} = \frac{1}{|M(I)|} \sum_{e\in E} x^*_e = \frac{1}{|M(I)|} LP(I) \geq \frac{1}{|M(I)|} |M_{OPT}(I)| = A(I).$$

Hence,

$$\min\left(\sum_{c\in C} g_c f_c \alpha_c, \sum_{(l,r)\in B} \beta_{lr}\right) \geq A(I),$$

so we have demonstrated a feasible solution of FLP with objective value at least $A(I)$. The maximum value of FLP is therefore also an upper bound for $A(I)$. $\square$

5.1 Adding More Classes

The optimal solution value of FLP as stated above is 3/2, and this is not surprising, since it really isn't capturing any insight beyond the fact that there are no augmenting 3-paths. To improve it, we need to add more explicit knowledge about the structure of augmenting 5-paths, so that it can deduce (like the work of Iwama et al.) that it is not possible to form an instance I admitting only augmenting 5-paths.

We therefore add a new class c representing the unique structure of an augmenting 5-path to C, with $g_c = 2$ and $f_c = 3/2$. Accordingly, we reduce f_c to 4/3 for all classes $c \in C_R$ – since we are now explicitly representing augmenting 5-paths as their own class, the classes in C_R now represent the "left-over" edges in M that are at worst now part of augmenting 7-paths. Solving FLP now yields an optimal solution of value 25/17, matching the result of Iwama et al. In fact, we can think of FLP in this case as automating the proof of Iwama et al. Instead of starting with known inequalities (e.g., the constraints from the LP relaxation) and combining them in an "ad hoc" series of lemmas and theorems, FLP automatically finds the best combination.

To improve further, we can add more classes to C that explicitly represent all possible arrangements of augmenting 7-paths, reducing f_c to 5/4 for all $c \in C_R$. This yields an optimal solution of value $19/13 \approx 1.4615$. One might hope that this process would continue reducing the upper bound on our approximation ratio in an asymptotic manner, since each time we explicitly include longer augmenting paths, the optimum solution of FLP cannot increase.

Lemma 4. *Let FLP_k denote FLP when up through augmenting k-paths are explicitly represented. Then $FLP_{k+2} \leq FLP_k$.*

Proof. This follows from the fact that FLP_k is a relaxation of FLP_{k+2}. That is, the optimal solution (α, β) for FLP_{k+2} can be naturally mapped to a feasible solution (α', β') for FLP_k (by assigning all weight from classes representing augmenting $(k+2)$-paths to the corresponding classes in C_R), for which $FLP_k \geq FLP_k(\alpha', \beta') \geq FLP_{k+2}(\alpha, \beta) = FLP_{k+2}$. $\square$

Unfortunately, when we solve FLP after explicit inclusion of augmenting 9-paths and beyond, we still obtain an optimal solution of 19/13, and indeed this is the best result we can obtain, since when we build FLP up to just augmenting 7-paths, we find an optimal solution in which $\alpha_c = 0$ for all $c \in C_R$.

Lemma 5. *If an optimal solution of FLP_k assigns $\alpha_c = 0$ for all $c \in C_R$, then $FLP_k = FLP_{k'}$ for all $k' \geq k$.*

Proof. The fact that $FLP_k \geq FLP_{k'}$ comes from Lemma 4. To show the reverse inequality, observe that any optimal solution (α, β) for FLP_k with $\alpha_c = 0$ for all $c \in C_R$ can be naturally mapped to an equivalent solution (α', β') of $FLP_{k'}$ with $FLP_k = FLP_k(\alpha, \beta) = FLP_{k'}(\alpha', \beta') \leq FLP_{k'}$, since only the classes $c \in C_R$ change their meaning and local approximation ratio f_c when moving from k to k'. $\square$

To ensure that our results are accurate (due to the highly complex nature of FLP), we built three independent code bases for generating and solving FLP, all of which agreed in their output.

Acknowledgments

The authors received support from NSF CAREER award CCF-0845593.

References

[1] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[2] Gagan Goel and Aranyak Mehta. Online budgeted matching in random input models with applications to adwords. In *SODA*, pages 982–991, 2008.

[3] Gagan Goel and Pushkar Tripathi. Matching with our eyes closed. In *FOCS*, pages 718–727, 2012.

[4] Magnús M. Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Improved approximation results for the stable marriage problem. *ACM Transactions on Algorithms*, 3(3), 2007.

[5] Chien-Chung Huang and Telikepalli Kavitha. An improved approximation algorithm for the stable marriage problem with one-sided ties. In *IPCO*, pages 297–308, 2014.

[6] Robert W. Irving and David Manlove. Approximation algorithms for hard variants of the stable marriage and hospitals/residents problems. *Journal of Combinatorial Optimization*, 16(3):279–292, 2008.

[7] Kazuo Iwama, David Manlove, Shuichi Miyazaki, and Yasufumi Morita. Stable marriage with incomplete lists and ties. In *In Proceedings of ICALP 99: the 26th International Colloquium on Automata, Languages and Programming*, pages 443–452. Springer-Verlag, 1999.

[8] Kazuo Iwama, Shuichi Miyazaki, and Naoya Yamauchi. A 1.875: approximation algorithm for the stable marriage problem. In *SODA*, pages 288–297, 2007.

[9] Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. A 25/17-approximation algorithm for the stable marriage problem with one-sided ties. In *ESA (2)*, pages 135–146, 2010.

[10] Zoltán Király. Better and simpler approximation algorithms for the stable marriage problem. In *ESA*, pages 623–634, 2008.

[11] Zoltán Király. Linear time local approximation algorithm for maximum stable marriage. *Algorithms*, 6(3):471–484, 2013.

[12] David F. Manlove, Robert W. Irving, Kazuo Iwama, Shuichi Miyazaki, and Yasufumi Morita. Hard variants of stable marriage. *Theoretical Computer Science*, 276(1-2):261–279, April 2002.

[13] Eric McDermid. A 3/2-approximation algorithm for general stable marriage. In *ICALP (1)*, pages 689–700, 2009.

[14] Katarzyna E. Paluch. Faster and simpler approximation of stable matchings. In *WAOA*, pages 176–187, 2011.

[15] András Radnai. Approximation algorithms for the stable matching problem. Master's thesis, Eötvös Loránd University, 2014.

[16] Alvin E. Roth, Uriel G. Rothblum, and John H. Vande Vate. Stable matchings, optimal assignments, and linear programming. *Mathematics of Operations Research*, 18(4):803–828, November 1993.

[17] Chung-Piaw Teo and Jay Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23(4):874–891, 1998.

[18] H Yanagisawa. *Approximation algorithms for stable marriage problems*. PhD thesis, Kyoto University, 2007.

Dynamic Task Assignments: An Online Two Sided Matching Approach

Sujit Gujar and Boi Faltings
Artifitial Intelligence Laboratory
École polytechnique fédérale de Lausanne

Abstract

For the task assignment problem in an expert crowdsourcing platform, we propose that the dynamically arriving workers report their preferences for the tasks as ordinal preferences to the platform. We model then the task assignment problem as a dynamic two sided matching problem. In this paper we study the dynamic two sided matching when the men (the workers) side of the market is arriving dynamically and the women (the requesters) side is available since beginning. We assume strict preferences of the agents. Using a deferred acceptance algorithm as a building block, we first develop f^{APODA}, a class of strategy-proof online mechanisms. We design f^{APODA} and f^{ThODA} in this class. As no mechanism can achieve stability in our settings, we propose a weaker notion of stability, namely, progressive stability. We introduce an online mechanism f^{RODA} that achieve the progressive stability. For achieving good rank-efficiency, we design an online matching mechanism f^{BOMA}. We study all the four mechanisms empirically for stability and rank-efficiency.

1 Introduction

The term *crowdsourcing* was introduced by Howe [11]. Since then, crowdsourcing has become popular to get tasks done in the form of an open call over Internet. Currently there are thousands of websites, also called as *platforms*, available for crowdsourcing.[1] In the crowdsourcing market, there are two types of users of the platform, the one who posts the task is called *a requester* and the one who seeks to work on the task is called *a worker*. Crowdsourcing of complex macro tasks is referred to as *expert crowdsourcing* [21]. For example, oDesk, topcoder are expert crowdsourcing platforms. Consider the following scenario of an expert crowdsourcing as shown in the Figure 1.

Example 1 *On a Monday morning three requesters login to a crowdsourcing platform with their tasks. These tasks are to develop software modules and are having deadlines in two weeks. Rewards for these tasks are \$600, \$700 and \$650 respectively.* m_1, m_2 *and* m_3 *are eligible and interested workers for these tasks. The requester of task* w_1 *prefers* m_3*, then* m_1 *and then* m_2*. The worker* m_3 *prefers to work on* w_1 *then* m_2, m_3*. Similarly the other workers (and also the requesters) have preferences over the tasks (workers).* m_1 *is available from Monday morning till Tuesday evening for the task assignment, where as* m_2 *and* m_3 *are present only on Monday and Tuesday respectively. The goal is to optimally assign the tasks to the dynamic workers.*

As seen in the above example, in expert crowdsourcing market, the requesters have preferences over the workers who are assigned to their tasks and the workers have preferences over the tasks they are assigned. For the workers, one of the important advantages is to select tasks of their own choice. Hence in such task assignments, it is very important to cater to the preferences of the workers to retain them with the platform.

[1] http://crowdsourcing.org

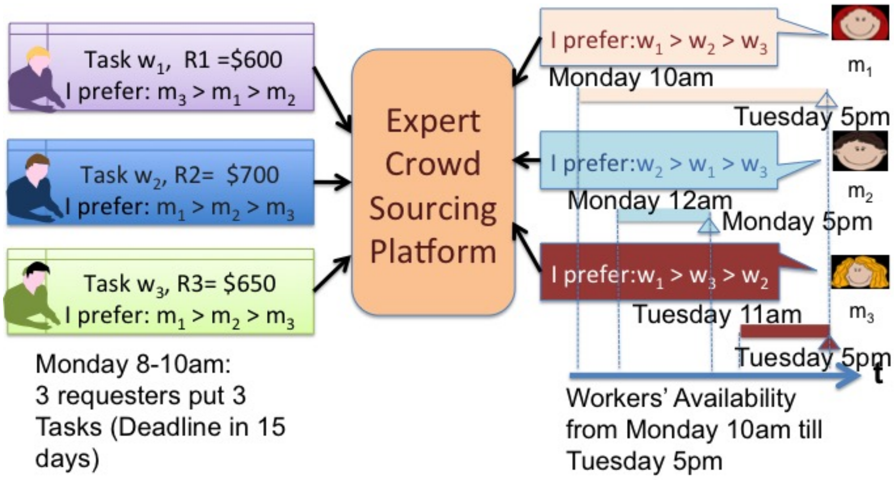


Figure 1: Crowdsourcing scenario

We model the task assignment problem as a matching in two sided market: the requesters as one side and the workers as another, both the sides having preferences for the match they obtain. We refer to the requesters as *Women* and the workers as *Men*. Two sided matching problem is extensively studied using game theory in *static settings*, that is, all men and women are simultaneously available for matching. However, this need not be the case in the real world applications. As seen in the above example, each worker arrives dynamically and needs to get the task before he leaves. Matching in such dynamic settings is called as *online matching*. The strategic men may manipulate a matching mechanism by misreporting their preferences. Gujar and Parkes [8] addressed the online matching in two sided markets in a game theoretic setting. However, in their setting men are static where as in our settings men are dynamic. The authors [8] assume external pool of men available as substitutes. In this paper we do not assume such pool of men or women and construct online matching mechanisms. In particular, the following are our contributions.

Contributions We propose to use dynamic two sided matching for the task assignment problem of an expert crowdsourcing platform.

- First we develop a class of online matching mechanisms, by partitioning dynamically arriving men, which we call as *Partition Online Deferred Acceptance*, f^{PODA}. These mechanisms are truthful.
  - We design two partition mechanisms, *Arrival Priority Online Deferred Acceptance*, f^{APODA} and *Threshold Online Deferred Acceptance* f^{ThODA} .
- It is impossible to achieve stability in online settings. Hence, we introduce a notion of *progressive stability* in online matching. We propose a f^{RODA}, an online matching mechanism that achieves the progressive stability at the cost of truthfulness. We believe that f^{RODA} satisfies weaker notion of truthfulness, namely ex-ante truthful.
- To obtain a good rank efficiency, that is average rank of a matching proposed by a mechanism, we devise an online matching mechanism f^{BOMA}.

The rest of the paper is organized as follows.

Organization First we review the related research in Section 2. In Section 3, we explain our model and notation. We design truthful mechanisms in Section 4. We propose mechanisms to improve stability and rank-efficiency in Section 5. We study our mechanisms empirically and discuss the properties achieved by these mechanisms in Section 6. We conclude the paper in Section 7.

2 Related Work

Task Assignment in Crowdsourcing [10, 9, 4, 13, 17, 12, 21, 2] addressed the task assignment problem in crowdsourcing. However, most of them are concerned only about the quality of the answers received and how to assign tasks to workers so as to meet a requester's goals. Difallah *et. al.* [6] proposed to push the tasks to the workers based on their preferences. These are only categorical preferences and not the workers' preferences for the requesters or for any specific tasks. The authors did not address the requesters' preferences. Moreover, the workers may be strategic in reporting their preferences which is not addressed in [6].

Akbarpour *et. al.* [1] designed algorithms for dynamic matching markets. The goal in [1] is to assign a maximum number of matches in large markets with dynamic population. The authors did not consider the preferences of the participants.

Two Sided Matching In their seminal work, Gale and Shapley abstracted two sided matching as a marriage problem [7]. The authors introduced a notion of *stability* and proposed an algorithm *Deferred Acceptance* (DA). Since then two sided matching problem has been extensively studied and applied in many real world settings. Roth [18] proved that there is no strategy-proof mechanism that achieves stability. Majumdar [15] proved that, under certain conditions, the DA satisfies a weaker notion of strategy-proofness, namely Bayesian Incentive Compatibility. If we allow the participants to report weak preferences, that is, indifference among alternatives, there are exponential number of stable matchings and selecting a stable matching in such settings poses algorithmic challenges. For details about static two sided matching, we refer the readers to [20, 19, 16].

In online settings, [14] designed algorithms assuming agents are truthful. Compte and Jehiel [5] considered a different dynamic matching problem to the one studied here. In their model, all men and women are static, but the men and women experience a preference shock and are interested in re-match. It imposes an individual-rationality constraint across periods so that no man or woman becomes worse off as the match changes in response to a shock. The authors demonstrated how to modify the deferred acceptance algorithm to their problem. The dynamic matching was addressed by [8] for the case of static men.

3 Preliminaries and Notation

Motivated by the task assignment problem in an expert crowdsourcing, we make certain assumptions.

Assumptions

- The men side is dynamic where as the women side is static. (In Expert crowdsourcing scenario: the tasks are complex and need much longer duration to complete them. We restrict to time window during which these tasks needs to be completed once posted. The workers log-in at different times.)
- The Men are strategic and the women are honest in reporting the preferences. (In our example, preferences of the requesters can be derived from skills required for the tasks and performance of the workers in similar tasks in the past where as the workers may be strategic in reporting their preferences.)

- The preferences of the men and women are strict.[2]
- The men do not lie about their arrival departure periods. Such settings are called as *exogenous*.

We use the following notation in the rest of the paper.

3.1 Notation

In the market there are n men (M) at one side and n women (W) at the other. Every agent is interested in obtaining a match at the other side. Let $\succ_i$ be strict preference order of an agent $i \in M \cup W$ over other side of the market. A preference profile of the agents is denoted as: $\succ = (\succ_i, \succ_{-i})$, where $\succ_{-i}$ is the preferences of all the remaining agents apart from i. m_j arrives into the market in period a_j and is available for matching till d_j. We denote the schedules of arrival and departure of men by $\rho = \{(a_j, d_j) : j \in M\}$ We denote a match by μ where $\mu(m) \in W \cup \{\phi\}$ and $\mu(w) \in M \cup \{\phi\}$. Our notation is summarized in Table 1.

n	Total number of men (women)
M	Set of Men
W	Set of Women
a_j, d_j	Arrival time and departure time for a man m_j
ρ	$= \{(a_j, d_j)_{m_j \in M}\}$ Arrival-Departure Schedule of Men
$\succ_i$	Preference of $i \in M \cup W$
$\succ$	$= (\succ_i, \succ_{-i})$ Preference profile of all agents
$W(t)$	Women that are not matched till t.
$M(t)$	$\{m_j \ni a_j \leq t \leq d_j$ and m_j is not matched.$\}$
$AM(t)$	$\{m_j \mid a_j = t\}$ Set of men arriving in time slot t
$DM(t)$	$\{m_j \mid d_j = t\}$ Set of men departing in time slot t
f	Matching mechanism
μ	$= f(\succ, \rho)$. A matching

Table 1: Notation

An *online matching mechanism* f selects a matching $\mu = f(\succ, \rho)$ [8]. A matching mechanism f should be feasible. That is, it should match each $m_j \in M$ before d_j, that is before he leaves the system. Important desirable properties of a matching mechanism are *truthfulness*, *stability* and *rank-efficiency*.

Definition 1 (Strategy-Proof) *Online mechanism f is* strategy-proof *(or* truthful*) for men if for each man m, for all arrival-departure schedules ρ, and for all preferences $\prec_{-m}$, and for all $\succ'_m \neq \succ_m$,*

$$\mu'(m) \not\succ \mu(m),$$

where $\mu' = f(\succ'_m, \succ_{-m}, \rho)$.

Definition 2 (Stability) *We say a pair $(m, w) \in M \times W$ blocks a matching μ if, $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. If there is no blocking pair, we say the matching is stable. And if a matching mechanism always produce stable matching, we say the mechanism is stable.*

[2] We further assume that the agents are able to identify their preferences. In general crowdsourcing, it may be in-feasible for agents to know own preferences over large number of tasks. However, in an expert crowdsourcing, the number of tasks in which the worker is interested, is limited.

To evaluate performance of an online mechanism, we also consider its rank-efficiency. Rank-efficiency of an online mechanism is an expected rank for each agent that it assigns to its match. We, following Budish and Cantillon [3], assume risk neutral agents with a constant difference in utility across the matches that are adjacent in their preference list. The rank of an agent i for a matching μ, written $rank_i(\mu)$, is the rank order of the agent with whom he or she is matched. A match by i with the most-preferred agent in $\succ_i$ receives rank order 1 and with the least-preferred receives rank order n. If $\mu(i) = \phi$ then the rank-order is $n + 1$. Based on this, the rank of a matching μ is $rank(\mu) = \frac{1}{2n} \sum_{i \in M \cup W} rank_i(\mu)$.

To define the rank-efficiency of a mechanism we assume a distribution function Φ on $(\succ, \rho)$ and compute the expected rank over the induced distribution on matches:

Definition 3 (Rank-efficiency) *The rank-efficiency of an online mechanism f, given distribution function Φ, is*

$$rank^f = \mathbb{E}_{(\succ,\rho)\sim\Phi}[rank(f(\succ, \rho))].$$

We first describe a celebrated algorithm, Deferred Acceptance, in the next subsection.

3.2 Static Matching Mechanisms

Lets say $\rho = \{(1,1)_{m_i} \forall m_i \in M\}$. We refer to this as *static settings*. Gale and Shapley [7] proposed a deferred acceptance algorithm when all the agents are static.

Definition 4 (Man-proposing DA) *Each man states his most preferred woman. Each woman keeps the best match and rejects other men. All rejected men then propose to their next preferred women. The procedure continues until there are no more rejections.*

We denote matching produced by the above algorithm as $DA(M, W)$. This mechanism is strategy-proof for men and always selects a man-optimal stable matching [19]. That is no other stable matching is preferred by all the men. Similarly we can define a woman-proposing DA.[3] In general, it is not possible to have a matching mechanism that is stable and strategy-proof for both sides of the market [18]. Hence, we address incentive constraints only for the men side. Since a DA has interesting game theoretic properties, we design strategy-proof mechanisms for the dynamic settings using the DA as a building block.

4 Dynamic Matching: Truthful Mechanisms

In dynamic settings, either one of the two sides is dynamic or both the sides are dynamic. In this paper, we do not address the case when both the sides are dynamic.

4.1 Dynamic Matching: The Static Men and The Dynamic Women

Gujar and Parkes [8] consider a setting when the strategic side (men) is static and the honest side (women) is dynamic. The authors propose a matching mechanism GSODAS. In GSODAS, the authors propose to use $DA(M, W(t))$ in each period t where $W(t)$ is a set of women available in t. If a man gets matched with a woman better than his previous period match, he skips the previous match. If his previously matched woman has already left the market, she gets a substitute for him. The authors assume an external pool of men available as substitutes, may be in a secondary market. The authors prove that the GSODAS is stable and strategy-proof for the men.

Though their setting resembles to our setting, in our setting the men side is dynamic where as in [8] it is static.

[3]Unless otherwise stated, when we say DA in this paper, we mean man-proposing DA.

4.2 Dynamic Matching: The Dynamic Men and The Static Women

As seen in Section 3, a DA is very simple, fast algorithm and always leads to stable matching in static settings. It is strategy-proof for men. Hence, we use it as a building block in designing matching mechanisms. We make some observations for dynamic settings.

(1) We do not assume possibility of substitutes. For example, consider the expert crowdsourcing market described in the Introduction. Once a worker is assigned a task and he starts working on it, the platform cannot preempt it from him and request him to get another task in another market. Thus, if a woman w is matched with a man m and if m leaves the system in period time t, she is not available for matching for all time periods $> t$.

(2) Suppose we execute DA at t_1 and $t_2 > t_1$ with all the men available in the system at those periods. If a man m participates in both the DAs, he gets better or worse match than the first DA. Due to possible collisions across multiple DAs, it is not feasible to guarantee him the best match across multiple DAs. This can potentially create misreports by allowing him to get better match in the end. This motivates our mechanisms in the next section.

4.3 Partition Online Deferred Acceptance (PODA)

Every man is matched using DA at one period in his availability and this match is final. No man is considered for DA more than once. This induces a partition among the men based on the time of their matching. Let $\Pi = \{M_1, \dots, M_k\}$ be collection of subsets of M such that (i) $M_i \cap M_j = \phi \forall i \neq j$, (ii) $\cup M_i = M$, (iii) $t_1, \dots, t_k$ such that $\forall m_j \in M_i, a_j \leq t_i \leq d_j$, and (iv) this partition is is independent of the preferences of men.

With this we propose a class of online matching mechanisms which we call as *Partition online DA* (PODA).

ALGORITHM 1: Matching Mechanism f^{PODA}

Input: Preferences of Men and Women $(\succ)$, ρ, Π, $\{t_1, \dots, t_k\}$
Output: A matching μ

1 $t = 1$, $W(1) = W$
2 **if** $t \in \{t_1, \dots, t_k\}$ **then**
3     $\mu^{t_i} = DA(M_i, W(t))$
4 $t \leftarrow t + 1$
5 $W(t) \leftarrow W(t-1) \setminus \{w : \mu^t(w) \neq \phi\}$
6 **if** $W(t) == \phi$ **then**
7     $\mu = \cup_{t_i} \mu^{t_i}$
8   STOP.
9 GO TO STEP 2.

Proposition 1 *f^{PODA} is strategy-proof for men.*

In f^{PODA}, each man is part of a DA only once. A DA is strategy-proof for men, hence no man can benefit by misreporting his preference when he is matched using a DA. As the partition is independent of the preferences of the men, no man's preference can influence his competitors in DA. That implies that no man can benefit by misreporting. Hence, f^{PODA} is strategy-proof for men.

□

Examples of f^{PODA}

f^{PODA} assumes Π is given. We show with two examples how to create partition without knowing any details about the men arriving in the future and their schedule.

4.3.1 Arrival Priority Online Deferred Acceptance (f^{APODA})

In this mechanism we consider the partition of men induced by their arrival periods. That is, in each period t if $AM(t) \neq \phi$ we execute $DA(AM(t), W(t))$. Its very simple and easy to implement. However, it may lead to many blocking pairs. To optimize for stability, we propose the following mechanism.

4.3.2 Threshold Online Deferred Acceptance (f^{ThODA})

In this mechanism, we induce a partition of the men greedily and based on a given parameter threshold (Th). The basic idea is to accumulate more men for matching every time we execute the DA. The parameter Th can be optimized for a given stochastic process of arrival-departure of the men. The proposed matching mechanism is follows:

ALGORITHM 2: Matching Mechanism f^{ThODA}

Input: Preferences of Men and Women ($\succ$), ρ, Th
Output: A matching μ
1 $t = 1$, $M(1) = AM(1), W(1) = W$
2 **if** $DM(t) == \phi$ **then**
3   DO NOTHING
4 **if** $DM(t) \neq \phi$ *and* $|M(t)| > Th$ **then**
5     $\mu^t = DA(M(t), W(t))$
6 **if** $DW(t) \neq \phi$ *and* $|M(t)| \leq Th$ **then**
7     $\mu^t = DA(DM(t), W(t))$
8 $t \leftarrow t + 1$
9 $W(t) \leftarrow W(t-1) \setminus \{w : \mu^t(w) \neq \phi\}$,
$M(t) \leftarrow \{M(t-1) \cup AM(t)\} \setminus \{m : \mu^t(m) \neq \phi\} \cup DM(t)$
10 **if** $W(t) == \phi$ **then**
11     $\mu = \cup_t \mu^t$
12   STOP.
13 GO TO STEP 2.

Both the above mechanisms are truthful. We now discuss the stability of these mechanisms in the following subsection.

4.4 Stability in Partition DA

In general, no mechanism can predict preferences of the agents yet to arrive. So, it is not fair to expect stability in dynamic settings. Note that GSODAS can achieve stability with external pool of men in secondary markets. We do not assume possibility of substitutes. It follows from Proposition 3.1 in [8] that stability is impossible in our settings. Hence, in this paper, we propose a weaker notion of stability, namely *progressive stability*. The idea is, at each instance of time, there is no blocking pair in the system. We do not allow women, that are matched with men who are no longer available for matching, to form a blocking pair. More formally, progressive stability is defined as follows.

Definition 5 (Progressive Stability) *A pair (m, w) is said to be blocking pair at time t if (i) m, w both are present in the system at t, and not matched with each other, (ii) prefer to match with each other than their current match, (iii) their current matches are also present in the system. In each time period, if no such pair exists, we say a matching is progressively stable.*

PODA mechanisms cannot achieve progressive stability. The partition is independent of the preferences. Consider the following instance of preferences. Let $t_1 < t_2$ and w matched with $m_1 \in M_1$ with $d_1 > t_2$. For some $m_2 \in M_2$, if $m_2 \succ_w m_1$ and m_2 prefers w most, (m_2, w) is blocking pair at t_2.

In the next section, we look for non-strategy proof mechanisms to improve stability and rank-efficiency.

5 Dynamic Mechanisms: Progressive Stability and Rank-Efficiency

5.1 Repeated Online Deferred Acceptance (RODA)

We run the DA in every period and only the matches involving the departing men are final.

ALGORITHM 3: Matching Mechanism f^{RODA}

Input: Preferences of Men and Women $(\succ)$, ρ

Output: A matching μ

$t = 1$, $M(1) = AM(1), W(1) = W$

$\mu^t = DA(M(t), W(t))$

for $m \in DM(t)$ **do**

  $\mu(m) = \mu^t(m)$ and $\mu(\mu^t(m)) = m$

$t \leftarrow t + 1$

$M(t) \leftarrow \{M(t-1) \cup AM(t)\} \setminus DM(t)$, $W(t) \leftarrow \{W(t-1) \setminus \{w : \mu^t(w) \in DM(t)\}$

if $W(t) == \phi$ **then**

  STOP.

GO TO STEP 2.

In f^{RODA}, men get match only at their departure. Hence, men may have strong incentive for early departure. However, in this paper, we focus only on exogenous settings. We illustrate f^{RODA} with the following example.

Example 2 *Consider the example shown in Figure 1. Let their preferences be as in Table 2. We refer to Monday as $t = 1$ and Tuesday as $t = 2$. Workers m_1, m_2 arrive at $t = 1$ and m_2 leaves in the same slot. m_3 arrives at $t = 2$. In first round, (m_1, w_1) and (m_2, w_2) are matched using the DA. As m_2 has to leave by end of slot 1, that match is final. m_1 has deadline $t = 2$ and hence does not start working on T_1. In second period, using the DA (m_1, w_3) and (m_3, w_1) are matched and as all workers have appeared in the system, this match is final.*

Workers (Men)		Requesters (Women	
m_1	$w_1 \succ w_2 \succ w_3$	w_1	$m_3 \succ m_1 \succ m_2$
m_2	$w_2 \succ w_1 \succ w_3$	w_2	$m_1 \succ m_2 \succ m_3$
m_3	$w_1 \succ w_3 \succ w_2$	w_3	$m_1 \succ m_2 \succ m_3$

Table 2: f^{RODA} Example Preferences

Theorem 5.1 *f^{RODA} is progressively stable.*

Proof: We prove by induction. At $t = 1$, we are using the static version of a DA and hence the matching at $t = 1$ are stable. Say f^{RODA} is progressively stable till $t = \tau$. We show that f^{RODA} cannot introduce any blocking pair at $t = \tau + 1$. Let's assume (m, w) is blocking at $\tau + 1$. Say at $\tau + 1$, m is matched with w' and w is matched with m'. For $m : w \succ w'$, and for $w : m \succ m'$. In the time slot $\tau + 1$, m must have proposed to w before w'. If w is matched with m' from previous round and is still present in the system, w would had rejected him and be matched with m. Thus, w must be matched with some worker who already left the system. Hence, w cannot be part of blocking pair in progressive stability. Thus, f^{RODA} is progressively stable.

□

f^{RODA} achieves progressive stability but is manipulable:

Claim 1 *f^{RODA} is not strategy-proof.*

Proof: In Example (2), m_1 can report preference to be $w_2 \succ w_1 \succ w_3$. He gets matched with w_2 in period 1 and this does not change in period 2. By misreporting, he obtains a preferable match than w_3.

□

Note that, (i) such manipulation requires information about the preference of a man yet to arrive. (ii) By simulations with uniform preferences, we observed, Prob$(rank_i(f^{RODA}(\succ, \rho)) = k)$ decreases with k. Hence we believe that f^{RODA} is ex-ante strategy-proof for uniform preferences. (iii) If beliefs of men about the preferences of men yet to arrive are uniform even after observing their own preference, it need not give any incentive to men to misreport in f^{RODA}.

5.2 Rank-Efficiency

In the static settings, the DA selects a man-optimal stable matching and hence, it is Pareto-efficient. However, as we consider ordinal preferences, to measure efficiency, we use a rank-efficiency of a matching mechanism. To achieve stability, the DA mechanism has to compromise on rank-efficiency.

For example, say there are three men m_1, m_2, m_3 and three women w_1, w_2, w_3. Let $\succ_{m_1} = w_1 \succ w_2 \succ w_3$ and $\succ_{m_2} = \succ_{m_3} = w_1 \succ w_3 \succ w_2$. All three women have preference $m_1 \succ m_2 \succ m_3$. Assume $a_m = d_m = 1$ for all three men. The DA will produce matching μ^D : $(m_1, w_1), (m_2, w_3), (m_3, w_2)$ with $rank(\mu^D) = 2$ and a matching $\mu : (m_1, w_2), (m_2, w_1), (m_3, w_3)$ has $rank(\mu) = \frac{11}{6}$.

With this example, to improve rank-efficiency, we propose f^{BOMA}, an online matching mechanism that uses maximum weight bipartite matching for matching the men.

5.2.1 Bipartite Online Matching Algorithm

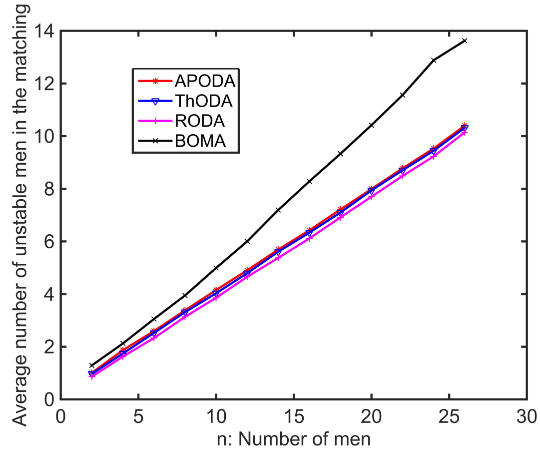
f^{BOMA} is same as f^{ThODA} except that in lines 5 and 7 of Algorithm 2, we use Max-wt-Bipartite$(M(t), W(t))$ and Max-wt-Bipartite$(DM(t), W(t))$ respectively instead of the DA. Max-wt-Bipartite(A, B) gives maximum weight bipartite matching between men A and women B. An edge between a man m and woman w has weight $2n + 2 - rank_m(w) - rank_w(m)$.

We now perform an empirical study of the proposed mechanisms in the next section.

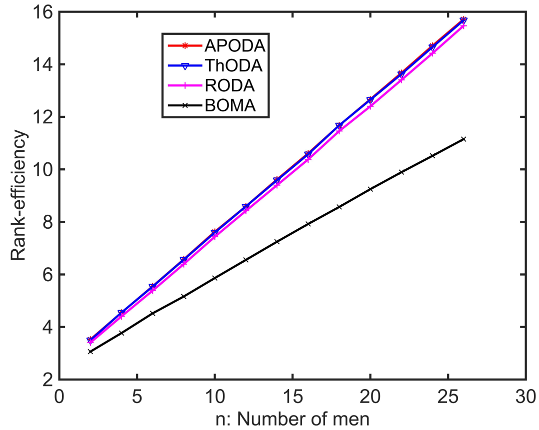
6 Evaluation of the Mechanisms

6.1 Empirical Evaluation

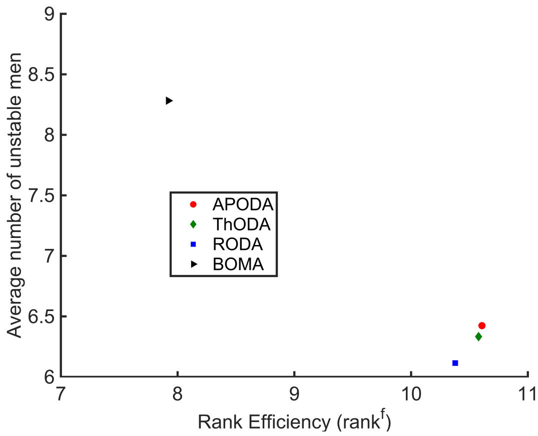
In this section, we describe the simulations that we carried to measure stability and rank-efficiency of the proposed mechanisms. We assume the men arrive into the system according a



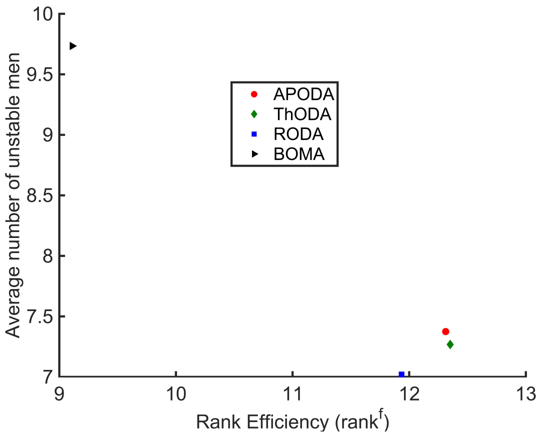
(a) Stability of the four mechanisms for various n with $\lambda = 3, \mu = 0.5$



(b) Rank-efficiency of the four mechanisms for various sn with $\lambda = 3, \mu = 0.5$



(c) Scatter Plot for rank-efficiency and stability of the four mechanisms for $n = 20, \lambda = 3, \mu = .05$



(d) Scatter Plot for rank-efficiency and stability of the four mechanisms for $n = 24, \lambda = 5, \mu = .05$

Figure 2: Comparison of $f^{APODA}, f^{ThODA}, f^{RODA}, f^{BOMA}$

Poisson process with parameter λ and wait in the system according to a exponential distribution with mean $\mu = 0.5$. The preferences of every man and woman are drawn uniformly at random from all possible preference profiles. With these setup, we generated 5,000 instances of matching problem for each of various n, λ combinations and measured stability and rank-efficiency across these instances.

To study stability, we measure the average number of unstable men produced by the mechanism. We say a man is *unstable* if he is involved in at least one blocking pair.

Figure 2 (a) and (b) shows how the stability and rank-efficiency changes for all the mechanism as we increase n by fixing $\lambda = 3$ Figure 2 (c) is a scatter plot for stability vs rank-efficiency of all the mechanisms for $n = 20, \lambda = 3$ and (d) is a scatter plot for $n = 24, \lambda = 6$.

We discuss compare all the proposed mechanisms in the next subsection.

6.2 Discussion

We designed four mechanisms for two sided matching problem when the men side is dynamic. Based on a design goal, we can choose which mechanism to be used. If strategy-proofness is important, we propose to use any mechanism from a class f^{PODA}. These are partition mechanism in which a feasible partition of the men is given. If men are impatient or if it is difficult to know how to partition, one can use $f^{APODA} \in f^{PODA}$. If underlying stochastic model is available, we propose to use $f^{ThODA} \in f^{PODA}$. It improves the stability by 2%-5% over f^{APODA}.

If stability is of utmost important, we recommend to use f^{RODA} which is progressively stable and performs better than other mechanisms on conventional stability. If the design goal is to improve rank-efficiency, we recommend to use f^{BOMA} that improves rank-efficiency by 25% but the average number of unstable men increases by 35%. This is summarized in Table 3. The numbers in a row indicate ranking among the four mechanisms on respective performance measure.

	f^{APODA}	f^{ThODA}	f^{RODA}	f^{BOMA}
Strategy-proof	Y	Y	N	N
Stability	3	2	1	4
Progressive Stability	N	N	Y	N
Rank-efficiency	3	4	2	1

Table 3: Notation

7 Conclusion

Motivated by an expert crowdsourcing market, in this paper, we addressed two sided dynamic matching problem when one side, the men side, is dynamically arriving to the market where as the other side, the women side, is available for matching from the start. We focused on exogenous men settings. We first proposed strategy-proof mechanisms, f^{PODA}, f^{APODA}, f^{ThODA}. As it is impossible to achieve stability in online settings, we introduced a weaker notion of stability, namely progressive stability. We proposed a mechanism f^{RODA} that achieves the progressive stability if all the agents are truthful. However, f^{RODA} is not strategy-proof. We also proposed a mechanism f^{BOMA} to improve the rank-efficiency, but it has poor stability and is not strategy-proof. In the previous section we compared all the mechanisms empirically. Based on a design goal, one can choose an appropriate matching mechanism.

References

[1] Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. Dynamic matching market design. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 355–355. ACM, 2014.

[2] Satyanath Bhat, Swaprava Nath, Onno Xoeter, Sujit Gujar, Yadati Narahari, and Chris Dance. A mechanism to optimally balance cost and quality of labeling tasks outsourced to strategic agents. In *Thirtheenth International Conference on Autonomous Agents and Multiagent Systems*, pages 917–924, 2014.

[3] E. Budish and E. Cantillon. Strategic behavior in multi-unit assignment problems: Lessons for market design. Technical report, Harvard Business School, 2009.

[4] Xi Chen, Qihang Lin, and Dengyong Zhou. Optimistic knowledge gradient policy for optimal budget allocation in crowdsourcing. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 64–72, 2013.

[5] O. Compte and P. Jehiel. Voluntary participation and re-assignment in two-sided matching. Working papers, Paris School of Economics, October 2008.

[6] Djellel Eddine Difallah, Gianluca Demartini, and Philippe Cudré-Mauroux. Pick-a-crowd: tell me what you like, and i'll tell you what to do. In *Proceedings of the 22nd international*

conference on World Wide Web, pages 367–374. International World Wide Web Conferences Steering Committee, 2013.

[7] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, pages 9–15, 1962.

[8] Sujit Gujar and David C Parkes. Dynamic matching with a fall-back option. In *ECAI*, pages 263–268, 2010.

[9] Chien-Ju Ho, Shahin Jabbari, and Jennifer W Vaughan. Adaptive task assignment for crowdsourced classification. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 534–542, 2013.

[10] Chien-Ju Ho and Jennifer Wortman Vaughan. Online task assignment in crowdsourcing markets. In *AAAI*, 2012.

[11] Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006.

[12] Panagiotis G. Ipeirotis, Foster Provost, and Jing Wang. Quality management on amazon mechanical turk. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, HCOMP '10, pages 64–67, New York, NY, USA, 2010. ACM.

[13] David R Karger, Sewoong Oh, and Devavrat Shah. Budget-optimal crowdsourcing using low-rank matrix approximations. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 284–291. IEEE, 2011.

[14] R. M. Karp, U. V. Vazirani, and V. V. Vazirani. An optimal algorithm for on-line bipartite matching. In *Proc. STOC'90*, pages 352–358, 1990.

[15] Dipjyoti Majumdar. Ordinally bayesian incentive compatible stable matchings. Working Papers 05001, Concordia University, Department of Economics, August 2003.

[16] David Manlove. *Algorithmics of matching under preferences.* World Scientific Publishing, 2013.

[17] Vikas C. Raykar, Shipeng Yu, Linda H. Zhao, Gerardo Hermosillo Valadez, Charles Florin, Luca Bogoni, and Linda Moy. Learning from crowds. *J. Mach. Learn. Res.*, 11:1297–1322, August 2010.

[18] Alvin E Roth. The economics of matching: Stability and incentives. *Mathematics of operations research*, 7(4):617–628, 1982.

[19] Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis.* Number 18. Cambridge University Press, 1992.

[20] Tayfun Sönmez and M Utku Ünver. Matching, allocation, and exchange of discrete resources. *Handbook of Social Economics*, 1:781–852, 2011.

[21] Long Tran-Thanh, Sebastian Stein, Alex Rogers, and Nicholas R Jennings. Efficient crowdsourcing of unknown experts using multi-armed bandits. In *European Conference on Artificial Intelligence*, pages 768–773, 2012.

A Local Search Algorithm for SMTI and its extension to HRT Problems

Danny Munera[1], Daniel Diaz[1], Salvador Abreu[2], Francesca Rossi[3], Vijay Saraswat[4], and Philippe Codognet[5]

[1] University of Paris 1 / CRI
Danny.Munera@malix.univ-paris1.fr, Daniel.Diaz@univ-paris1.fr
[2] University of Évora / LISP / CRI
spa@di.uevora.pt
[3] University of Padova / Harvard University
frossi@math.unipd.it
[4] IBM TJ Watson Research Center
vsaraswa@us.ibm.com
[5] JFLI-CNRS / UPMC / University of Tokyo
codognet@is.s.u-tokyo.ac.jp

Abstract. Hospitals/Residents with Ties (HRT) forms a class of problems with many applications, some of which are of considerable size. Solving these problems has been shown to be NP-hard. In previous work, we developed a local search algorithm which displays very high performance in solving Stable Matching with Ties and Incomplete lists (SMTI) problems. In this paper, we propose a method to tackle HRT problems with a slightly modified version of our SMTI solver. We describe our method and provide an initial performance assessment, which turns out to show that the resulting solver can deal with significant HRT problems, providing optimal solutions in most cases, in a very short time.

1 Introduction

In 1962, Gale and Shapley introduced the Stable Matching (SM) problem [4]. An SM instance of size n involves a set of n men and a set of n women, each of whom has ranked all members of the other set in strict order of preference. Solving such a problem consists of finding a *matching*, i.e. a one-to-one matching between the men and the women. In addition the matching must be *stable*, meaning that there is no man-woman pair where both would rather marry each other than their current partner – such a pair is called a *blocking pair*. Gale and Shapley proved that such a stable matching always exists and proposed an $O(n^2)$ algorithm (called GS in what follows) to find one.

However, requiring *each* member to rank *all* members of the opposite sex in a *strict* order is too strong a restriction for many real-life, large-scale applications. A natural variant of SM is the *Stable Matching with Ties and Incomplete Lists* (SMTI) problem [14, 17]. In SMTI, the preference lists may include ties (to express indifference) and may be incomplete (to express that some partners are unacceptable). More formally, an SMTI instance of size n consists of n men and n women, and a preference list for each of them, which contains some of the people of the other gender. Such preference lists are weak orders, that is, total orders possibly containing ties. Given an SMTI instance, a *matching* M is a set of pairs (m, w) representing a (possibly partial) one-to-one matching of men and women. If a man m is not matched in M (i.e. for no w is it the case that $(m, w) \in M$), we say that m is *single* in M (similarly for women). The *size* of a matching M is the cardinality of M, denoted $|M|$.

With the introduction of ties in the preference lists, three different notions of stability may be used [10, 17, 14]. As we consider only *weak stability* (the most challenging), we simply call it *stability*. In the context of M, a pair (m, w) is a *Blocking Pair* (BP) iff (a) m and w accept each other and (b) m is either single in M or strictly prefers w to his current wife, and (c) w is either single in M or strictly prefers m to her current husband. A matching M is stable iff it has no blocking pairs.

A (weakly) stable matching always exists and can be found with variants of the GS algorithm. Since any given SMTI instance may have stable matchings of different sizes, a natural requirement is to find one of maximum cardinality. This optimization problem (called MAX-SMTI) has many real-life applications [17, 14] and has attracted a lot of research in recent years because of that: car sharing or bipartite market sharing, job markets and social networks. Many of these applications involve very large sets. Unfortunately, the MAX-SMTI problem has been shown to be NP-hard, even for very restricted cases (e.g. only men declare ties, ties are of length two, the whole list is a tie) [13, 17].

We have recently proposed a Local Search (LS) algorithm for the SMTI problem [19]. For this, an SMTI problem is first modeled as a permutation problem and then solved by the *Adaptive Search* (AS) method [1, 2]. Basically, starting from a random matching, our algorithm iteratively tries to improve the current matching by performing a swap between two variables (i.e. two men exchange their partner). For this, a limited neighbourhood is explored and the most promising swap is selected based on a heuristic which selects the most significant blocking pair to fix and/or a single man to marry. The algorithm stops when a *perfect matching* is found (a stable matching with no singles) or when a given timeout is reached (in which case the best matching found so far is returned). This algorithm turned out to have very high performance and is able to optimally solve several large instances.

Another very useful variant of SM is the *Hospitals / Residents problem with Ties* (HRT) [12, 11, 17]. An HRT instance consists of two sets: the residents $R = \{r_1, \dots r_{n_1}\}$ who apply to the hospitals $H = \{h_1, \dots h_{n_2}\}$. The preference list of a resident $r_i \in R$ consists of the ordered list of *acceptable* hospitals (a subset of H). The preference list of a hospital $h_j \in H$ contains the ordered list of residents (a subset of R) who consider h_j acceptable. All preference lists are allowed to contain ties. In addition, each hospital $h_j \in H$ has a capacity c_j indicating the maximum number of positions it offers.

The problem consists of finding a stable matching between residents and hospitals satisfying both the preference lists (the matching must be stable) and the capacities (each resident being assigned to at most one hospital and the number of residents assigned to any hospital h_j must not exceed c_j). At any stage during the matching process, a hospital h_j with a_j assignees is said to be *over-subscribed* if $a_j > c_j$, *full* if $a_j = c_j$, and *under-subscribed* if $a_j < c_j$.

The previously discussed notion of weak stability can be adapted to HRT: in the context of M, a pair (r, h) give rise to a blocking pair iff (a) r and h accept each other and (b) r is either unassigned in M or strictly prefers h to his assigned hospital, and (c) h is either under-subscribed or strictly prefers r to the worst resident assigned to it. As for SMTI, a matching M is stable iff it has no blocking pairs.

HRT problem has many practical applications, e.g. assignment of applicants to positions in job markets. In the medical employment domain, there are national programs in various countries such as the Scottish Foundation Allocation Scheme (SFAS), the Canadian Resident Matching Service (CARMS) or the National Resident Matching Program (NRMP) in the USA. Obviously, such programs involve very large sets. Unfortunately, as for SMTI, the problem of finding a stable matching of maximum cardinality (called MAX-HRT) for a given instance of HRT is NP-hard (even for restricted cases, e.g. if the ties are only allowed on one side). Finding an efficient algorithm to solve HRT problems is thus a true challenge with many real applications.

We deem it interesting to see if we can attack the HRT problem with our LS algorithm. While SMTI is a special case of HRT (where each hospital has capacity one) we chose to adopt a reverse approach, considering an HRT as a special case of SMTI so as to keep the main lines of our algorithm (permutation-based, which ensures a compact memory representation and an implicit modeling of the *all-different* constraint). To this end, we can use the so-called *cloning technique* [9] which basically consists of creating c_j copies of the hospital h_j, each of capacity 1, and to use these copies (inside a tie) each time this hospital is referenced in a resident's preference list. Strictly speaking, the resulting problem instance is not exactly an STMI instance since the resulting sets (residents and cloned hospitals) can have different sizes but it is trivial to add dummy elements (residents or hospitals). The extension of our algorithm to deal with this feature is very simple. All of this makes HRT and SMTI equivalent problems.

This *RISC*-like approach is analogous to what occurs with SAT modeling: the object formulation is often voluminous and cumbersome but its resolution by the best SAT solvers is very efficient – often faster than what is obtained with dedicated solvers which take higher-level formulations, such as CSP or constraint programming. Upon dealing with hard problem instances, we try to improve the solver at low level, meaning that the techniques which we may come up with to better solve HRT will also benefit SMTI, in the general case. We stress that our LS algorithm gets much better performance than complete methods (i.e. enumerative, branch and bound, linear and integer programming, etc.), even though we do not always reach the optimum solution.

As the rest of this paper will show, we manage to get very competitive performance on real-world data sets of considerable size and difficulty. We are also convinced that this will further benefit from the efficiency improvements we explored in [19], namely parallelism.

2 A Local Search Method for SMTI

Local search is a meta-heuristic method for solving optimization problems. It requires a cost function to evaluate the quality of a given assignment of variables (i.e. a *configuration*). The method also needs a transition function which defines, for each configuration, a set of neighbours. The simplest Local Search algorithm starts from a random configuration, explores the neighbourhood, selects a promising neighbour and moves to it. This iterative process continues until a solution is found. In this paper, we use a Local Search method developed by our team: the Adaptive Search method [1].

Adaptive Search (AS) is a generic, domain-independent, constraint-based local search method. This meta-heuristic takes advantage of the modeling of the problem in terms of constraints and variables, in order to guide the search more precisely than a single global cost function.

The error function in AS is a heuristic value which stands for the degree of satisfaction of the constraints. The method combines the error for each constraint to obtain a global cost and then, for each variable, AS projects constraint errors on the involved variables. AS repairs the *worst* variable (highest error) with the *best* (most promising) available value.

AS also includes an adaptive memory inspired by Tabu Search [8] in which each variable leading to a local minimum is marked and cannot be selected for the next few iterations. A local minimum is a configuration for which none of the neighbours improve the current cost. Finally, the algorithm also includes partial resets in order to escape stagnation around local minima.

For this work we use a particular implementation of AS, specialized for permutation problems. In this case all n variables have the same initial domain of size n and are subject to an implicit *all-different* constraint.

2.1 AS Model for SMTI Problems

We recently developed an efficient Adaptive Search model to solve SMTI problems (AS-SMTI). In this section we sketch the main features of the modeling; the interested reader may refer to [19] for details.

To use AS, we model the SMTI problem as a permutation problem: we define a sequence of n variables $(X_1 \dots X_n)$ which take for values permutations of the vector $(1 \dots n)$. $X_i = j$ is interpreted as either $(m_i, w_j) \in M$, or m_i is single if w_j is not on its preference list. Note that this interpretation remains valid when the values of any two variables are swapped (this is how value assignment is implemented in permutation problems).

The AS method seeks to improve the stability of a matching by removing blocking pairs (BPs). Some BPs may be useless in that fixing them does not improve things since the man involved remains part of another BP. To avoid this, the method focuses only on the so-called *undominated blocking pairs* [15, 5]. Let (m, w) and (m, w') be BPs. BP (m, w) dominates (from the men's point of view) BP (m, w') iff m prefers w to w'. A BP (m, w) is *undominated* iff there is no other BP dominating (m, w). In the following we only consider *undominated* BPs, which we simply call BPs.

Adaptive Search relies on a global objective function (called cost function) to measure the degree of error of a configuration. The cost function of a matching M is defined as follows:

$$\mathit{cost}(M) = \#BP(M) \times n + \#\mathit{Singles}(M)$$

where $\#BP(M)$ is the number of BPs in M, and $\#\mathit{Singles}(M)$ is the number of single men in M. The number of BPs is weighted with n to prioritize stable matchings over matchings with fewer singles. A matching M is stable iff $cost(M) < n$, and *perfect* iff $cost(M) = 0$. AS stops as soon as the cost function reaches 0 or when a given time limit is hit, in which case it returns the best matching found so far.

The AS-SMTI modeling defines the function $R(w, m)$ as the *rank* of m in the preference list of w, ranging over $1..(n+1)$, with $i < j$ implying w prefers (man with rank) i to (man with rank) j, and $R(w, m) = n + 1$ iff m is not in the preference list of w. The implementation does some straightforward pre-computation to avoid the linear cost of recomputing $R(w, m)$. The algorithm computes the BPs in the match M, going through all men in the problem. For each men m, let w the current partner of m such that $(m, w) \in M$, AS-SMTI verifies in the preference list of m if there is a woman w' with a higher level of preference than w. If w' exists, let m' the current partner of w' such that $(m', w') \in M$, the algorithm checks in the preference list of w' if the man m has a higher level of preference than m'. If this happens, the algorithm has found the BP (m, w'). The associated error for this BP is determined by the following expression: $R(w', m') - R(w', m)$. Thus, the further the assigned man is from the BP (in the preference list of w'), the larger the error.

It is worth noticing that while (undominated) BPs are considered from the men point of view, the associated errors are computed from the women point of view. Since preference lists can include ties, a man can be implied in several undominated BPs. For efficiency and simplicity reasons, AS-SMTI only considers the first encountered BP for a given man m and computes its error as explained above. Other strategies exist to aggregate the error associated to all BPs (select the maximum error, the average error, randomly one error, . . .).

At each iteration, AS selects the "worst" variable from the current matching M to improve it (the man involved in the BP with the largest error as explained above). In case of several man have the same highest cost, one is selected randomly. AS then fixes the culprit by swapping X_m and $X_{m'}$. In short, AS considers all BPs, chooses the variable corresponding to the worst one, fixes it by moving to a new configuration and re-evaluates the cost of the resulting matching. This heuristic avoids the cost of fixing all BPs, one by one.

In most cases, the resulting matching improves on the current one, and AS continues iteratively. When this is not the case, AS has reached a minimum (global or local). As AS has no way of knowing when the optimum has been reached (except when the cost is 0) it handles both cases similarly trying to escape the minimum invoking a *"reset"* procedure. This procedure slightly alters the current assignment of variables, trying to fix the 2 worst BPs and/or to assign a woman to a single man. The *reset* procedure is stochastic; it will *also* fix the second worst variable with a probability p: good results are obtained with a high value, e.g. $p \simeq 0.98$. This procedure turns out to be very effective: while preserving most of the configuration (no more than 2 swaps are performed), it enables AS to escape all local minima and reach very good solutions.

AS stops when one of the following conditions is reached: (a) a perfect solution has been found (i.e. cost = 0), (b) a given *target cost* T is reached (it is possible to ask the solver to find a solution with at most T singles) or (c) a timeout is reached (in which case the best solution found so far is returned).

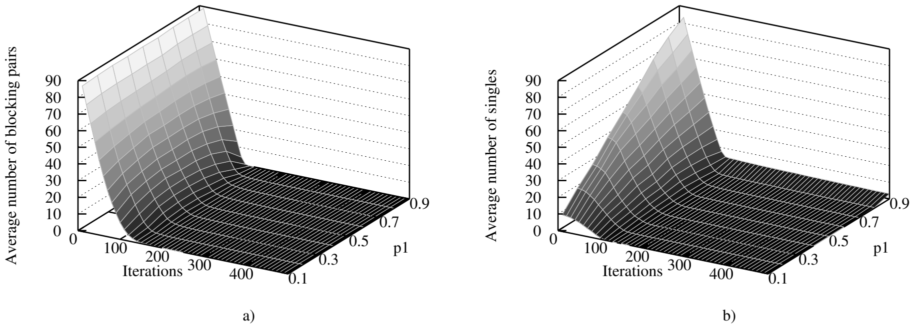


Fig. 1: a) average number of BP b) average number of singles when solving SMTI problems ($n = 100$, $p2 = 0.5$, varying $p1$)

2.2 Performance evaluation on SMTI problems

We present here an evaluation of our AS-SMTI algorithm. For this, we used a dataset composed of random problems created using the generator described in [6] which takes three parameters: the size (n), the probability of incompleteness ($p1$) and the probability of ties ($p2$). Given a triple, $(n, p1, p2)$, a SMTI problem instance with n men and n women is generated as follows: for each man and woman, the algorithm generates a random permutation of size n, as a preference list. Then, the algorithm iterates over each object in the preference lists, and with a probability $p1$, this object is *deleted* from the preference list. Finally, the algorithm iterates again over each remaining object (in the men and women preference lists) and with a probability $p2$, a *tie* is created between the current object and the previous one.

For a given combination of $p1$ and $p2$ 100 different problems were generated. Since AS is a stochastic procedure, each problem is solved 50 times and results are averaged. We used an X10 implementation of AS running sequentially on an AMD Opteron 6380 clocked at 2.5 GHz, i.e. using only one core.

The first experiment analyzes the number of iterations needed to solve an SMTI problem. Every 10 iterations, the solver reports the number of BPs and the number of singles of the current configuration. Due to space limitation, we here consider problems of size 100, for different values of $p1$ in $[0.1, 0.9]$ and for $p2 = 0.5$ (ties may appear in both sides). Figure 1 presents the averaged results of this experimentation. It appears that the average number of BPs quickly decreases. For instance, on average, after 200 iterations, a stable matching is already reached in $99,88\%$ of the cases. It is worth noticing that some difficult instances can require more iterations (the maximum observed has been 460 iterations for a problem generated with $p1 = 0.5$). Figure 1 b) shows the evolution of the number of singles with respect to the number of iterations. Again, this number quickly decreases. It is worth noticing that when the incompleteness of the problem is high (e.g. when $p1 = 0.9$), some problems do not have a perfect solution and the number of singles does not fall below some boundary value.

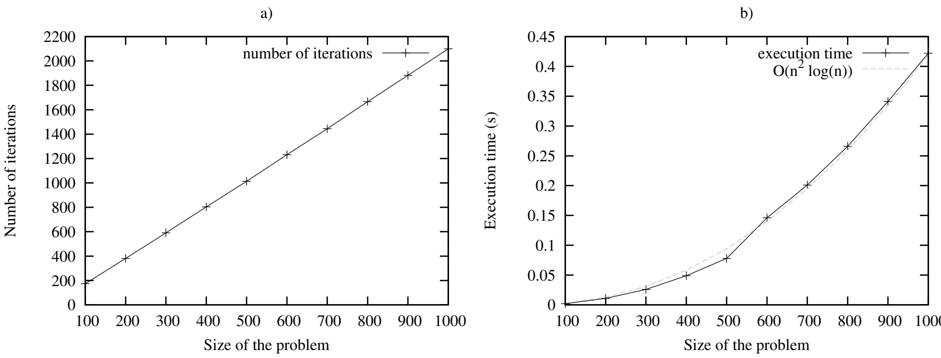
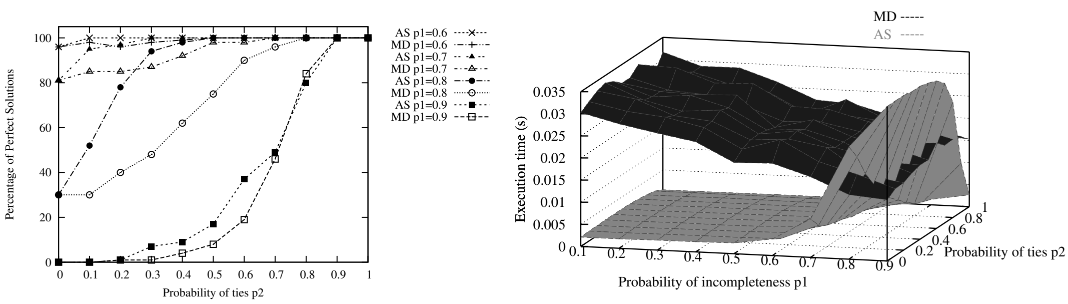


Fig. 2: Runtime behaviour of AS-SMTI: a) number of iterations, b) execution time (varying the size of the problem, p1=0.5 p2=0.5)

The second experiment analyzes the scalability of the AS-SMTI algorithm. For this we fixed the parameters $p1$ and $p2$ to 0.5, and we varied the size n of the problem in the range of $[100, 1000]$ using steps of 100. Figure 2 shows the curves corresponding to the number of iterations and to the runtime when varying n. We can observe that the number of iterations to obtain a perfect solution for SMTI problems varies linearly from 200 iterations for $n = 100$ to 2000 iterations for $n = 1000$, with a corresponding runtime in $O(n^2 \log n)$. Obviously, the results of this test cannot be generalized and we plan to extend the experimentation with other values for $p1$ and $p2$.

Finally, we compared our AS solver with McDermid's method (MD) [18], a very efficient 3/2-approximation algorithm, as implemented in [20]. For this test we used a data set composed of SMTI problems of size $n = 100$, with $p1$ ranging over $[0.1, 0.9]$ and $p2$ over $[0, 1]$, with step 0.1. With MD, for each $(p1, p2)$ pair, we solved the 100 instances once and averaged the the execution time.



(a)

(b)

Fig. 3: AS vs. MD: a) quality of solutions b) execution time.

Figure 3a compares the quality of solutions. The percentage of perfect stable matchings found by the AS algorithm is considerably higher than those found by MD, in particular using a probability of ties $p2 \in [0.1..0.8]$.

Figure 3b compares the execution times, as a 3D chart. In many cases, AS is up to an order of magnitude faster than MD. With higher probability of incompleteness (e.g. $p1 = 0.9$), MD outperforms AS. This can be explained by the time-complexity of MD which is proportional to the total length of the preference lists, i.e. it linearly decreases as $p1$ increases. The MD algorithm seems to perform faster than our AS approach only when $p1 = 0.9$. We note that MD always returns the same, single and (sub)optimal solution, while AS will yield more than one solution, with observably better quality. Moreover, a *solution quality* vs. *performance* trade-off is always possible in AS, by tweaking the timeout parameter. A complete comparison of the AS-SMTI algorithm with a state-of-the-art Local Search method [5] and a SAT encoding of the SMTI problem [7] may be found in [19].

3 Solving HRT Problems

In this section, we propose an algorithm to solve the HRT problem based on the algorithm presented above (we call this extension AS-HRT). Our goal is to obtain an HRT solver with *minimum changes* in the AS-SMTI algorithm. For this purpose, AS-HRT resorts to the "cloning" technique described in [3, 9, 21]. The main idea in cloning an HRT problem is to define a match between residents and *positions* instead of hospitals. A position being a single post offered by a hospital (a hospital h_j can offer c_j positions), each position can only be assigned to only one resident (capacity equal to one). To convert an HRT problem into an SMTI formulation, we create a new set of positions composed of the single positions offered by the hospitals. Each position has the same preference list as its "root" hospital. In the residents' preference lists, each hospital h_j is replaced by a sequence composed of the associated c_j positions (all forming a tie). The resulting equivalent SMTI problem consists of matching residents and positions. The main drawback of the cloning process is a significant increase in the size of the problem (but this remains manageable with modern computers).

Using cloning we may convert an HRT problem into an (asymmetric) SMTI problem, in a polynomial time. We use the term "asymmetric" because the resulting SMTI problem can have sets of different sizes. More formally, an *asymmetric SMTI* problem is specified by (a) two sets M and W of cardinality m and n respectively, (b) a ranking function $R : M \times W \to \{1 \ldots n + 1\}$ and $R : W \times M \to \{1 \ldots m + 1\}$ (we use the same name R for the ranking function for men and women). Note that the only point of generality over the standard SMTI formulation is that m is not required to be identical to n. From the AS-SMTI algorithm this comes down handling a vector of $max(m, n)$ values with some "dummy" values for missing elements.

It is interesting to characterize the c_j clones in the resulting SMTI problem since they play the same role (they are interchangeable). Given an SMTI instance, $m_1, m_2 \in M$ are said to be *equivalent* (written $m_1 \sim m_2$) if $\forall w, R(m_1, w) = R(m_2, w)$. Similarly, $w_1, w_2 \in W$ are said to be equivalent if $\forall m, R(w_1, m) = R(w_2, m)$. Note that $\sim$ is an equivalence relation.

A given SMTI problem may have equivalent elements or not. When we translate an HRT problem into SMTI we get equivalences: the c_j elements corresponding to a hospital h_j are all equivalent to each other.

When walking through several matchings, it would be nice to avoid exploring equivalent matchings (i.e. those whose difference only concerns equivalent elements). Indeed, only one matching from an equivalence class needs to be examined. It is worth observing that given a matching M if (m,w) and $(m',w') \in M$ give rise to a blocking pair (m,w'), then $m \not\sim m'$ and $w \not\sim w'$. Therefore, the swap executed by the main loop of the AS-SMTI algorithm to fix a BP already ensures that a matching is not changed to an equivalent matching. However, it is not the case of the reset procedure invoked to escape a local minimum which performs some random swaps. We have not yet improved this point (if a reset procedure swaps two equivalent elements the local minimum is not escaped and another reset will occur resulting in a waste of time). To be fully aware of equivalences in the current AS-SMTI algorithm, we should first compute equivalences upfront (this is an easy linear operation) and avoid swaps between equivalent elements in the reset procedure. We plan to do this in a second version of the algorithm.

3.1 Implementation

To implement the extension to HRT problems, we developed a pre-processor and a post-processor (see Figure 4). The pre-processor converts an HRT problem instance I into an equivalent SMTI problem instance I'. The post-processor systems takes the match found by the AS-SMTI solver for I' and converts back into HRT match form. The resulting HRT match is the solution to the initial problem.

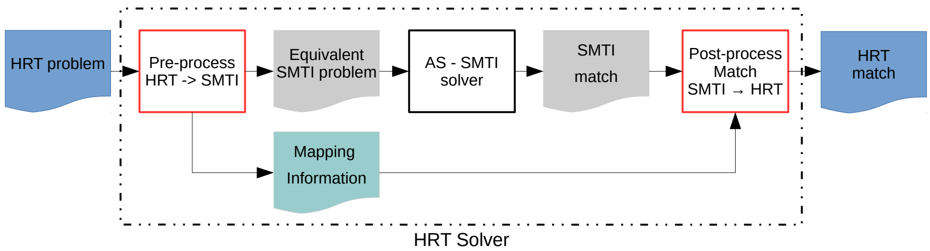


Fig. 4: Description of the solver extension to solve HRT problems

3.2 Preliminary Performance Evaluation

We did preliminary experiments to assess the performance of the HRT extension to AS-SMTI. We used a data set composed by randomly generated problems,[6] with the the same parameters as in [16]:

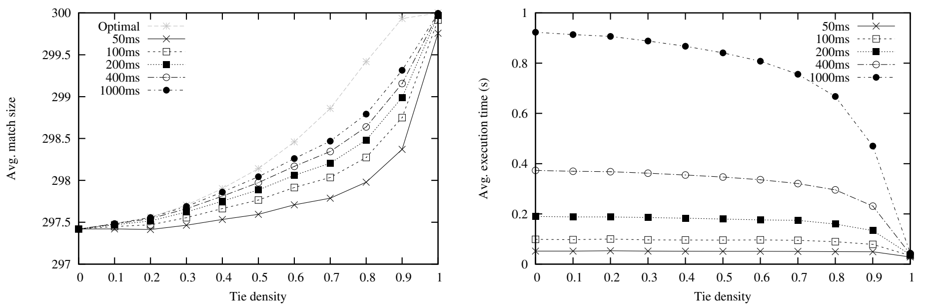
– Number of residents $n1 = 300$ (size of the problem).

[6] The data set was kindly provided by Augustine Kwanashie and David F. Manlove from the University of Glasgow, UK, and is the same as that in [16].

– Number of hospitals $n2 = 21$.
– Length of the residents' preference list $l = 5$.
– Total number of positions $C = 300$.
– Tie density td ranging over $[0, 1]$ with step 0.1.

For each td value, we used 10000 instances. The experiment was carried out on a machine with 4×16-core AMD Opteron 6380 CPUs, running at 2.5 GHz and 128 GB of RAM, using only 1 core. We solve each problem instance once, collecting the maximum size of the match and the execution time. We tested the timeout for AS-SMTI solver at $50, 100, 200, 400$ and $1000ms$.

Figure 5a shows the maximum size of the match found by AS-HRT, varying the tie density and using different timeout limits. We also include the optimal match size found with the Integer Programming method developed in [16]. AS-HRT almost reaches the optimal match size when using low values of the tie density ($td < 0.6$) and when the tie density is 1. These results are obtained even with low timeout values, i.e. $50ms$. When the tie density is between 0.6 and 0.9, AS-HRT does not always reach the optimal solution: yet in these cases, the minimum ratio to the optimal size we obtained was 0.998 times (for the $td = 0.9$).



(a) (b)

Fig. 5: AS-HRT: a) match size and b) execution time (size=300, varying td)

Figure 5b presents the average execution time of AS-HRT. The results show that when using $td = 1$, the execution time to get the optimal solution is *almost* constant, at about $40ms$. When the tie density decreases, the average execution time tends to be the same as the chosen timeout value. This behaviour can be explained because, when the optimal solution of a problem instance is not perfect ($|M| < 300$), the AS-SMTI method, which doesn't know this, will keep trying to improve on it until the timeout is reached.

4 Conclusions and Future Work

We recently developed a Local Search solver for SMTI problems based on the Adaptive Search method. To use this, we need to model problems as permutations and provide some heuristics based on the study of relevant blocking pairs to guide the search process, in order to iteratively improve the current matching. A reset procedure is invoked when the solver is trapped in local minima. This solver is very efficient and can solve optimally large instances quickly. For the most difficult problems, it is possible to tune the trade-off between solution quality and solving performance by tweaking the timeout parameter.

On the top of this solver, we have built a solver for HRT which basically maps an HRT problem to an equivalent SMTI problem using the *cloning technique*. This required a slight modification to the core solver. We also characterized the resulting clones with the notion of *equivalence* which captures the fact that clones are interchangeable and thus it is not desirable to replace a matching by another which is equivalent (i.e. one that only differs in equivalent elements). We showed that the core LS algorithm mainly satisfies this property, save in the reset procedure. We have plans to improve on this situation.

We presented a preliminary experimental evaluation based on thousands of problems of size 300. The solver already performs very well: using a timeout of $1s$, it reaches the optimal solution for most of the instances. For the most difficult instances, when the optimum is not reached within this timeout, the returned solutions are very good with size within a factor 0.998 w.r.t. the optimal solution. We plan to experiment with other dataset reflecting a more realistic case of the HRT problem, e.g. modeling the popularity or unpopularity of the hospitals. Moreover, we also plan to use larger problem instances, as a future development.

There are several avenues for improvement, of which we name a few: we already mentioned how to take equivalence into account in the reset procedure. It is also possible to see if other problem reduction techniques are fruitful (e.g. those mentioned in [16] for the IP formulation). It would also be interesting to start from a pertinent solution, instead of a pure random assignment: this could be done with the help of a very fast approximation algorithm. The size of this solution could be also used as a lower bound.

Finally, as our base method is amenable to massive parallelisation, we will explore parallelism to tackle both hard and large problem instances, as we already did in [19].

Acknowledgments

We would like to thank Augustine Kwanashie and David F. Manlove from the University of Glasgow, for kindly providing us with the extensive HRT data sets.

References

1. P. Codognet and D. Diaz. Yet Another Local Search Method for Constraint Solving. In K. Steinhöfel, editor, *Stochastic Algorithms: Foundations and Applications*, pages 342–344. Springer Berlin Heidelberg, London, 2001.
2. P. Codognet and D. Diaz. An Efficient Library for Solving CSP with Local Search. In *5th international Conference on Metaheuristics*, pages 1–6, Kyoto, Japan, 2003.
3. L. Dubins and D. Freedman. Machiavelli and the Gale-Shapley Algorithm. *American mathematical monthly*, 88(7):485–494, 1981.
4. D. Gale and L. Shapley. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.
5. M. Gelain, M. Pini, F. Rossi, K. Venable, and T. Walsh. Local Search Approaches in Stable Matching Problems. *Algorithms*, 6(4):591–617, Oct. 2013.

6. I. Gent and P. Prosser. An Empirical Study of the Stable Marriage Problem with Ties and Incomplete Lists. In *in ECAI 2002*, pages 141–145. IOS Press, 2002.
7. I. Gent, P. Prosser, B. Smith, and T. Walsh. SAT Encodings of the Stable Marriage Problem with Ties and Incomplete Lists. In *SAT*, volume 8, pages 133–140, Cincinnati, USA, 2002.
8. F. Glover and M. Laguna. *Tabu Search.* Kluwer Academic Publishers, July 1997.
9. D. Gusfield and R. Irving. *The Stable Marriage Problem: Structure and Algorithms.* MIT press, 1989.
10. R. Irving. Stable Marriage and Indifference. *Discrete Applied Mathematics*, 48(3):261–272, Feb. 1994.
11. R. Irving and D. Manlove. Finding Large Stable Matchings. *ACM Journal of Experimental Algorithmics*, 14(September), 2009.
12. R. Irving, D. Manlove, and S. Scott. The Hospitals/Residents Problem with Ties. In *7th Scandinavian Workshop on Algorithm Theory, SWAT'00 Lecture Notes in Computer Science*, number i, pages 259–271, Berlin, 2000.
13. K. Iwama, D. Manlove, S. Miyazaki, and Y. Morita. Stable Marriage with Incomplete Lists and Ties. In *In Proceedings of ICALP 99: the 26th International Colloquium on Automata, Languages and Programming*, number ii, pages 443–452. Springer-Verlag, 1999.
14. K. Iwama and S. Miyazaki. A Survey of the Stable Marriage Problem and its Variants. In *International Conference on Informatics Education and Research for Knowledge-Circulating Society , ICKS'08*, number i, pages 131–136. IEEE Press, 2008.
15. F. Klijn and J. Masso. Weak Stability and a Bargaining Set for the Marriage Model. *Games and Economic Behavior*, 42(1):91–100, 2003.
16. A. Kwanashie and D. Manlove. An Integer Programming Approach to the Hospital/Residents Problem with Ties. *CoRR*, abs/1308.4:1–10, Aug. 2013.
17. D. Manlove, R. Irving, K. Iwama, S. Miyazaki, and Y. Morita. Hard Variants of Stable Marriage. *Theoretical Computer Science*, 276(1-2):261–279, 2002.
18. E. McDermid. A 3/2-approximation Algorithm for General Stable Marriage. In *International Colloquium on Automata, Languages and Programming, ICALP'2009*, pages 689–700, Rhodes, Greece, 2009.
19. D. Munera, D. Diaz, S. Abreu, F. Rossi, V. Saraswat, and P. Codognet. Solving Hard Stable Matching Problems via Local Search and Cooperative Parallelization. In *AAAI*, 2015.
20. A. Podhradsky. Stable Marriage Problem Algorithms. Master's thesis, Masaryk University, 2010.
21. A. Roth and O. Sotomayor. Two Sided Matching: A Study in Game-Theoretic Modeling and Analysis. *Econometric Society Monographs*, 18, 1990.

Fairness and Efficiency in a Random Assignment: Three Impossibility Results

Alexander S. Nesterov *

February 18, 2015

Abstract

We consider the problem of allocating N indivisible objects among N agents according to their preferences when transfers are absent. We study the tradeoff between fairness and efficiency in the class of strategy-proof mechanisms. The main finding is that for strategy-proof mechanisms the following efficiency and fairness criteria are mutually incompatible: (1) Ex-post efficiency and envy-freeness, (2) ordinal efficiency and weak envy-freeness and (3) ordinal efficiency and equal division lower bound. Result 1 is the first impossibility result for this setting that uses ex-post efficiency; results 2 and 3 are more practical than similar results in the literature. In addition, for $N = 3$ we provide two characterizations of the celebrated random serial dictatorship mechanism: it is the unique strategy-proof, ex-post efficient mechanism that (4) eliminates strict ordinal envy among agents with the same ordinal preferences, or (5) eliminates cardinal envy among agents with the same cardinal preferences (by providing these agents with assignments of equal expected utility). Result 4 strengthens the characterization by Bogomolnaia and Moulin (2001), and result 5 implies the impossibility result by Zhou (1990).

This paper is a short version of Nesterov (2014) and omits all the proofs.

JEL Classification: *C78; D71; D78*

Key words: random assignment; random serial dictatorship; strategy-proofness; ex-post efficiency; weak envy-freeness; equal division lower bound.

1 Introduction

In this paper we study the assignment problem, where a set of indivisible objects is allocated to a set of agents according to their preferences so that each agent receives precisely one object.[1]

Since the formal introduction of the assignment problem by Hylland and Zeckhauser (1979) there has been a search for "nice" mechanisms that would satisfy three major properties: incentive compatibility, efficiency and fairness. Hylland and Zeckhauser (1979) themselves propose a pseudo-market mechanism that optimally satisfies the latter two properties: their mechanism is ex-ante

*WZB, Reichpietschufer 50, 10785, Berlin, Germany; e-mail: alexander.nesterov@wzb.eu

[1] The assignment problem is also known as the one-sided matching problem and the house allocation problem.

efficient and envy-free. However, the reliability of these efficiency and fairness properties is doubtful because the mechanism does not satisfy strategy-proofness − it is not incentive compatible. Therefore, it may not satisfy these properties under *true* preferences.

The further search for a reliably "nice" mechanism gave rise to a series of negative results. It was Gale (1987) who for the first time conjectured that for an assignment problem with at least three agents, no mechanism can satisfy ex-ante Pareto efficiency, strategy-proofness and *anonymity*. (Anonymity is weak notion of fairness that requires that any two agents with identical reported utilities get the same individual random assignment.) Later, Zhou (1990) showed a slightly stronger result, where instead of anonymity he used *symmetry*. (Symmetry is implied by anonymity, it requires that any two agents with identical reported utilities get the same *expected utility* and not necessarily the same random assignments).

Subsequently, in their seminal paper Bogomolnaia and Moulin (2001), hereinafter referred to as BM, show a similar but logically independent impossibility result. BM consider agents with strict *ordinal* preferences over objects (as opposed to utilities in the papers mentioned above). Based on these preferences, BM redefine the efficiency concept: they call a random assignment *ordinally efficient* if it is not stochastically dominated by any other random assignment for all agents simultaneously.[2] Using this criterion, they show the following impossibility result: for the assignment problem with at least four agents, no mechanism can satisfy strategy-proofness, ordinal efficiency and *equal treatment of equals*. (The latter is a weak fairness criterion that requires that agents with the same ordinal preferences get the same random assignments.)[3]

The goal of the current paper is to further study the feasibility set of the "nice" mechanisms. There are five main results in this paper: three impossibilities and two characterizations.

The first result states a general impossibility regarding ex-post efficiency. We show that when there are at least three agents, there is no ex-post efficient, envy-free and strategy-proof mechanism. In fact, Lemma 1 shows an even stronger result, in which envy-freeness and strategy-proofness are substituted by a pair of weaker properties. This result is most relevant for deterministic assignment mechanisms that are usually required to be Pareto efficient and strategy-proof, but they are very unfair ex-post. [4] That is why modifications of these mechanisms may involve randomization in order to restore fairness ex-ante. However, as implied by the impossibility result, in these modifications envy-freeness can only be achieved at the cost of either ex-post efficiency or strategy-proofness.

The second result states that if there are at least four agents, there is no weak envy-free, ordinally efficient and strategy-proof mechanism. (A random assignment is weak envy-free if, for each agent, her own assignment is not strictly stochastically dominated by any other agent's assignment.) Together with the previous impossibility result, it shows the tradeoff between efficiency and fairness in terms of envy. Precisely, given strategy-proofness, when relaxing the fairness criterion from envy-

[2] Ordinal efficiency is also often referred to as *sd-efficiency*. Ex-post efficiency is implied by ordinal efficiency, which in turn is implied by ex-ante efficiency.

[3] Equal treatment of equals implies anonymity but is logically independent from symmetry, since the latter does not require equal random assignments for equals unlike the other two notions.

[4] In fact, ex-post fairness is an extremely restrictive property, as shown Kesten and Yazici (2012).

freeness to weak envy-freeness, the feasibility threshold in terms of efficiency shifts from ex-post efficiency to ordinal efficiency.

This result is very close to the impossibility result in BM, namely the mutual incompatibility of strategy-proofness, ordinal efficiency, and equal treatment of equals. Equal treatment of equals and weak envy-freeness are logically independent, but the latter is arguably more relevant in practice for two main reasons. First, weak envy-freeness applies to the full set of preference profiles, while equal treatment of equals restricts assignments of agents with identical preferences. Second, when equal treatment of equals is applicable, it excessively restricts the random assignment, similarly to envy-freeness. [5]

The third result of the paper is the characterization of the *random serial dictatorship mechanism* (RSD) for the case of three agents. For the case with three agents, BM also characterize RSD as the unique ex-post efficient, strategy-proof mechanism that satisfies equal treatment of equals. We strengthen this result by showing that RSD is the unique strategy-proof and ex-post efficient mechanism that eliminates sd-envy between agents with identical preferences, the property that we call *weak envy-freeness among equals*.[6]

This result implies the characterization by BM. Another implication of our result is that RSD can be characterized as a unique mechanism that is strategy-proof, ex-post efficient and weak envy-free (for all agents).

The fourth result is another characterization of RSD; it is strongly related to the previous result, though logically independent. This time, for $N = 3$ RSD is characterized as the only mechanism that is ex-post efficient, strategy-proof and that satisfies symmetry, the fairness notion used by Zhou (1990). Symmetry is quite similar to weak envy-freeness among equals—agents with identical preferences receive assignments that do not dominate one another—but symmetry is defined in cardinal terms. Since RSD is not ex-ante efficient, this charaterization result implies the impossibility by Zhou (1990). This result also implies the characterization in BM (since symmetry is weaker than equal treatment of equals).

In the last part of the paper we focus on the second most important approach to fairness: the so-called "fare share guaranteed". Here, the agents' assignments are compared not one to another, as in envy-freeness, but to the equal division assignment such that each agent receives each object with equal probability $\frac{1}{N}$. A random assignment that ordinally dominates the equal division is said to satisfy *equal division lower bound*.[7]

In the fifth result of the paper we show that there is no strategy-proof and ordinally efficient mechanism that satisfies equal division lower bound. This result is important for a large class of mechanisms that satisfy equal division lower bound by construction. In these mechanisms, agents always have the opportunity to get at least the equal division assignment. For example, in the

[5] Indeed, one can see equal treatment of equals as envy-freeness for a limited set of agents − only for the agents with identical preferences.

[6] Weak envy-freeness among equals can be seen as a natural relaxation of either the equal treatment of equals or the weak envy-freeness.

[7] An extensive review on comparison to equal division and other notions of fairness for allocation rules can be found in Moulin (2014) and Thomson (2007).

Table 1: Summary of results

		Strategy-proof mechanisms			
		Envy-free	Weak envy-free	Equal division lower bound	Equal treatment of equals
Ex-post efficient	$N = 3$	$\emptyset$ (Theorem 1, BM*)	RSD! (Corollary 2)	RSD	RSD! (Corollary 1, BM)
	$N > 3$	$\emptyset$ (Theorem 1)	RSD (BM)	RSD	RSD
Ordinally efficient	$N > 3$	$\emptyset$	$\emptyset$ (Theorem 2)	$\emptyset$ (Theorem 3)	$\emptyset$ (BM)

Exclamation mark denotes uniqueness, BM stands for Bogomolnaia and Moulin (2001).
*The case of three agents is also mentioned by BM, p.310, though informally.

pseudo-market mechanism proposed in Hylland and Zeckhauser (1979) the agents have equal budgets with which they purchase probability shares of objects at competitive equilibrium prices. As a result, in any feasible random assignment, the budget is sufficient to purchase the assignment that is at least as good as the equal division. Therefore, such mechanism inevitably lacks either ordinal efficiency or strategy-proofness (in fact, the latter is the case since the mechanism is ex-ante efficient, which implies ordinal efficiency).

Despite the negative results presented in this paper we, however, can still hope to find a strategy-proof, fair, and efficient mechanism in some relevant cases. For large markets in which every object has an increasing number of copies (for example, in the school choice setting, one can think of seats in one school as copies of a unique seat; the number of seats grows large while the number of schools remains the same), Che and Kojima (2010) show that RSD is asymptotically ordinally efficient. For a similar large market, Kojima and Manea (2010) show that the probabilistic serial mechanism is asymptotically strategy-proof. Therefore, the impossibility results presented here do not hold asymptotically for these types of large markets.

Some of the results of this paper are limited by the nature of the standard framework that is used. In a more general setting where the number of houses may be higher than the number of agents (especially in the case with a null object), the agents have a richer strategy set and thus one cannot directly transfer the results to that setting. For instance, in such settings RSD is no longer ex-post efficient for some preference profiles; it can also be dominated by another strategy-proof mechanism (see Erdil (2014) for these and other results in the general setting). However, the negative results must hold, since the standard setting is a special case of the general setting.

Table 1 summarizes the main findings of this paper as well as the relevant results of BM.

The paper proceeds as follows: Section 2 introduces the framework, section 3 presents the first impossibility result (Theorem 1), section 4 covers two characterization results (Proposition 1 and Proposition 2), section 5 presents the second impossibility result (Theorem 2), section 6 — the third impossibility result (Theorem 3), and section 6 concludes by discussing the implications of the findings and the remaining open questions.

2 The Model

In this section we introduce the framework: define the assignment problem, the random assignment mechanism and its properties..

Let $A = \{a_1, a_2, ..., a_N\}$ be the set of N agents and $H = \{h_1, h_2, ..., h_N\}$ be the set of N houses. Each agent $a \in A$ is endowed with a strict preference relation $\succ_a$ on H with a corresponding weak preference relation $\succcurlyeq_a$. A set of individual preferences of all agents constitutes a **preference profile** $\succ = (\succ_a)_{a \in A}$. Let $\mathcal{R}$ be the set of all possible individual preferences, and $\mathcal{R}^N$ be the set of all possible preference profiles. In what follows we assume that the sets A and H are fixed and that the house allocation problem is defined by the preference profile $\succ$ only.

Each assignment problem has either a deterministic solution, called matching, or a probabilistic solution, called random assignment. A **random assignment** P is a doubly stochastic matrix of size N. Each element $P_{a,h}$ of the matrix P represents a probability of agent a to be assigned house h. Let $\mathcal{P}$ be a set of all possible random assignments P.

A **matching** μ is a random assignment whose elements can only be zeros or ones, so that μ precisely prescribes which agent receives which house. Let $\mathcal{M}$ be a set of all possible matchings μ. According to the Birkhoff-von Neumann theorem, any random assignment P can be represented as a lottery over the set of matchings $\mathcal{M}$ (but this representation is not necessarily unique). For this reason and since agents care only about their own assignment, we can concentrate on random assignments without specifying the exact matchings that these random assignments correspond to.

In order to be able to compare different random assignments we need the following definitions. A set of houses that agent a weakly prefers to some house h is the **upper contour set of house** h **at** $\succ_a$: $U(\succ_a, h) = \{h' \in H : h' \succcurlyeq_a h\}$. For example, the upper contour set of the most preferred house is always this same house, of the second most preferred house $-$ the two best houses and so forth.

Given the individual random assignment P_a, the overall probability of agent a being assigned some house that is at least as good as house h is her **surplus at** h **under** P_a: $F(\succ_a, h, P_a) = \sum_{h' \in U(\succ_a, h)} P_{a,h'}$. In other words, the surplus at h is the probability of being assigned some object from the upper contour set of h.

An individual random assignment P_a **ordinally dominates** another individual random assignment P'_a **at** $\succ_a$ (denoted by $P_a \geq_a P'_a$) if it first order stochastically dominates it. The equivalent condition is that all surpluses of P_a weakly exceed the surpluses of P'_a: for each $h \in H$ $F(\succ_a, h, P_a) \geq F(\succ_a, h, P'_a)$. A **strict ordinal domination** (denoted by $P_a >_a P'_a$) occurs under the additional condition that the two random assignments are not identical. Finally, a random assignment P is said to **dominate** another random assignment P' if it dominates for all agents simulatneously; P **strictly dominates** P' if it just dominates P' and the assignments are not identical.

2.1 Axioms

Below we introduce the properties of random assignments and *mechanisms* $-$ the systematic procedures that associate each preference profile $\succ \in \mathcal{R}^N$ with some random assignment $P \in \mathcal{P}$: $P = \varphi(\succ)$, where φ denotes a mechanism.

Efficiency. For a matching there is a single definition of efficiency: a matching is **(Pareto) efficient at some preference profile** if it is not dominated by any other matching at this preference profile. A random assignment is **ex-post efficient (ExPE) at a preference profile** if it can be represented as a lottery over efficient matchings. A random assignment is **ordinally efficient (OE) at a preference profile** if it is not stochastically dominated by any other random assignments at this preference profile. A mechanism is said to be **ex-post efficient (ordinal efficient)** if for any preference profile it results in an ex-post efficient (ordinally efficient) random assignment.

Strategy-proofness. A mechanism φ is **strategy-proof (SP)** if at any preference profile no agent can benefit by misreporting her preferences: for each $a \in A$, for each $\succ \in \mathcal{R}^N$ and for each $\succ'_a \in \mathcal{R}$ the following holds: $\varphi(\succ) \geq_a \varphi_a(\succ'_a, \succ_{-a})$. In other words, under a strategy-proof mechanism, truth-telling is always a dominant strategy for every agents.

Now we introduce an auxiliary notion of incentive compatibility which is weaker than strategy-proofness; we use this property for the first impossibility result below. This notion restricts the set of (potentially) profitable strategies for agents. A mechanism is **upper shuffle-proof (USP)** if no agent a can change her surplus at some object h by "shuffling" the objects that are *strictly* better than h (or misreporting the preferences within the upper contour set of h excluding h itself). Formally, for each $a \in A, h \in H$, and for each $\succ \in \mathcal{R}^N, \succ'_a \in \mathcal{R}$ such that $U(\succ_a, h) = U(\succ'_a, h)$, the following holds: $F(\succ_a, h, \varphi_a(\succ)) - \varphi_{ah}(\succ) = F(\succ_a, h, \varphi_a(\succ')) - \varphi_{ah}(\succ')$ (the difference represents the sum of assignment probabilities for houses that are strictly better than h).[8]

Fairness. A random assignment P is **envy-free (EF)** if every agent prefers her assignment to any other agent's assignment: for each $a, a' \in A$ $P_a \geq_a P_{a'}$. A random assignment P is **weak envy-free (wEF)** if no agent strictly prefers some other agent's assignment: there do not exist $a, a' \in A$ such that $P_{a'} >_a P_a$. Another widely used notion of fairness is the **equal treatment of equals (ETE)**: for each $a, a' \in A$ with $\succ_a = \succ_{a'}$ the individual random assignments are identical: $P_a = P_{a'}$. A weaker combination of the previous two properties is called **weak envy-freeness among equals**. A random assignment P is weak envy-free among equals if for any two agents a, a' with identical preferences $\succ_a = \succ_{a'}$ none of them strictly prefers the assignment of the other: $P_{a'} \not>_a P_a$.

Another approach to fairness is the so-called "fair-share guaranteed". It requires that each agent weakly prefers her individual assignment to the equal division assignment where each agent receives $1/N$ of each object. Formally, P satisfies **equal division lower bound (EDLB)** if $P \geq ED$, where ED denotes the equal division random assignment.

[8] For example, if $N = 3$ upper shuffle-proofness requires that no agent can benefit—in terms of the sum of assignment probabilities for the top two houses—by swapping these two houses. The agents could still possibly benefit: either in some other respect (not in terms of the surplus of the second best object), or from using other strategies (that involve other swaps).

Next, we introduce two auxiliary notions of fairness. A random assignment P is **upper envy-free (UEF)** if any two agents with identical upper contour sets of some house h receive equal assignment probabilities of h: for each $a, a' \in A, h \in H$ such that $U(\succ_a, h) = U(\succ_{a'}, h)$ it follows that $P_{ah} = P_{a'h}$. The other fairness notion is a generalization of equal treatment of equals. A random assignment P satisfies the **strong equal treatment of equals (SETE**) if any two agents with identical preferences from the top house down to some particular house receive identical assignments from the top down to that house.

Finally, we introduce two fairness notions and one efficiency notion for the cardinal framework.[9] Assume that each agent $a \in A$ reports her utility $u_a = \{u_{ah}\}_{h \in H} : u_{ah} \in \mathbb{R}$ for each object $h \in H$. A random assignment P is **symmetric** if every two agents a and a' with the same reported utilities $u_a = u_{a'}$ receive equal expected utility: $\sum u_{ah} P_{ah} = \sum u_{a'h} P_{a'h}$. A random assignment P is **anonymous** if the same two agents in addition receive identical random assignments: $P_a = P_{a'}$. A random assignment P is **ex-ante efficient at utility** $U = \{u_a\}_{a \in A}$ if there does not exist any other random assignment P' such that for each agent a assignment P provides at least as high (expected) utility as assignment P': $\sum P_{ah} u_{ah} \geq \sum P'_{ah} u_{ah}$, and at least for one of the agents the inequality is strict.

The fairness notions presented above can be logically ordered as follows.
(i) Envy-freeness $\Longrightarrow$ upper envy-freeness $\Longrightarrow$ strong equal treatment of equals $\Longrightarrow$ equal treatment of equals $\Longrightarrow$ anonymity $\Longrightarrow$ symmetry;
(ii) envy-freeness $\Longrightarrow$ weak envy-freeness;
(iii) envy-freeness $\Longrightarrow$ equal division lower bound. Weak envy-freeness, equal division lower bound and upper envy-freeness (as well as strong equal treatment of equals and equal treatment of equals) are logically independent.

We have now prepared all necessary definitions and their logical relations to study the first impossibility result presented in the next section.

3 First Impossibility Result

We begin by studying the tradeoff between the properties of a mechanism when fairness is of a higher concern than efficiency. The following theorem considers the set of strategy-proof mechanisms that are moderately efficient (at least ex-post efficient) and very fair (envy-free, which implies all other fairness criteria). The set of such mechanisms turns out to be empty:

Theorem 1. *For $N \geq 3$ there does not exist a mechanism that is ex-post efficient, strategy-proof, and envy-free.*

The result above is a direct corollary to a stronger result of Lemma 1:

[9] We need the cardinal framework only for the second characterization, apart from that we use the ordinal framework.

Lemma 1. *There does not exist a mechanism that is ex-post efficient, upper-shuffle-proof, and upper-envy-free. We first prove the claim for $N = 3$ and we do it by contradiction. Suppose there exists a mechanism φ satisfying ex-post efficiency, upper shuffle-proofness an upper envy-freeness.*

All three assumptions in the lemma are necessary. Should we drop the ex-post efficiency requirement, a uniform lottery mechanism satisfies strategy-proofness and envy-freeness (and, therefore, upper shuffle-proofness and upper envy-freeness). If we drop the strategy-proofness requirement, then the probabilistic serial mechanism satisfies ex-post efficiency and envy-freeness (and upper envy-freeness). Finally, RSD is a natural benchmark to discuss the fairness requirement. It is easy to show that RSD is always SETE because of the underlying dictatorship procedure: the assignment probabilities for every house depend only on the preferences for the corresponding upper contour set.[10] In the same time RSD is not upper envy-freeness, which is true, for instance, for the preference profile $\succ$ in the proof above. The lemma shows that this gap between strong equal treatment of equals and upper envy-freeness is so big, that even the certain compromise on strategy-proofness (requiring upper shuffle-proofness instead of strategy-proofness) is not enough to close it.

Lemma 1 can be seen as a generalization of the statement in BM (p. 310) about the incompatibility of ex-post efficiency, strategy-proofness, and no envy for the case of three agents. Here we show the incompatibility of ex-post efficiency and two weaker properties: upper strategy-proofness and upper envy-freeness for any number of agents.[11]

In the following section we interchange the fairness and efficiency requirements: we relax the fairness criterion and strengthen the efficiency criterion in order to obtain a different but closely related impossibility result.

4 Two Characterizations

We begin by characterizing the RSD mechanism as a unique strategy-proof, ex-post efficient, and weak envy-free mechanism for a problem with three agents.

Proposition 1. *(First characterization of RSD) For $N = 3$ a mechanism is strategy-proof, ex-post efficient, and weak envy-free for equals if and only if it is RSD.*

We get two immediate corollaries from the proposition by relaxing the weak envy-freeness among equals requirement.

Corollary 1. *(BM) For $N = 3$, a mechanism is strategy-proof, ex-post efficient, and satisfies the equal treatment of equals if and only if it is RSD.*

The second corollary follows from the fact that RSD satisfies weak envy-freeness (shown in BM):

[10] This property is defined as a *weak invariance* in Hashimoto et. al (2014) and plays a central role in their characterization of the probabilistic serial mechanism.

[11] Perhaps BM did not show this impossibility result for the general case since they had a different focus: "For problems involving four agents and more, the impossibility result is more severe" (p.310). However, the result they show (the incompatibility of strategy-proofness, ordinal efficiency and equal treatment of equals) is logically independent from Theorem 1 and especially from Lemma 1 since ordinal efficiency is stricter than ex-post efficiency.

Corollary 2. *For $N = 3$, a mechanism is strategy-proof, ex-post efficient and weak envy-free if and only if it is RSD.*

We now complement this result by another characterization result in which we use a slightly different fairness criterion: symmetry. Symmetry is defined using cardinal terms. [12] For a moment, assume that agents report their utilities and not just their ordinal preferences. Then a random assignment is symmetric if any two agents with identical utilities receive the assignments of the same expected utility.

Although symmetry is related to weak envy-freeness (equal agents receive individual assignments such that they do not dominate one another), these two properties are logically independent: symmetry applies to a smaller subset of utility domain, but for this set of utilities it also has stricter implications.

Proposition 2. *(Second characterization of RSD) For $N = 3$, a mechanism is strategy-proof, ex-post efficient and symmetric if and only if it is RSD.*

One of the consequences of the proposition is that for the case of three agents RSD can also be characterized using anonymity, a property used by Gale (1987) in his conjecture, and also using a stronger equal treatment of equals (Corollary 1). Since RSD has this property and since anonymity implies symmetry, any mechanism that is strategy-proof, ex-post efficient and anonymous is equivalent to RSD.

Another immediate consequence of the characterization is the impossibility (for the case of three agents) to find a mechanism that would cardinally dominate RSD and in the same time be symmetric and strategy-proof.

Corollary 3. *For $N = 3$, if a mechanism is strategy-proof and symmetric, it cannot dominate RSD.*

Given this result we can show the famous impossibility result from Zhou (1990): ex-ante efficiency, strategy-proofness and symmetry are mutually incompatible.

Corollary 4. *(Zhou 1990) For $N \geq 3$, there does not exist a mechanism that is strategy-proof, ex-ante efficient and symmetric.*

For the case $N = 3$ the proof is just the combination of the Corollary 3 and the fact that RSD is not ex-ante efficient but only ex-post efficient. For the general case $N \geq 3$, as it is done in the second part of the proof of the Theorem 1, we construct a preference profile such that any ex-post efficient mechanism cannot be cardinally dominated for any of the agents except the first three agents. For this preference profile the problem is effectively reduced to the size of three.

In the next section we use Corollary 2 for the second impossibility result.

[12] Since the focus of the paper is on ordinal properties, we describe the cardinal definitions and results rather informally.

5 Second Impossibility Result

In the previous two sections we mostly discussed the problems with only three agents. For these cases ex-post efficiency mechanism cannot be ordinally dominated, hence, ex-post efficiency coincides with ordinal efficiency. This changes when the number of agents is four or higher: an ex-post efficient mechanism such as RSD can be first-order stochastically dominated for some preference profiles. In the following two sections we further study the trade-off between fairness and efficiency, where we put a higher weight on the latter and require ordinal efficiency and not just ex-post efficiency.

The next result shows the loss in fairness required to satisfy ordinal efficiency: any ordinally efficient mechanism must be either non-strategy-proof or cannot eliminate sd-envy.

Theorem 2. *For $N \geq 4$ there does not exist a mechanism that is ordinally-efficient, strategy-proof, and weak envy-free.*

It is easy to see the independence of axioms in Theorem 2. First, let us weaken the ordinal efficiency requirement and demand ex-post efficiency. Then there exist at least one ex-post efficient, strategy-proof, weak envy-free mechanism: random serial dictatorship mechanism. Next, let us drop the weak-envy-freeness requirement. Then there exists at least one strategy-proof, ordinally efficient mechanism: serial dictatorship mechanism. Finally, the probabilistic serial mechanism is an example of an ordinally efficient, (weak) envy-free mechanism.

Overall, when the fairness of a random assignment is judged by comparing the individual assignments between each other, weak envy-freeness is arguably a reasonable minimum fairness requirement. In the following section, we discuss a different approach to fairness, where the individual assignments are compared to some alternative "fair" assignment such as equal division.

6 Third Impossibility Result

The last impossibility result also uses a strong notion of efficiency and a weak notion of fairness, but this time fairness is defined by the equal division lower bound.

Theorem 3. *For $N \geq 4$ there does not exist a mechanism that is ordinally-efficient, strategy-proof, and satisfies the equal division lower bound.*

We can easily check the independence of the axioms in this result. First, a pure lottery mechanism is strategy-proof and satisfies equal division lower bound, but is not ordinally efficient. Second, a serial dictatorship mechanism is strategy-proof and ex-ante efficient (and therefore ordinally efficient), but does not satisfy equal division lower bound. Finally, the probabilistic serial mechanism is ordinally efficient and envy-free (and therefore satisfies equal division lower bound), but is not strategy-proof.

From a theoretical point of view, equal division lower bound is related more to how efficient rather that how equitable the assignment is, as compared to weak envy-freeness and equal treatment of equals. Unlike the other two notions, EDLB does not compare the individual assignments to each other but to the (usually inefficient) equal division benchmark. Therefore, EDLB does not require the assignment to be fair in the egalitarian sense, but only that this assignment *dominates* the most egalitarian assignment − equal division.

Another essential feature of the equal division lower bound is that several popular mechanisms satisfy this property. One of these mechanisms is RSD. Indeed, in the RSD procedure each agent has an equal chance to be the first in the ordering (and thus receive her first best house), the second (and thus receive at least her second best) and so on. Therefore, under the RSD assignments all agents are weakly better off than under the uniform lottery. Hence, an important implication of Theorem 3 is the restriction that it puts on the feasibility set of mechanisms that dominate RSD.

Corollary 5. *For $N > 3$ any ordinally efficient mechanism dominating RSD is not strategy-proof.*

The corollary, however, does not restrict the set of mechanisms that dominate RSD *without* being ordinally efficient. Thus, in the set of strategy-proof mechanisms there might still be room for improvement upon RSD.

7 Conclusions

This paper considers the standard random assignment problem of assigning N indivisible objects to N agents and shows the impossibility for a strategy-proof mechanism to be simultaneously fair and efficient (in three specific ways). Theorem 1 shows the impossibility to combine a weak notion of efficiency − ex-post efficiency, with a strong notion of fairness − envy-freeness; it is the first known impossibility result in the related literature that involves ex-post efficiency. Theorem 2 shows the impossibility for the opposite set of properties: a weak notion of fairness − weak envy-freeness and a strong notion of efficiency − ordinal efficiency. Finally, Theorem 3 shows a similar impossibility result with a different weak fairness notion: equal division lower bound.

The paper also shows that for the case of three agents the trinity of strategy-proofness, ex-post efficiency, and weak envy-freeness for agents with identical preferences uniquely defines the random serial dictatorship mechanism. Alternatively, if we use symmetry—a cardinal fairness notion—instead of the ordinal weak envy-freeness among equals, we get the same characterization of the random serial dictatorship mechanism.

It, however, remains unclear, what combination of properties characterizes RSD for the general case. The characterization result in this paper cannot be directly generalized even for the case of four agents (however, there are also no counter examples found). The reason for this complication is that weak envy-freeness (and especially weak envy-freeness among equals) is not handy enough as compared to the equal treatment of equals. For instance, for two agents with identical preferences weak envy-freeness gives precise implications only in case these agents receive identical probabilities for all but two objects. Then the two agents have to have the same random assignment for the remaining objects as well. Equal treatment of equals, on the contrary, has implications for the assignment probabilities of all objects. Therefore, I believe, generalizing this characterization result would be more difficult than the result that uses equal treatment of equals.

Another open question is to what extent one of the three properties can be satisfied should the other two be taken at their extreme. For instance, if ordinal efficiency and envy-freeness are satisfied, then the probabilistic serial mechanism appears to be the "most" strategy-proof mechanism since it is weakly invariant (limits the set of profitable deviations) and weak strategy-proof (which means that no agent can receive a stochastically dominant assignment by manipulating). Similarly, one could be interested in the "most fair" mechanism that satisfies strategy-proofness and ordinal efficiency (since the only known SD mechanism is very unfair), and in the "most" efficient mechanism that satisfies strategy-proofness and envy-freeness (again, the only known equal division or pure lottery mechanism disregards preferences and therefore is almost always inefficient).

References

1. Bogomolnaia, A., & Moulin, H. (2001). A new solution to the random assignment problem. Journal of Economic Theory, 100(2), 295-328.

2. Che, Y. K., & Kojima, F. (2010). Asymptotic equivalence of probabilistic serial and random priority mechanisms. Econometrica, 78(5), 1625-1672.

3. Erdil, A. (2014). Strategy-proof stochastic assignment. Journal of Economic Theory, 151, 146-162.

4. Gale, D. (1987). College course assignments and optimal lotteries. University of California at Berkeley.

5. Hashimoto, T., Hirata, D., Kesten, O., Kurino, M., & Ünver, M. U. (2014). Two axiomatic approaches to the probabilistic serial mechanism. Theoretical Economics, 9(1), 253-277.

6. Hylland, A., & Zeckhauser, R. (1979). The efficient allocation of individuals to positions. The Journal of Political Economy, 293-314. ISO 690

7. Kesten, O., & Yazici, A. (2012). The Pareto-dominant strategy-proof and fair rule for problems with indivisible goods. Economic Theory, 50(2), 463-488.

8. Kojima, F., & Manea, M. (2010). Incentives in the probabilistic serial mechanism. Journal of Economic Theory, 145(1), 106-123.

9. Moulin, H. (2014). Cooperative Microeconomics: A Game-Theoretic Introduction.

10. Nesterov, A. S. (2014). Fairness and efficiency in a random assignment: Three impossibility results (No. SP II 2014-211). WZB Discussion Paper.

11. Thomson, W. (2007). Fair allocation rules. Handbook of social choice and welfare, 2, 393-506.

12. Zhou, L. (1990). On a conjecture by Gale about one-sided matching problems. Journal of Economic Theory, 52(1), 123-135.

It's Not Easy Being Three: The Approximability of Three-Dimensional Stable Matching Problems

Rafail Ostrovsky * Will Rosenbaum †

February 15, 2015

Abstract

In 1976, Knuth [14] asked if the stable marriage problem (SMP) can be generalized to marriages consisting of 3 genders. In 1988, Alkan [2] showed that the natural generalization of SMP to 3 genders (3GSM) need not admit a stable marriage. Three years later, Ng and Hirschberg [16] proved that it is NP-complete to determine if given preferences admit a stable marriage. They further prove an analogous result for the 3 person stable assignment (3PSA) problem.

In light of Ng and Hirschberg's NP-hardness result for 3GSM and 3PSA, we initiate the study of approximate versions of these problems. In particular, we describe two optimization variants of 3GSM and 3PSA: *maximally stable marriage/matching* (*MSM*) and *maximum stable submarriage/submatching* (*MSS*). We show that both variants are NP-hard to approximate within some fixed constant factor. Conversely, we describe a simple polynomial time algorithm which computes constant factor approximations for the maximally stable marriage and matching problems. Thus both variants of MSM are APX-complete.

1 Introduction

1.1 Previous Work

Since Gale and Shapley first formalized and studied the stable marriage problem (SMP) in 1962 [6], many variants of the SMP have emerged (see, for example, [7, 14, 15, 20]). While many of these variants admit efficient algorithms, two notably do not[1]: (1) incomplete preferences with ties [10], and (2) 3 gender stable marriages (3GSM) [16].

In the case of incomplete preferences with ties, it is NP-hard to find a maximum cardinality stable marriage [10]. The intractability of exact computation for this problem led to the study of approximate versions of the problem. These investigations have resulted in hardness of approximation results [9, 21] as well as constant factor approximation algorithms [12, 13, 18, 21].

In 3GSM, players are one of three genders: women, men, and dogs (as suggested by Knuth). Each player holds preferences over the set of *pairs* of players of the other two genders. The goal is to partition the players into families, each consisting of one man, one woman, and one dog, such that no triple mutually prefer one another to their assigned families. In 1988, Alkan showed that for this natural generalization of SMP to three genders, there exist preferences

*University of California, Los Angeles (Departments of Computer Science and Mathematics). Work supported in part by NSF grants 09165174, 1065276, 1118126 and 1136174, US-Israel BSF grant 2008411, OKAWA Foundation Research Award, IBM Faculty Research Award, Xerox Faculty Research Award, B. John Garrick Foundation Award, Teradata Research Award, and Lockheed-Martin Corporation Research Award. This material is based upon work supported by the Defense Advanced Research Projects Agency through the U.S. Office of Naval Research under Contract N00014-11-1-0392. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

†University of California, Los Angeles (Department of Mathematics).

[1] Assuming, of course, P$\neq$NP!

which do not admit a stable marriage [2]. In 1991, Ng and Hirschberg showed that, in fact, it is NP-complete to determine if given preferences admit a stable marriage [16]. They further generalize this result to the three person stable assignment problem (3PSA). In 3PSA, each player ranks all pairs of other players without regard to gender. The goal is to partition players into disjoint triples where again, no three players mutually prefer each other to their assigned triples.

Despite the advances for stable marriages with incomplete preferences and ties (see [15] for an overview of relevant work), analogous approximability results have not been obtained for 3 gender variants of the stable marriage problem. In this paper, we achieve the first substantial progress towards understanding the approximability of 3GSM and 3PSA.

1.2 Overview of our results

1.2.1 3 gender stable marriages (3GSM)

We formalize two optimization variants of 3GSM: maximally stable marriage (3G-MSM) and maximum stable submarriage (3G-MSS). For 3G-MSM, we seek a perfect (3 dimensional) marriage which minimizes the number of unstable triples—triples of players who mutually prefer each other to their assigned families. For 3G-MSS, we seek a largest cardinality *sub*marriage which contains no unstable triples among the married players. Exact computation of both of these problems is NP-hard by Ng and Hirschberg's result [16]. Indeed, exact computation of either allows one to detect the existence of a stable marriage.

We obtain the following inapproximability result for 3G-MSM and 3G-MSS.

Theorem 1.1 (Special case of Theorem 3.1)**.** There exists an absolute constant $c < 1$ such that it is NP-hard to approximate 3G-MSM and 3G-MSS to within a factor c.

In fact, we prove a slightly stronger result for 3G-MSM and 3G-MSS. We show that the problem of determining if given preferences admit a stable marriage or if all marriages are "far from stable" is NP-hard. See Section 2.1 and Theorem 3.1 for the precise statements. In the other direction, we describe a polynomial time constant factor approximation algorithm for 3G-MSM.

Theorem 1.2. There exists a polynomial time algorithm, **AMSM**, which computes a $\frac{4}{9}$-factor approximation to 3G-MSM.

Corollary 1.3. 3G-MSM is APX-complete.

1.2.2 Three person stable assignment (3PSA)

We also consider the three person stable assignment problem (3PSA). In this problem, players rank all pairs of other players and seek a (3 dimensional) matching—a partition of players into disjoint triples. Notions of stability, maximally stable matching, and maximum stable submatching are defined exactly as the analogous notions for 3GSM. We show that Theorems 1.1 and 1.2 have analogues with 3PSA:

Theorem 1.4. There exists a constant $c < 1$ such that it is NP-hard to approximate 3PSA-MSM and 3PSA-MSS to within a factor c.

Theorem 1.5. There exists a polynomial time algorithm, **ASA**, which computes a $\frac{4}{9}$-factor approximation to 3PSA-MSM.

Remark 1.6. We remark that the hardness of approximation of 3PSA-MSM bears a strong resemblance to the work of Abraham, Biró, and Manlove [1], who prove a similar hardness of approximation result for the two person stable assignment problem. The authors prove that finding a matching which minimizes the number of "blocking pairs" is NP-hard, as is approximating the minimum number of blocking pairs to within an additive error of $n^{1/2-\varepsilon}$.

Our proofs of the lower bounds in Theorems 1.1 and 1.4 use a reduction from the 3 dimensional matching problem (3DM) to 3G-MSM. Kann [11] showed that Max-3DM is Max-SNP complete. Thus, by the PCP theorem [3, 4] and [5], it is NP-complete to approximate Max-3DM to within some fixed constant factor. Our hardness of approximation results then follow from a reduction from 3DM to 3G-MSM.

Theorems 1.2 and 1.5 follow from a simple greedy algorithm. Our algorithm constructs marriages (or matchings) by greedily finding triples whose members are guaranteed to participate in relatively few unstable triples. Thus, we are able to efficiently construct marriages (or matchings) with a relatively small fraction of blocking triples.

2 Background and Definitions

2.1 3 Gender Stable Marriage (3GSM)

In the 3 gender stable marriage problem, there are disjoint sets of ***women***, ***men***, and ***dogs*** denoted by A (for Alice), B (for Bob), and D (for Dog), respectively. We assume $|A| = |B| = |D| = n$, and we denote the collection of ***players*** by $V = A \cup B \cup D$. A ***family*** is a triple abd consisting of one woman $a \in A$, one man $b \in B$, and one dog $d \in D$. A ***submarriage*** S is a set of pairwise disjoint families. A ***marriage*** M is a maximal submarriage—that is, one in which every player $v \in V$ is contained in some (unique) family so that $|M| = n$. Given a submarriage S, we denote the function $p_S : V \to V^2 \cup \{\varnothing\}$ which assigns each player $v \in V$ to their partners in S, with $p_S(v) = \varnothing$ if v is not contained in any family in S.

Each player $v \in V$ has a ***preference***, denoted $\succ_v$ over pairs of members of the other two genders. That is, each woman $a \in A$ holds a total order $\succ_a$ over $B \times D \cup \{\varnothing\}$, and similarly for men and dogs. We assume that each player prefers being in some family to having no family. For example, $bd \succ_a \varnothing$ for all $a \in A$, $b \in B$ and $d \in D$. An instance of the ***three gender stable marriage problem*** (3***GSM***) consists of A, B, and D together with preferences $P = \{\succ_v | v \in V\}$ for all of the players $v \in V$.

Given a submarriage S, a triple abd is an ***unstable triple*** if a, b and d each prefer the triple abd to their assigned families in S. That is, abd is unstable if and only if $bd \succ_a p_S(a)$, $ad \succ_b p_S(b)$, and $ab \succ_d p_S(d)$. A triple abd which is not unstable is ***stable***. In particular, abd is stable if at least one of a, b and d prefers their family in S to abd. Let A_S, B_S and D_S be the sets of women, men and dogs (respectively) which have families in S. A submarriage S is ***stable*** if there are no unstable triples in $A_S \times B_S \times D_S$.

Unlike the two gender stable marriage problem, this three gender variant arbitrary preferences need not admit a stable marriage. In fact, for some preferences, every marriage has many unstable triples (see Section 3.1). Thus we consider two optimization variants of the three gender stable marriage problem.

2.1.1 Maximally Stable Marriage (3G-MSM)

The ***maximally stable marriage problem*** (3G-MSM) is to find a marriage M with the maximum number of stable triples with respect to given preferences P. For fixed preferences P and marriage M, the ***stability*** of M with respect to P is the number of stable triples in $A \times B \times D$:

$$\mathrm{stab}(M) = |\{abd \,|\, abd \text{ is stable}\}| \,.$$

Thus, M is stable if and only if $\mathrm{stab}(M) = n^3$. Dually, we define the ***instability*** of M by $\mathrm{ins}(M) = n^3 - \mathrm{stab}(M)$. For fixed preferences P, we define

$$\mathrm{MSM}(P) = \max \{\mathrm{stab}(M) \,|\, M \text{ is a marriage}\} \,.$$

For preferences P and fixed $c < 1$, we define ***Gap***$_c$-3***G-MSM*** to be the problem of determining if $\mathrm{MSM}(P) = n^3$ or $\mathrm{MSM}(P) \leq cn^3$.

2.1.2 Maximum Stable Submarriage

The ***maximum stable submarriage problem*** (3G-MSS) is to find a maximum cardinality stable submarriage S. We denote

$$\mathrm{MSS}(P) = \max\{|S| \mid S \text{ is a stable submarriage}\}$$

Note that P admits a stable marriage if and only if $\mathrm{MSS}(P) = n$. For fixed $c < 1$, we define $\boldsymbol{Gap_c}$**-3*G-MSS*** to be the problem of determining if $\mathrm{MSS}(P) = n$ or if $\mathrm{MSS}(P) \le cn$.

2.2 Three person stable assignment (3PSA)

In the ***three person stable assignment problem*** (3***PSA***), there is a set U of $|U| = 3n$ players who wish to be partitioned into n disjoint triples. For a set $C \subseteq U$, we denote the set of k-subsets of C by $\binom{C}{k}$. A ***submatching*** is a set $S \subseteq \binom{U}{3}$ of disjoint triples in U. A ***matching*** M is a maximal submatching—a submatching with $|M| = n$. Given a submatching S, U_S is the set of players contained in some triple in S:

$$U_S = \{u \in U \mid u \in t \text{ for some } t \in S\}.$$

Each player $u \in U$ holds preferences among all pairs of potential partners. That is, each $u \in U$ holds a linear order $\succ_u$ on $\binom{U \setminus \{u\}}{2} \cup \{\varnothing\}$. We assume that each player prefers every pair to an empty assignment. Given a set P of preferences for all the players and a submatching S, we call a triple $uvw \in \binom{U_S}{3}$ ***unstable*** if each of u, v and w prefer the triple uvw to their assigned triples in S. Otherwise, we call uvw ***stable***. A submatching S is ***stable*** if it contains no unstable triples in $\binom{U_S}{3}$. We define the ***stability*** of S by

$$\mathrm{stab}(S) = \left| \left\{ uvw \in \binom{U_S}{3} \,\middle|\, uvw \text{ is stable} \right\} \right|.$$

Dually, the ***instability*** of S is $\mathrm{ins}(S) = \binom{|S|}{3} - \mathrm{stab}(S)$.

The ***maximally stable matching problem*** (3***PSA-MSM***) is to find a matching M which maximizes $\mathrm{stab}(M)$. The ***maximum stable submatching problem*** (3***PSA-MSS***) is to find a stable submatching S of maximum cardinality.

Remark 2.1. We may consider a variant of 3PSA with ***unacceptable partners***. In this variant, each player $u \in U$ ranks only a subset of $\binom{U \setminus \{u\}}{2}$, and prefers being unmatched to unranked pairs. 3GSM is a special case of this variant where $U = A \cup B \cup D$ and each player ranks precisely those pairs consisting of one player of each other gender. This observation will make our hardness results for 3GSM easily generalize to 3PSA.

2.3 Hardness of Gap_c-3DM-3

Our proofs of Theorems 1.1 and 1.4 use a reduction from the three dimensional matching problem (3DM). In this section, we briefly review 3DM, and state the approximability result we require for our lower bound results.

Let W, X and Y be finite disjoint sets with $|W| = |X| = |Y| = m$. Let $E \subseteq W \times X \times Y$ be a set of edges. A ***matching*** $M \subseteq E$ is a set of disjoint edges. The ***maximum*** 3 ***dimensional matching problem*** (***Max-***3***DM***) is to find (the size of) a matching M of largest cardinality in E. ***Max-***3***DM-***3 is the restriction of Max-3DM to instances where each element in $W \cup X \cup Y$ is contained in at most 3 edges. For a fixed constant $c < 1$, we define $\boldsymbol{Gap_c}$**-3*DM*-3** to be the problem of determining if an instance I of Max-3DM-3 has a perfect matching (a matching M of size m) or if every matching has size at most cm.

Theorem 2.2. There exists an absolute constant $c < 1$ such that Gap_c-3DM-3 is NP-hard.

Kann showed that Max-3DM-3 is Max-SNP complete[2] by giving an L-reduction from Max-3SAT-B to Max-3DM-3 [11]. By the celebrated PCP theorem [3, 4] and [5], Kann's result immediately implies that Max-3DM-3 is NP-hard to approximate to within some fixed constant factor. However, Kann's reduction gives a slightly weaker result than Theorem 2.2. In Kann's reduction, satisfiable instances of 3SAT-B do not necessarily reduce to instances of 3DM-3 which admit perfect matchings. In the extended version of this paper [17], we describe a straightforward alteration of Kann's reduction such that satisfiable instances of 3SAT-B admit perfect matchings, while far-from-satisfiable instances are still far from admitting perfect matchings.

3 Hardness of Approximation

In this section, we prove the main hardness of approximation results. Specifically, we will prove the following theorems.

Theorem 3.1. There exists and absolute constant $c < 1$ such that Gap$_c$-3G-MSM and Gap$_c$-3G-MSS are NP-hard.

Theorem 3.2. There exists an absolute constant $c < 1$ such that Gap$_c$-3PSA-MSM and Gap$_c$-3PSA-MSS are NP-hard.

3.1 Preferences for 3GSM with Many Unstable Triples

In this section, we construct preferences P for 3GSM such that any marriage M induce many unstable triples with respect to P.

Theorem 3.3. There exist preferences P for 3GSM and a constant $c < 1$ for which MSM$(P) \leq cn^3$.

We first consider the case where $n = 2$. We denote $A = \{a_1, a_2\}$, $B = \{b_1, b_2\}$, and $D = \{d_1, d_2\}$. Consider preference lists P as described in the following table, where most preferred partners are listed first.

player	preferences		
a_1	b_1d_1	b_2d_2	$\cdots$
a_2	b_2d_1	$\cdots$	
b_1	a_1d_1	$\cdots$	
b_2	a_1d_2	a_2d_1	$\cdots$
d_1	a_2b_2	a_1b_1	$\cdots$
d_2	a_1b_2	$\cdots$	

The ellipses indicate that the remaining preferences are otherwise arbitrary. Suppose M is a stable marriage for P. We must have either $a_1b_1d_1 \in M$ or $a_1b_2d_2 \in M$, for otherwise the triple $a_1b_2d_2$ is unstable. However, if $a_1b_1d_1 \in M$, then $a_2b_2d_1$ is unstable. On the other hand, if $a_1b_2d_2 \in M$ then $a_1b_1d_1$ is unstable. Therefore, no such stable M exists. In particular, every marriage M contains at least one unstable triple.

The idea of the proof of Theorem 3.3 is to choose preferences P such that when restricted to many sets of two women, two men and two dogs, the relative preferences are as above. Thus any marriage containing families consisting of these players must induce unstable triples.

Proof of Theorem 3.3. We partition the sets A, B and D each into two sets of equal size: $A = A_1 \cup A_2$, $B = B_1 \cup B_2$, $D = D_1 \cup D_2$. Consider the preferences P described in Figure 1. We will prove that for P, every matching M contains at least $n^3/128$ unstable triples. Let M be an arbitrary marriage, and suppose ins$(M) < n^3/128$. We consider two cases separately.

[2]The complexity class Max-SNP was introduced by Papadimitriou and Yannakakis in [19], where the authors also show that Max-3SAT-B is Max-SNP complete.

player	preferences		
$a_1 \in A_1$	$B_1 D_1$	$B_2 D_2$	$\cdots$
$a_2 \in A_2$	$B_2 D_1$	$\cdots$	
$b_1 \in B_1$	$A_1 D_1$	$\cdots$	
$b_2 \in B_2$	$A_1 D_2$	$A_2 D_1$	$\cdots$
$d_1 \in D_1$	$A_2 B_2$	$A_1 D_1$	$\cdots$
$d_2 \in D_2$	$A_1 B_2$	$\cdots$	

Figure 1: Preferences P inducing many blocking triples. Assuming n is even, we partition each gender into two equal sized sets $A = A_1 \cup A_2$, $B = B_1 \cup B_2$, and $D = D_1 \cup D_2$. The sets appearing in the preferences indicate that the player prefers all pairs in that set (in any order) followed by the remaining preferences. For example, all $a_1 \in A_1$ prefer all partners $bd \in B_1 \times D_1$, followed by all partners in $B_2 \times D_2$, followed by all other pairs in arbitrary order. Within $B_1 \times D_1$ and $B_2 \times D_2$, a_1's preferences are arbitrary.

Case 1: $|M \cap (A_1 \times B_1 \times D_1)| \leq n/4$. Let A_1', B_1' and D_1' be the subsets of A_1, B_1 and D_1 respectively of players not in triples contained in $A_1 \times B_1 \times D_1$. By the hypothesis, $|A_1'|, |B_1'|, |D_1'| \geq n/4$. Let $d_1 \in D_1'$. Notice that if $p_M(d_1) \notin A_2 \times B_2$, then $a_1 b_1 d_1$ is unstable for all $a_1 \in A_1'$, $b_2 \in B_1'$. Since fewer than $n^3/128$ triples in $A_1' \times B_1' \times D_1'$ are unstable, at least $3n/8$ dogs $d_1 \in D_1'$ must have families $a_2 b_2 d_1 \in A_2 \times B_2 \times D_1'$.

Since $|M \cap (A_2 \times B_2 \times D_1)| \geq 3n/8$, we must have $|M \cap (A_1 \times B_2 \times D_2)| \leq n/8$. Thus, there must be at least $n/8$ women $a_1 \in A_1$ with partners not in $(B_1 \times D_1) \cup (B_2 \times D_2)$. However, every such a_1 forms an unstable triple with every $b_2 \in B_2$ and $d_2 \in D_2$ which are not in families in $A_1 \times B_2 \times D_2$. Since there at least $3n/8$ such b_2 and d_2, there are at least

$$\left(\frac{n}{8}\right)\left(\frac{3n}{8}\right)\left(\frac{3n}{8}\right) > \frac{n}{128}$$

blocking triples, a contradiction.

Case 2: $|M \cap (A_1 \times B_1 \times D_1)| > n/4$. In this case, we must have $|M \cap (A_2 \times B_2 \times D_1)| < n/4$. This implies that

$$|M \cap (A_1 \times B_2 \times D_2)| > 3n/8 \tag{1}$$

for otherwise triples $a_2 b_2 d_1 \in (A_2 \times B_2 \times D_1)$ with $p_M(b_2) \notin A_1 \times D_1$ form more than $n^3/128$ unstable triples. But (1) contradicts the Case 2 hypothesis, as $|A_1| = n/2$.

Since both cases lead to a contradiction, we may conclude that any M contains at least $n^3/128$ unstable triples, as desired. □

3.2 The Embedding

We now describe an embedding of 3DM-3 into 3G-MSM. Our embedding is a modification of the embedding described by Ng and Hirschberg [16]. Let I be an instance of 3DM-3 with ground sets W, X, Y and edge set E. We assume $|W| = |X| = |Y| = m$. We will construct an instance $f(I)$ of 3G-MSM with sets A, B and D of women, men and dogs of size $n = 6m$ and suitable preferences P. We divide each gender into two sets $A = A^1 \cup A^2$, $B = B^1 \cup B^2$ and $D = D^1 \cup D^2$ where $|A^j| = |B^j| = |D^j| = 3m$ for $j = 1, 2$. Let $W = \{a_1, a_2, \ldots, a_m\}$, $X = \{b_1, b_2, \ldots, b_m\}$ and $Y = \{d_1, d_2, \ldots, d_m\}$, and denote

$$E = \bigcup_{i=1}^{n} \{a_i b_{k_1} d_{\ell_1},\ a_i b_{k_2} d_{\ell_2},\ a_i b_{k_3} d_{\ell_3}\}.$$

Without loss of generality, we assume each a_i is contained in exactly 3 edges by possibly increasing the multiplicity of edges containing a_i. The idea of the embedding $f(I)$ is that each $a_i \in W$ is mapped to 6 players $a_i^1[1], a_i^1[2], a_i^1[3] \in A^1$ and $a_i^2[1], a_i^2[2], a_i^2[3] \in A^2$. These three players in A^1 and A^2 correspond to the three edges in E which contain a_i. We choose preferences in such a way that at most one family from $a_i^1[1]b_k^1 d_\ell^1$, $a_i^1[2]b_k^1 d_\ell^1$, and $a_i^1[3]b_k^1 d_\ell^1$ can be in any (sub)marriage, where $a_i b_k d_\ell \in E$. Such a family corresponds to a choice of an edge containing a_i to include in a maximal matching. We then show that if I admits a perfect matching, then $f(I)$ admits a stable marriage. On the other hand, if I is far from admitting a perfect matching, then our choice of preferences ensure that any marriage induces many unstable triples by appealing to Theorem 3.3.

For $j = 1, 2$, we form the sets

$$A^j = \left\{ a_i^j[k] \,\middle|\, i \in [n],\ k \in [3] \right\}, \quad B^j = \left\{ b_i^j,\ w_i^j,\ y_i^j \,\middle|\, i \in [n] \right\}, \quad D^j = \left\{ d_i^j,\ x_i^j,\ z_i^j \,\middle|\, i \in [n] \right\}.$$

We now define preferences for each set of players, beginning with those in A.

$$\begin{array}{r|llllll}
a_i^1[m] & w_i^1 x_i^1 & y_i^1 z_i^1 & b_{k_m}^1 d_{\ell_m}^1 & B^1 D^1 & B^2 D^2 & \cdots \\
a_i^2[m] & w_i^2 x_i^2 & y_i^2 z_i^2 & b_{k_m}^2 d_{\ell_m}^2 & B^2 D^1 & \cdots &
\end{array}$$

The players in B have preferences given by

$$\begin{array}{r|llllll}
w_i^1 & a_i^1[1] x_i^1 & a_i^1[2] x_i^1 & a_i^1[3] x_i^1 & A^1 D^1 & \cdots & \\
y_i^1 & a_i^1[1] z_i^1 & a_i^1[2] z_i^1 & a_i^1[3] z_i^1 & A^1 D^1 & \cdots & \\
b_i^1 & A^1 D^1 & \cdots & & & & \\
w_i^2 & a_i^2[1] x_i^2 & a_i^2[2] x_i^2 & a_i^2[3] x_i^2 & A^1 D^2 & A^2 D^1 & \cdots \\
y_i^2 & a_i^2[1] z_i^2 & a_i^2[2] z_i^2 & a_i^2[3] z_i^2 & A^1 D^2 & A^2 D^1 & \cdots \\
b_i^2 & A^1 D^2 & A^2 D^1 & \cdots & & &
\end{array}$$

The preferences for D are given by

$$\begin{array}{r|llllll}
x_i^1 & a_i^1[3] w_i^1 & a_i^1[2] w_i^1 & a_i^1[1] w_i^1 & A^2 B^2 & A^1 B^1 & \cdots \\
z_i^1 & a_i^1[3] y_i^1 & a_i^1[2] y_i^1 & a_i^1[1] y_i^1 & A^2 B^2 & A^1 B^1 & \cdots \\
d_i^1 & A^2 B^2 & A^1 B^1 & \cdots & & & \\
x_i^2 & a_i^2[3] w_i^2 & a_i^2[2] w_i^2 & a_i^2[1] w_i^2 & A^1 B^2 & \cdots & \\
z_i^2 & a_i^2[3] y_i^2 & a_i^2[2] y_i^2 & a_i^2[1] y_i^2 & A^1 B^2 & \cdots & \\
d_i^2 & A^1 B^2 & \cdots & & & &
\end{array}$$

The sets A^j, B^j and D^j in the preferences described above indicate that all players in these sets appear consecutively in some arbitrary order in the preferences. Ellipses indicate that all remaining preferences may be completed arbitrarily. For example, $a_1^1[1]$ most prefers $w_1^1 x_1^1$, followed by $y_1^1 z_1^1$ and $b_{k_m}^1 d_{\ell_m}^1$. She then prefers all remaining pairs in $B^1 D^1$ in any order, followed by all pairs in $B^2 D^2$, followed by the remaining pairs in any order.

Lemma 3.4. The embedding $f : \text{3DM-3} \longrightarrow \text{3GSM}$ described above satisfies

1. If $\mathrm{opt}(I) = m$—that is, I admits a perfect matching—then $f(I)$ admits a stable marriage (i.e. $\mathrm{MSM}(P) = n^3$).

2. If $\mathrm{opt}(I) \leq cm$ for some $c < 1$, then there exists a constant $c' < 1$ depending only on c such that $\mathrm{MSM}(P) \leq c' n^3$.

Proof. To prove the first claim assume, without loss of generality, that $M' = \{a_i b_{k_1} d_{\ell_1} \mid i \in [n]\}$ is a perfect matching in E. It is easy to verify the marriage

$$M = \left\{a_i^j[1] b_{k_1}^j d_{\ell_1}^j\right\} \cup \left\{a_i^j[2] w_i^j x_i^j\right\} \cup \left\{a_i^j[3] y_i^j z_i^j\right\}$$

contains no blocking triples, hence is a stable marriage.

For the second claim, let M be an arbitrary marriage in $A \times B \times D$. We observe that there are at least $(1-c)m$ players $a^1 \in A^1$ and $(1-c)m$ players $a^2 \in A^2$ that are not matched with pairs from their top three choices. Suppose to the contrary that $\alpha > (2+c)m$ players $a^1 \in A^1$ are matched with their top 3 choices. This implies that more than cm women $a^1 \in A^1$ are matched in triples of the form $a^1 b_k^1 d_\ell^1$ with $ab_k d_\ell \in E$, implying that E contains a matching of size $\alpha - 2m > cm$, a contradiction. Thus at least $2(1-c)m$ women in $A^1 \cup A^2$ are matched below their top three choices.

Let A' denote the set of women matched below their top three choices, and B' and D' the sets of partners of $a \in A'$ in M. By the previous paragraph, $|A'| \geq 2(1-c)m = (1-c)m/6$. Further, the relative preferences of players in A', B' and D' are precisely those described in Theorem 3.3. Thus, by Theorem 3.3, any marriage M among these players induces at least $|A'|^3/128$ blocking triples. Hence M must contain at least

$$\frac{|A'|}{128} \geq \frac{(1-c)^3}{3456} n^3$$

blocking triples. $\square$

Proof of Theorem 3.1. The reduction f : 3DM-3 $\longrightarrow$ 3GSM is easily seen to be polynomial time computable. Thus, by Lemma 3.4, f is a polynomial time reduction from Gap$_c$-3DM-3 to Gap$_{c'}$-3G-MSM where $c' = 1 - (1-c)^3/3456$. The NP hardness of Gap$_c$-3G-MSM is then an immediate consequence of Theorem 2.2.

The hardness of Gap$_c$-3G-MSS is a consequence of the hardness Gap$_c$-3G-MSM. Consider an instance of 3GSM with preferences P. We make the following observations.

1. MSM$(P) = n^3$ if and only if MSS$(P) = n$.

2. If MSM$(P) \leq (1-3\varepsilon)n^3$ for $\varepsilon > 0$, then MSS$(P) \leq (1-\varepsilon)n$.

The first observation is clear. To prove the second, suppose that MSS$(P) > (1-\varepsilon)n$, and let S be a maximum stable submarriage. We can form a marriage M by arbitrarily adding εn disjoint families to S. Since each new family can induce at most $3n^2$ blocking triples, M has at most $3\varepsilon n^3$ blocking triples, hence MSM$(P) > (1-3\varepsilon)n^3$. The two observations above imply that any decider for Gap$_{(1-\varepsilon)}$-3G-MSS is also a decider for Gap$_{(1-3\varepsilon)}$-3G-MSM. Thus, the NP-hardness of Gap$_c$-3G-MSM immediately implies the analogous result for Gap$_c$-3G-MSS. $\square$

Here we sketch a proof of the analogous lower bounds for 3PSA given in Theorem 3.2.

Proof sketch of Theorem 3.2. As noted in Remark 2.1, we may view 3GSM as a special case of 3PSA with incomplete preferences. The NP-hardness of approximation of 3PSA with incomplete preferences is analogous to the proof of Theorem 3.1. Given an instance I of 3GSM with sets A, B, and D and preferences P, take $U = A \cup B \cup D$ and form 3PSA preferences P' by appending the remaining pairs to P arbitrarily. Analogues of Theorem 3.3 and Lemma 3.4 hold for this instance of 3PSA, whence Theorem 3.2 follows. We leave details to the reader. $\square$

4 Approximation Algorithms

4.1 3GSM approximation

In this section, we describe a polynomial time approximation algorithm for MSM, thereby proving Theorem 1.2. Consider an instance of 3GSM with preferences P, and as before $A = \{a_1, a_2, \dots, a_n\}$, $B = \{b_1, b_2, \dots, b_n\}$, and $D = \{d_1, d_2, \dots, d_n\}$. Given a triple $a_i b_j d_k$, we define its ***stable set*** S_{ijk} to be the set of (indices of) triples which cannot form unstable triples with $a_i b_j d_k$. Specifically, we have

$$S_{ijk} = \left\{\alpha\beta\delta \in [n]^3 \,\middle|\, b_\beta d_\delta \preceq_{a_i} b_j d_k,\ \alpha = i\right\} \cup \left\{\alpha\beta\delta \in [n]^3 \,\middle|\, a_\alpha d_\delta \preceq_{b_j} a_i d_k,\ \beta = j\right\}$$
$$\cup \left\{\alpha\beta\delta \in [n]^3 \,\middle|\, a_\alpha b_\beta \preceq_{d_k} a_i b_j,\ \delta = k\right\}$$

The idea of our algorithm is to greedily form families that maximize $|S_{ijk}|$. Pseudocode is given in Algorithm 1.

Algorithm 1 AMSM(A, B, D, P)

---

find $ijk \in [n]^3$ which maximize $|S_{ijk}|$
$A' \leftarrow A \setminus \{a_i\}$, $B' \leftarrow B \setminus \{b_j\}$, $D' \leftarrow D \setminus \{d_k\}$
$P' \leftarrow P$ restricted to A', B', and D'
return $\{a_i b_j d_k\} \cup$**AMSM**(A', B', D', P')

It is easy to see that **AMSM** can be implemented in polynomial time. The naive algorithm for computing $|S_{ijk}|$ for fixed $ijk \in [n]^3$ by iterating through all triples $\alpha\beta\delta \in [n]^3$ and querying each player's preferences can be implemented in time $\tilde{O}(n^3)$. The maximal such $|S_{ijk}|$ can then be found by iterating through all $ijk \in [n]^3$. Thus the first step in **AMSM** can be accomplished in time $\tilde{O}(n^6)$. Finally, the recursive step of **AMSM** terminates after n iterations, as each iteration decreases the size of A, B, and D by one.

Lemma 4.1. For any preferences P, and sets A, B and D with $|A| = |B| = |D| = n$, there exists a triple $ijk \in [n]^3$ with

$$|S_{ijk}| \geq \frac{4n^2}{3} - n - 1. \tag{2}$$

Proof. We will show that there exists a triple $a_i b_j d_k$ such that at least two of a_i, b_j, and d_k respectively rank $b_j d_k$, $a_i d_k$, and $a_i b_j$ among their top $n^2/3 + 1$ choices. Note that this occurs precisely when at least two of the the following inequalities are satisfied

$$\left|\left\{\beta\delta \in [n]^2 \,\middle|\, b_\beta d_\delta \succeq_{a_i} b_j d_k\right\}\right| \leq \frac{n^2}{3} + 1, \quad \left|\left\{\alpha\delta \in [n]^2 \,\middle|\, a_\alpha d_\delta \succeq_{b_j} a_i d_k\right\}\right| \leq \frac{n^2}{3} + 1,$$
$$\text{and} \quad \left|\left\{\alpha\beta \in [n]^2 \,\middle|\, a_\alpha b_\beta \succeq_{d_k} a_i b_j\right\}\right| \leq \frac{n^2}{3} + 1$$

Mark each triple $a_i b_j d_k$ which satisfies one of the above inequalities. Each a_i induces $\frac{n^2}{3} + 1$ marks, so we get $\frac{n^3}{3} + n$ marks from all $a \in A$. Similarly, we get $\frac{n^3}{3} + n$ marks from B and D. Thus, marks are placed on at least $n^3 + 3n$ triples. By the pigeonhole principle, at least one triple is marked twice.

We claim that the triple $a_i b_j d_k$ satisfying two of the above inequalities satisfies equation (2). Without loss of generality, assume that $a_i b_j d_k$ satisfies the first two equations. Thus, a_i and b_j must each contribute at least $\frac{2n^3}{3} - 1$ stable triples with respect to $a_i b_j d_k$. Further, at most $n - 1$ such triples can be contributed by both a_i and b_j, as such triples must be of the form $a_i b_j d_\delta$ for $\delta \neq k$. Thus (2) is satisfied, as desired. $\square$

We are now ready to prove that **AMSM** gives a constant factor approximation for the maximally stable marriage problem.

Proof of Theorem 1.2. Let M be the marriage found by **AMSM**, and suppose

$$M = \{a_1b_1d_1,\ a_2b_2d_2,\ \ldots,\ a_nb_nd_n\}$$

where $a_1b_1d_1$ is the first triple found by **AMSM**, $a_2b_2d_2$ is the second, et cetera. By Lemma 4.1, $|S_{111}| \geq \frac{4}{3}n^2 - O(n)$. Therefore, the players a_1, b_1, and d_1 can be contained in at most $\frac{5}{3}n^2 + O(n)$ unstable triples in any marriage containing the family $a_1b_1d_1$. Similarly, for $1 \leq i \leq n$ the ith family $a_ib_id_i$ can contribute at most $\frac{5}{3}(n-i+1)^2 + O(n)$ new unstable triples (not containing any a_j, b_j, or d_j for $j < i$). Thus, the total number of unstable triples in M is at most

$$\sum_{i=1}^{n} \left(\frac{5}{3}(n-i+1)^2 + O(n) \right) = \frac{5}{9}n^3 + O(n^2).$$

Thus, we have $\mathrm{stab}(M) \geq 4n^3/9 - O(n^2)$ as desired. □

4.2 3PSA approximation

AMSM can easily be adapted for 3PSA. Let U be a set of players with $|U| = 3n$, and let P be a set of complete preferences for the players in U. Given a triple $abc \in \binom{U}{3}$, we form the ***stable set*** S_{abc} consisting of triples that at least one of a, b, c does not prefer to abc. The approximation algorithm **ASA** for 3PSA is analogous to **AMSM**: form a matching M by finding a triple abc that maximizes $|S_{abc}|$, then recursing. The following lemma and its proof are analogous to Lemma 4.1.

Lemma 4.2. For any set U of players with $|U| = 3n$ and complete preferences P, there exists a triples $abc \in \binom{U}{3}$ such that

$$|S_{abc}| \geq 6n^2 - O(n).$$

Using Lemma 4.2, we prove Theorem 1.5 analogously to Theorem 1.2.

Proof of Theorem 1.5. Each triple abc can intersect at most $3\binom{3n}{2} \leq \frac{27}{2}n^2$ blocking triples. Thus, by Lemma 4.2, the total number blocking triples in the matching M found by **ASA** is at most

$$\sum_{i=0}^{n-1} \left(\frac{15}{2}(n-i+1)^2 + O(n) \right) = \frac{5}{2}n^3 + O(n^2).$$

Therefore,

$$\mathrm{stab}(M) \geq 2n^3 - O(n^2),$$

as the total number of triples in $\binom{U}{3}$ is $\frac{9}{2}n^3 - O(n^2)$. Hence M is a $\frac{4}{9}$-approximation to a maximally stable matching, as desired. □

5 Concluding Remarks and Open Questions

While **AMSM** gives a simple approximation algorithm for 3G-MSM, we do not generalize this result to 3G-MSS. Indeed, even the first two families output by **AMSM** may include blocking triples. We leave the existence of an efficient approximation for 3G-MSS as a tantalizing open question.

Open Problem 5.1. Is it possible to efficiently compute a constant factor approximation to 3G-MSS?

Finding an approximation algorithm for maximally stable marriage was made easier by the fact that any preferences admit a marriage/matching with $\Omega(n^3)$ stable triples. However, for 3G-MSS, it is not clear whether every preference structure admits stable submarriages of size $\Omega(n)$. We feel that understanding the approximability of 3G-MSS is a very intriguing avenue of further exploration.

Open Problem 5.2. How small can a maximum stable submarriage/submatching be? What preferences achieve this bound?

In our hardness of approximation results (Theorems 3.1 and 3.2), we do not state explicit values of c for which Gap$_c$-3G-MSM and Gap$_c$-3G-MSS (and the corresponding problems for three person stable assignment) are NP-complete. The value implied by our embedding of 3SAT-B via 3DM-3 is quite close to 1. It would be interesting to find a better (explicit) factor for hardness of approximation. Conversely, is it possible to efficiently achieve a better than 4/9-factor approximation for maximally stable marriage/matching problems?

Open Problem 5.3. For the maximally stable marriage/matching problems, close the gap between the 4/9-factor approximation algorithm and the $(1-\varepsilon)$-factor hardness of approximation.

Notice that in the preference structure described in the proof of Theorem 3.3 (upon which our hardness of approximation results rely), for any $i \neq j$, w_i^2 prefers $a_i^2[1]x_i^2$ to $a_i^1[1]x_i^2$, but prefers $a_i^1[1]x_j^2$ to $a_i^2[1]x_j^2$. Thus, depending on the the second player (x_j^2 or x_i^2), w_i^2 does not consistently prefer pairs involving $a_i^1[1]$ to $a_i^2[1]$ or vice versa. Ng and Hirschberg call such preferences as ***inconsistent***, and asked whether consistent preferences always admit a (3 gender) stable marriage. Huang [8] showed that consistent preferences need not admit stable marriages, and indeed it is still NP-complete to determine whether or not given consistent preferences admit a stable marriage.

Open Problem 5.4. Are MSM and MSS still hard to approximate if preferences are restricted to be consistent?

References

[1] David J Abraham, Péter Biró, and David F Manlove. almost stable matchings in the roommates problem. In *Approximation and online algorithms*, pages 1–14. Springer, 2006.

[2] Ahmet Alkan. Nonexistence of stable threesome matchings. *Mathematical Social Sciences*, 16(2):207–209, 1988.

[3] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof Verification and the Hardness of Approximation Problems. *J. ACM*, 45(3):501–555, 1998.

[4] Sanjeev Arora and Shmuel Safra. Probabilistic Checking of Proofs: A New Characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[5] Irit Dinur. The PCP theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12, 2007.

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

[7] Dan Gusfield and Robert W Irving. *The stable marriage problem: structure and algorithms*, volume 54. MIT press Cambridge, 1989.

[8] Chien-Chung Huang. Two's Company, Three's a Crowd: Stable Family and Threesome Roommates Problems. In Lars Arge, Michael Hoffmann, and Emo Welzl, editors, *Algorithms – ESA 2007*, pages 558–569. Springer Berlin Heidelberg, 2007.

[9] Robert W Irving, David F Manlove, and Gregg O'Malley. Stable marriage with ties and bounded length preference lists. *Journal of Discrete Algorithms*, 7(2):213–219, 2009.

[10] Kazuo Iwama, Shuichi Miyazaki, Yasufumi Morita, and David Manlove. Stable marriage with incomplete lists and ties. In *Automata, Languages and Programming*, pages 443–452. Springer, 1999.

[11] Viggo Kann. Maximum bounded 3-dimensional matching is MAX SNP-complete. *Information Processing Letters*, 37(1):27–35, 1991.

[12] Zoltán Király. Better and simpler approximation algorithms for the stable marriage problem. *Algorithmica*, 60(1):3–20, 2011.

[13] Zoltán Király. Linear time local approximation algorithm for maximum stable marriage. *Algorithms*, 6(3):471–484, 2013.

[14] Donald Ervin Knuth. *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*, volume 10. American Mathematical Soc., 1997.

[15] David Manlove. *Algorithmics of matching under preferences.* World Scientific Publishing, 2013.

[16] Cheng Ng and Daniel S Hirschberg. Three-Dimensional Stabl Matching Problems. *SIAM Journal on Discrete Mathematics*, 4(2):245–252, 1991.

[17] Rafail Ostrovsky and Will Rosenbaum. It's not easy being three: The approximability of three-dimensional stable matching problems. *CoRR*, abs/1412.1130, 2014.

[18] Katarzyna Paluch. Faster and simpler approximation of stable matchings. *Algorithms*, 7(2):189–202, 2014.

[19] Christos H Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 425–440, New York, NY, USA, 1991. ACM.

[20] Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis.* Cambridge University Press, 1992.

[21] Hiroki Yanagisawa. *Approximation algorithms for stable marriage problems.* PhD thesis, Kyoto University, Graduate School of Informatics, 2007.

Format B Paper Abstracts

Sticky Matching in School Choice

Mustafa Oğuz Afacan* Zeynel Harun Alioğulları† Mehmet Barlo‡

November, 2014

Abstract

We analyze the school choice model and introduce costly appeals against violations of students' priorities. If these costs are sufficiently high, then some of such appeals may not provide benefits to the parents even when their priorities are violated. Instead of working with cardinal notions, our construction elicits the relevant ordinal implications of these costs, the information about the least rank decrease a student would be appealing against a priority violation (his/her *stickiness degree*), from the students before the assignment is determined. Then the notion of stability, the main desiderata in school choice known to be at odds with efficiency, is weakened by disregarding priority violations not worth the cost and the notion of *sticky stability* is obtained. The first mechanism we introduce is "*efficiency improving deferred acceptance mechanism*" ($EIDA$) and we show that it is sticky stable and superior to the Gale and Shapley (1962)'s deferred acceptance mechanism (DA) in terms of efficiency and involves truthful revelations of the stickiness degrees. The $EIDA$ not maximally improving efficiency in the class of sticky stable solutions, leads us to design "*efficiency corrected deferred acceptance mechanism*" ($ECDA$) which turns out to be both sticky stable and efficient within the class of sticky stable mechanisms. While both mechanisms lack full incentive properties in the complete information case, in certain incomplete information settings, the former becomes immune to manipulations, whereas, the latter is still manipulable but with a diminished scope.

*Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; Phone: +90 216 483 9326; Fax: +90 216 483 9250 (CC. M. O. Afacan); email: mafacan@sabanciuniv.edu.

†Central Bank of the Republic of Turkey, Istiklal Cad. 10, Ulus, 06100 Ankara, Turkey; and FASS, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; email: zeynelharun@sabanciuniv.edu.

‡**Corresponding Author**: Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; Phone: +90 216 483 9284; Fax: +90 216 483 9250 (CC. M. Barlo); email: barlo@sabanciuniv.edu.

Generalized Three-Sided Assignment Markets: Consistency and the Core*

Ata ATAY[†1], Francesc Llerena[‡2], and Marina NÚÑEZ[§1]

[1]Department of Mathematical Economics, Finance and Actuarial Sciences, University of Barcelona. Av. Diagonal, 690, 08034 Barcelona, Spain

[2]Departament de Gestió d'Empreses, Universitat Rovira i Virgili-CREIP, Spain.

Abstract

A class of three-sided markets (and games) is considered, where value is generated by pairs or triplets of agents belonging to different sectors, as well as by individuals. For these markets we analyze the situation that arises when some agents leave the market with some payoff. To this end, we introduce the derived market (and game) and relate it to the Davis and Maschler (1965) reduced game. Consistency with respect to the derived market, together with singleness best and individual anti-monotonicity axiomatically characterize the core for these generalized three-sided assignment markets. These markets may have an empty core, but we define a balanced subclass, where the worth of each triplet is defined as the addition of the worths of the pairs it contains.

*The authors acknowledge the support from research grant ECO2011-22765 (Ministerio de Ciencia e Innovación and FEDER), 2014SGR40 and 2014SGR631 (Generalitat de Catalunya).

†E-mail: aatay@ub.edu

‡E-mail: francesc.llerena@urv.cat

§E-mail: mnunez@ub.edu

College Admission with Multidimensional Privileges: The Brazilian Affirmative Action Case

Orhan Aygün* Inácio Bó†

November, 2014

Abstract

In August of 2012 the Brazilian federal government enacted a law mandating the implementation of affirmative action policies in public federal universities for candidates from racial minorities, low income families and those coming from public high-schools. We show that by using the method proposed by the government, students who strategize over the privileges that they claim may improve their placement. Moreover, the choices made by the colleges will not satisfy a general fairness condition, implying that high achieving students target of the affirmative action policies may be rejected while low achieving high income majorities are accepted. Data from university admissions in more than 3,000 programs in 2013 show that the conditions for those unintended consequences are observed in more than 49% of those programs. We propose a choice function for the colleges that removes any gain from strategizing over the privileges claimed, is fair, satisfies the substitutes condition and under reasonable assumptions on the type distribution of the population fully satisfies the diversity objectives expressed by the law. We also propose a strategy-proof mechanism that matches students and colleges with the use of the proposed choice function.

JEL classification: C78, D63, D78, D82
Keywords: Mechanism Design, Matching with Contracts, College admissions, Affirmative Action, Diversity.

*Address: Boğaziçi University Department of Economics Natuk Birkan Building, Room 226B Bebek, Istanbul 34342; website: http://www.orhanaygun.com; e-mail: orhan.aygun@boun.edu.tr; phone: +90 (212) 359 4830.

†Address: WZB Berlin Social Science Center, Reichpietschufer 50, D-10785 Berlin, Germany; website: http://www.inaciobo.com; e-mail: inacio.bo@wzb.eu; phone: +49 (0)30-25491-292.

Manipulating the Probabilistic Serial Rule

Haris Aziz, Serge Gaspers, Simon Mackenzie, Nicholas Mattei

NICTA and UNSW, Kensington 2033, Australia

Nina Narodytska

Carnegie Mellon University, Pittsburgh, PA 15213-3891, USA

Toby Walsh

NICTA and UNSW, Kensington 2033, Australia

Abstract

The probabilistic serial (PS) rule is one of the most prominent randomized rules for the assignment problem. It is well-known for its superior fairness and welfare properties. However, PS is not immune to manipulative behaviour by the agents. We initiate the study of the computational complexity of an agent manipulating the PS rule. We show that computing an expected utility better response is NP-hard. On the other hand, we present a polynomial-time algorithm to compute a lexicographic best response. For the case of two agents, we show that even an expected utility best response can be computed in polynomial time. Our result for the case of two agents relies on an interesting connection with sequential allocation of discrete objects.

Keywords: Assignment problem, probabilistic serial mechanism, fair allocation
JEL: C62, C63, and C78

Email addresses: `haris.aziz@nicta.com.au` (Haris Aziz), `serge.gaspers@nicta.com.au` (Serge Gaspers), `Simon.Mackenzie@nicta.com.au` (Simon Mackenzie), `Nicholas.Mattei@nicta.com.au` (Nicholas Mattei), `ninan@cs.toronto.edu` (Nina Narodytska), `toby.walsh@nicta.com.au` (Toby Walsh)

A new solution concept for the roommate problem: The $\mathcal{Q}$-stable matchings*

Peter Biró† Elena Inarra‡ Elena Molis§

Abstract

It is well-known that roommate problems (Gale and Shapley, *AMM*, 1962) can have stable matchings (solvable problems) or not (unsolvable problems). The aim of this paper is to propose a new solution concept for the class of roommate problems with strict preferences. We believe that it is essential to require a solution concept which provides a stable matching when it exists and some matching otherwise. Thus, we focus on core consistent solutions. Several solution concepts have been proposed explicitly for dealing with unsolvable problems, but there has yet to be any in-depth discussion regarding comparisons between solution concepts and there is scope for new ones.

In our work we first introduce *maximum irreversible matchings.* These matchings incorporate pairings so stable that once they are formed they never split. At the interface between Economics and Computer Science two solution concepts have been proposed explicitly for dealing with unsolvable problems. *Almost stable matchings* (Abraham *et al.*, *AO-L Algorithms*, 2006) form a subclass of Pareto optimal matchings with the minimum number of blocking pairs. *Maximum stable matchings* (Tan, *BIT*, 1990) single out matchings with the largest set of pairs that are stable within themselves.

All three of the solution concepts mentioned show sufficient grounds for consideration as good candidates for solving roommate problems. Hence, it seems to make sense to consider a proposal that could conciliate most if not all of those solution concepts. However, we find that almost stable matchings are incompatible with the other two concepts. Hence, to solve the roommate problem we propose matchings that lie at the intersection of the maximum irreversible matchings and maximum stable matchings, which we call *Q- stable matchings.* We construct an efficient algorithm for computing one element of this set for any roommate problem.

For the roommate problem Inarra *et al.* (*GEB*, 2013). seek to determine which matchings a decentralized process may lead to. They consider a dynamic process in which a matching is adjusted when a blocking pair of agents mutually decide to become partners. If there are stable matchings the process eventually converges to one of them. Otherwise it leads to a set of matchings (an absorbing set) such that any matching in the set can be obtained from any other and it is impossible to escape from the matchings in that set. Therefore it is important to investigate whether our proposal, the $\mathcal{Q}$-stable matching, is achievable from a free interactions of agents, *i.e.* whether it belongs to an absorbing set. The answer is in the affirmative and we show that although not all $\mathcal{Q}$-stable matchings belong to an absorbing set, any matching determined by our algorithm does.

*This research is supported by the Spanish Ministry of Science and Innovation (ECO2012-31346 and ECO2013-44879-R), co-funded by ERDF, by Basque Government (IT-568-13), by the Government of Andalusia (Project for Excellence in Research SEJ1436) and by COST action ICI205 on Computational Social Choice. Péter Biró also acknowledges the support from the Hungarian Academy of Sciences under its Momentum Programme (LD-004/2010), and the Hungarian Scientific Research Fund, OTKA, grant no. K108673.

†Institute of Economics, Research Centre for Economic and Regional Studies, Hungarian Academy of Sciences, H-1112, Budaörsi út 45, Budapest, Hungary, and Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest. E-mail: peter.biro@krtk.mta.hu

‡BRiDGE, Foundation of Economic Analysis I, University of the Basque Country, Avenida Lehendakari Aguirre 83, E-48015 Bilbao, Spain. E-mail: elena.inarra@ehu.es

§GLOBE, Department of Economics, University of Granada, Campus Cartuja s/n 18071 Granada, Spain. E-mail: emolis@ugr.es

The Stable Fixtures Problem with Payments

Péter Biró[1,2⋆], Walter Kern[3],
Daniël Paulusma[4⋆⋆], and Péter Wojuteczky[1]

[1] Institute of Economics, Hungarian Academy of Sciences; H-1112, Budaörsi út 45, Budapest, Hungary, peter.biro@krtk.mta.hu, peter.wojuteczky@gmail.com
[2] Department of Economics, Stanford University
[3] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, P.O.Box 217, NL-7500 AE Enschede
w.kern@math.utwente.nl
[4] School of Engineering and Computing Sciences, Durham University, Science Laboratories, South Road, Durham DH1 3LE, UK,
daniel.paulusma@durham.ac.uk

Abstract. We introduce multiple partners matching games, which consist of a graph $G = (N, E)$, with an integer vertex capacity function b and an edge weighting w. The set N consists of a number of players that are to form a set $M \subseteq E$ of 2-player coalitions ij with value w_{ij}, such that each player i is in at most b_i coalitions. A payoff p is a vector with $p(i,j)+p(j,i) = w_{ij}$ if $ij \in M$ and $p(i,j) = p(j,i) = 0$ if $ij \notin M$. The pair (M, p) is called a solution. A pair of players i, j with $ij \in E \setminus M$ blocks a solution (M, p) if i, j can form, possibly only after withdrawing from one of their existing 2-player coalitions, a new 2-player coalition in which they are mutually better off. A solution is stable if it has no blocking pairs.

We give a polynomial-time algorithm that either finds that no stable solutions exists, or obtains a stable solution. This generalizes a known result of Sotomayor (1992) for multiple partners assignment games, where the underlying graph G is bipartite. We also characterize the set of stable solutions and initiate a study on the core of the corresponding cooperative game, where coalitions of any size may be formed.

⋆ Visiting professor at Stanford University until 31 January 2015, supported by the Hungarian Academy of Sciences under Momentum Programme LD-004/2010 and also by OTKA grant no. K108673.
⋆⋆ Supported by EPSRC Grant EP/K025090/1.

The most ordinally-egalitarian of random voting rules

Anna Bogomolnaia
University of Glasgow, Scotland, UK
e-mail: anna.bogomolnaia@glasgow.ac.uk

February 2015

Abstract

We consider a random social choice model, where n agents are to jointly choose one alternative from a given finite set A. Agents might have arbitrary, not necessarily strict, preferences over A. An ordinal random choice mechanism specifies a probability distribution on A for any given profile of agents' preferences over sure outcomes in A.

In this setting, Aziz and Stursberg ("A Generalization of Probabilistic Serial to Randomized Social Choice", *Proceedings of the 28-th AAAI Conference on Artificial Intelligence* (2014), 559–565) propose an "Egalitarian Simultaneous Reservation" rule (ESR), a generalization of Serial rule, one of the most discussed mechanisms in random assignment problem, to this more general random social choice domain.

We provide an alternative definition, or characterization, of ESR as the unique most ordinally-egalitarian one. Specifically, given a lottery p over alternatives, for each agent i we define $t_i^p(k)$ to be the total share in p of objects from her first k indifference classes. ESR is shown to be the unique one which leximin maximizes the vector of all such shares $(t_i^p(k))_{i,k}$. We submit that this is a proper way to apply the Egalitarian Principle for random mechanisms with ordinal input.

Serial rule is known to be characterized by the same property (see Bogomolnaia, A., "Random Assignment: Redefining the Serial Rule" (2014), mimeo). Thus, we provide an alternative way to show that ESR, indeed, coincides with Serial rule on the assignment domain. Moreover, since both rules are defined as the unique most ordinally-egalitarian ones, out result shows that ESR is "the right way" to think about generalizing Serial rule.

Keywords: Random Social choice, Random assignment, Serial Rule, Leximin

DRIVEN BY PRIORITIES MANIPULATIONS UNDER THE BOSTON MECHANISM

DAVID CANTALA AND JUAN SEBASTIÁN PEREYRA

Abstract. Inspired by real-life manipulations used when the Boston mechanism is in place, we study school choice markets where students submit preferences driven by priorities; that is, when students declare among the most preferred those schools for which they have high priority. Under this assumption, we first prove that the outcome of the Boston mechanism is the school-optimal stable matching. Moreover, the condition is necessary: if the outcome of the Boston mechanism is the school-optimal stable matching, then preferences are driven by priorities. Thus, under these manipulations, the final allocation of students may be purely shaped by schools' priorities. Second, we analyze a situation where the Boston mechanism is replaced either by a stable mechanism or by the top trading cycles mechanism, but it takes some periods before students begin to behave truthfully. If during this transition students try to manipulate the new mechanism as other students did before, we show that the new matching will not present large changes respect to previous allocations. Additionally, we run some computational simulations to show that the assumption of driven by priorities preferences can be relaxed by introducing an idiosyncratic preference component, and our main results hold for almost all students.

Keywords: Two-sided many-to-one matching; school choice; Boston algorithm; manipulation strategies; Deferred Acceptance algorithm; Top trading cycles.

JEL Classification: C72; D47; D78; D82.

We are grateful to Federico Echenique for suggesting the computational simulations. We also thank to Estelle Cantillon, Li Chen, Alvaro Forteza and Antonio Miralles for their comments and suggestions, as well as participants at the 7th Workshop Matching in Practice and dECON-Uruguay. Pereyra gratefully acknowledges financial support from ERC grant 208535. Cantala is affiliated with the Centro de Estudios Económicos at El Colegio de México; Pereyra is affiliated with ECARES-Universitè Libre de Bruxelles and F.R.S.-FNRS, emails `dcantala@colmex.mx`, `jpereyra@ulb.ac.be`.

A New Efficiency Criterion for Probabilistic Assignments

Battal Doğan[1] and Kemal Yıldız[2]

[1]Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland; battaldogan@gmail.com

[2]Department of Economics, Bilkent University, 06800 Ankara, Turkey; kemal.yildiz@bilkent.edu.tr

For probabilistic assignment of objects, when only ordinal preference information is available, we propose the following efficiency criterion: a probabilistic assignment dominates another assignment if, whenever the latter assignment is utilitarian efficient at a utility profile consistent with the ordinal preferences, the former assignment is utilitarian efficient too; and there is a utility profile consistent with the ordinal preferences at which the latter assignment is not utilitarian efficient but the former assignment is utilitarian efficient. We provide a simple characterization of this domination relation. We show that, if preferences are strict (no agent is indifferent between two different objects), an sd-efficient assignment π sw-dominates another sd-efficient assignment π' if and only if π has a finer support, i.e. the set of agent-object pairs assigned with positive probability in π is a proper subset of the set of agent-object pairs assigned with positive probability in π'. If preferences are weak (indifferences are allowed), we extend the support of an assignment so that it possibly includes a pair of agent and object that are not assigned with positive probability provided that there is an "equivalent assignment" that includes the pair. Then, we show that an sd-efficient assignment π sw-dominates another sd-efficient assignment π' if and only if π has a finer extended support. A consequence of these results is that when preferences are strict, the only sw-efficient assignments are the Pareto efficient deterministic assignments; and when the preferences are weak, the only undominated assignments are the sd-efficient assignments where each agent is indifferent among the objects that he is assigned with positive probability. We revisit an extensively studied probabilistic assignment mechanism, the Probabilistic Serial rule (Bogomolnaia and Moulin [JET, 2001]), and show that it can be improved in efficiency without sacrificing fairness.

The Secure Boston Mechanism

Umut Dur[†] Robert G. Hammond[‡] Thayer Morrill[§]

December, 2014

Abstract

The two primary objections to the Boston Mechanism (BM) are that it is not strategy-proof and that sophisticated students benefit at the expense of naive students. However, it is an attractive algorithm from an optimization standpoint. We introduce an intuitive modification of BM that secures any school a student was initially guaranteed but otherwise prioritizes a student at a school based upon how she ranks it. This algorithm is less manipulatable than BM and provides some protection for naive students. Our main result is to show that an equilibrium, in undominated strategies, of this new algorithm Pareto dominates the Deferred Acceptance algorithm when the student optimal stable assignment is Pareto inefficient. Key Words: Boston Mechanism, School Choice, Assignment.

JEL Classification: C78, D61, D78, I20

[†]Department of Economics, North Carolina State University. Contact: udur@ncsu.edu
[‡]Department of Economics, North Carolina State University. Contact: robert_hammond@ncsu.edu.
[§]Department of Economics, North Carolina State University. Contact: thayer_morrill@ncsu.edu.

The Curse of Stability: Designing the Appeals Round in School Choice

Umut Mert Dur[1]

North Carolina State University

Onur Kesten[2]

Carnegie Mellon University

Abstract

Almost all school choice plans feature an additional supplementary round where students who declare to be dissatisfied with the outcome in the main round can appeal to be re-assigned. The largest of these plans, the current New York City (NYC) high school matching system, assigns students to the schools via Gale and Shapley Deferred Acceptance (DA) mechanism in the main round which is next followed by a supplementary round employing the Top Trading Cycles (TTC) mechanism. Although both DA and TTC are strategy-proof, the current two-round system does not eliminate students' incentives to be strategic in their reported choices. In particular, a student may misreport his preferences in the main round in order to be assigned to a school which he can later trade with a more desirable school in the supplementary round. We ask whether we can design two-round school choice systems with good incentive properties and show that the answer to this questions does not only depend on the properties of the specific mechanisms used in each round, but also whether or not the non-appealing students are taken into account in the supplementary round. Our results show that the deficiency of the NYC system can be mitigated by either reversing the order of the two mechanisms, or by applying DA in both rounds together with allowing non-appealing students to passively participate in the supplementary round.

JEL Classification: C78, D61, H75, I28

Key Words: Matching Theory, Market Design, School Choice Problem

[1]Address: 2801 Founders Drive 4102 Nelson Hall, Raleigh, NC 27695, USA. E-mail: udur@ncsu.edu

[2]Address: 5000 Forbes Ave, Pittsburgh, PA 15213, USA. E-mail: okesten@andrew.cmu.edu.

School Choice with Neighbors[1]

Umut Mert Dur[2]
North Carolina State University

Thomas Wiseman[3]
University of Texas at Austin

Abstract

We consider the school choice problem where students may prefer to be assigned to the same school as a neighbor. In that setting, the set of stable matchings can be empty. Moreover, there does not exist a strategy-proof mechanism satisfying even a much weaker stability notion. Instead, we show that a variation on the Top Trading Cycles mechanism is both strategy-proof and Pareto efficient, and that it is in a well-defined sense one of the "most stable" strategy-proof mechanism. We also present a modified Deferred Acceptance algorithm with improved stability properties.

JEL Classification: C78, D61, H75, I28

Key Words: Matching Theory, Market Design, School Choice Problem

[1] We thank Bettina Klaus, Bumin Yenmez, Fran Flanagan, Scott Kominers, Flip Klijn, Onur Kesten, Azer Abizade, and Thayer Morrill for useful comments. We are also grateful for comments that we got from the audience during presentations at Conference on Economic Design '13 and Coalition Theory Network Workshop '14.

[2] Address: 2801 Founders Drive 4102 Nelson Hall, Raleigh, NC, 27695; e-mail: udur@ncsu.edu; web page: https://sites.google.com/site/umutdur/

[3] Address: The University of Texas at Austin, Department of Economics, 2225 Speedway, Austin, TX, 78712; e-mail: wiseman@austin.utexas.edu; web page: https://webspace.utexas.edu/wisemant/www/

Object Allocation via Deferred-Acceptance: Strategy-Proofness and Comparative Statics*

Lars Ehlers† Bettina Klaus‡

Abstract

We study the problem of assigning indivisible and heterogenous objects (*e.g.*, houses, jobs, offices, school or university admissions *etc.*) to agents. Each agent receives at most one object and monetary compensations are not possible. We consider mechanisms satisfying a set of basic properties (*unavailable type invariance*, *individual rationality*, *weak non-wastefulness*, or *truncation-invariance*).

In the house allocation problem, where at most one copy of each object is available, deferred-acceptance (DA)-mechanisms allocate objects based on exogenously fixed objects' priorities over agents and the agent-proposing deferred-acceptance-algorithm. For house allocation we show that DA-mechanisms are characterized by our basic properties and (i) *strategy-proofness* and *population-monotonicity* or (ii) *strategy-proofness* and *resource-monotonicity*.

Once we allow for multiple identical copies of objects, on the one hand the first characterization breaks down and there are unstable mechanisms satisfying our basic properties and (i) *strategy-proofness* and *population-monotonicity*. On the other hand, our basic properties and (ii) *strategy-proofness* and *resource-monotonicity*characterize (the most general) class of DA-mechanisms based on objects' fixed choice functions that are acceptant, monotonic, substitutable, and consistent. These choice functions are used by objects to reject agents in the agent-proposing deferred-acceptance-algorithm. Therefore, in the general model *resource-monotonicity* is the "stronger" comparative statics requirement because it characterizes (together with our basic requirements and *strategy-proofness*) choice-based DA-mechanisms whereas *population-monotonicity* (together with our basic properties and *strategy-proofness*) does not.

*We are grateful to two anonymous referees, Battal Doğan, and Jan Christoph Schlegel for their helpful comments and suggestions. Lars Ehlers acknowledges financial support from the SSHRC (Canada) and the FQRSC (Québec) and Bettina Klaus acknowledges financial support from the Swiss National Science Foundation (SNSF).

†Département de Sciences Économiques and CIREQ, Université de Montréal, Montréal, Québec H3C 3J7, Canada; e-mail: lars.ehlers@umontreal.ca.

‡Faculty of Business and Economics, University of Lausanne, Internef 538, CH-1015, Lausanne, Switzerland.; e-mail: bettina.klaus@unil.ch.

Polyhedral aspects of stable b-matching*

Pavlos Eirinakis[1], Dimitrios Magos[2], Ioannis Mourtos[1]
[1]Department of Management Science & Technology,
Athens University of Economics and Business, 76 Patission Ave., 104 34 Athens, Greece
email: {peir,mourtos}@aueb.gr
[2]Department of Informatics,
Technological Educational Institute of Athens, Ag. Spyridonos Str., 12210 Egaleo, Greece
email: dmagos@teiath.gr

The theory of matroid-kernels and their corresponding sets of blockers and antiblockers can be utilized to obtain a linear description of the stable b-matching problem (MM) [4]. We utilize the relation between antiblockers and rotations [2] to revisit that description and establish the dimension of the MM polytope. Moreover, we provide a minimal representation of the MM polytope by identifying its minimal equation system and facet-defining inequalities. This representation includes $O(m)$ constraints, m being the number of pairs, hence being significantly sparser than the existing one and linear with respect to the size of the problem. This minimal representation carries over to the stable admissions problem (SA), for which we also establish the facial correspondence of the linear representation based on matroid-kernels to the one based on combs, thus making the separation algorithm appearing in [1] obsolete.

Besides bringing a closure to the polyhedral study of the MM and SA polytopes, the minimal representation established here can be of practical importance in variants of the MM and SA involving additional constraints, e.g., couples in residency schemes; it provides a minimal linear relaxation which can admit additional constraints per variant and can be used efficiently in the framework of general solution methods (e.g., Branch & Cut) in cases where problem specific combinatorial algorithms [3] become useless, since not having the versatility of linear relaxations.

References

[1] Baïou M., Balinski M.: The stable admissions polytope. Math Program 87(3), 427-439 (2000).

[2] Eirinakis P., Magos D., Mourtos I.: Blockers and Antiblockers of Stable Matchings, Theor Comp Sci 524, 126-133 (2014).

[3] Eirinakis P., Magos D., Mourtos I., Miliotis P.: Finding all stable pairs and solutions to the many-to-many Stable Matching problem. INFORMS J Comp 24 (2), 245-259 (2012).

[4] Fleiner T.: On the stable b-matching polytope. Math Soc Sci 46 (2), 149-158 (2003).

*This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Thales. Investing in knowledge society through the European Social Fund.

Social welfare in one-sided matchings: Random priority and beyond*

Aris Filos-Ratsikas[1], Søren Kristoffer Stiil Frederiksen[1], and Jie Zhang[2]

[1] Department of Computer Science, Aarhus University, Abogade 34, 8000 Aarhus N, Denmark
{filosra,ssf}@cs.au.dk
[2] Department of Computer Science, University of Oxford, Parks Road, OX1 3QD Oxford, U.K.
jie.zhang@cs.ox.ac.uk

Abstract. We study the problem of approximate social welfare maximization (without money) in one-sided matching problems when agents have unrestricted cardinal preferences over a finite set of items. *Random priority* is a very well-known *truthful-in-expectation* mechanism for the problem. We prove that the approximation ratio of random priority is $\Theta(n^{-1/2})$ while no truthful-in-expectation mechanism can achieve an approximation ratio better than $O(n^{-1/2})$, where n is the number of agents and items. Furthermore, we prove that the approximation ratio of all ordinal (not necessarily truthful-in-expectation) mechanisms is upper bounded by $O(n^{-1/2})$, indicating that random priority is asymptotically the best truthful-in-expectation mechanism and the best ordinal mechanism for the problem.

* The authors acknowledge support from the Danish National Research Foundation and The National Science Foundation of China (under the grant 61061130540) for the Sino-Danish Center for the Theory of Interactive Computation, within which this work was performed. The authors also acknowledge support from the Center for Research in Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council. Jie Zhang also acknowledges support from ERC Advanced Grant 321171 (ALGAME).

On weighted kernels of two posets*

Tamás Fleiner† Zsuzsanna Jankó‡

Sands, Sauer and Woodrow in [4] proved an interesting generalization of the stable marriage theorem by Gale and Shapley in [3]. This result can be formulated in terms of partially ordered sets as follows. If $\preceq_1$ and $\preceq_2$ are two partial orders on the same ground set V then there is a common antichain K of these posets such that for any element $v \in V \setminus K$, there exists a vertex $k \in K$ such that $v \preceq_1 k$ or $v \preceq_2 k$ holds. This antichain is also called a kernel.

Fix a poset $P = (V, \preceq)$ and a demand function $w : V \to \mathbb{R}_+$. Weight function $f : V \to \mathbb{R}_+$ is *$\preceq$-tame (with respect to w)* if the total weight of no chain exceeds the demand of its minimal element unless this minimal element has zero weight. We say that element v of V is *$\preceq$-dominated* by f if there is a chain starting at v of total weight not less than the demand of v.

Let $P_1 = (V, \preceq_1)$ and $P_2 = (V, \preceq_2)$ be posets on ground set V and $w : V \to \mathbb{R}_+$ be a demand function. Weight function $f : V \to \mathbb{R}_+$ is a *weighted kernel* if f is both $\preceq_1$-tame and $\preceq_2$-tame and moreover each element v of V is $\preceq_1$-dominated or $\preceq_2$-dominated (or both). The main result of Aharoni, Berger and Gorelik [1] states that there always exists a weighted kernel.

We generalize this result with the help of choice functions. Function $\mathcal{F} : X \to X$ is a *choice function* on lattice $L = (X, \preceq)$ if $\mathcal{F}(x) \preceq x$ holds for any element x of X. Mapping $\mathcal{F} : X \to X$ is *antitone* if $x \preceq y$ implies $\mathcal{F}(y) \preceq \mathcal{F}(x)$. Function $\mathcal{D} : X \to X$ is a *determinant* of $\mathcal{F}$ if $\mathcal{F}(x) = x \wedge \mathcal{D}(x)$ holds for any element x of X. Choice function $\mathcal{F} : X \to X$ is called *substitutable* if there is an antitone determinant $\mathcal{A} : X \to X$ of $\mathcal{F}$, and $\mathcal{F} : X \to X$ is *path-independent* if $\mathcal{F}(x \vee y) = \mathcal{F}(x \vee \mathcal{F}(y))$ holds for any elements x, y of X. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are path-independent substitutable choice functions on lattice $L = (X, \preceq)$ then $s \in X$ is *$\mathcal{F}_1\mathcal{F}_2$-stable* if $\mathcal{F}_1(s) = \mathcal{F}_2(s) = s$ and $\mathcal{F}_1(s \vee x) \wedge \mathcal{F}_2(s \vee x) \preceq s$ holds for each element x of X.

Based on Tarski's fixed point theorem [5], we show that there always exist a $\mathcal{F}_1\mathcal{F}_2$-stable element. We generalize Blair's theorem [2] to our setting and prove that weighted kernels form a lattice under a certain natural partial order. To illustrate the robustness of our approach we indicate other possibilities for generalizing the result by Sands, Sauer and Woodrow that can be done by picking different path-independent substitutable choice functions.

References

[1] Aharoni, R., Berger, E., Gorelik, I.: Kernels in Weighted Digraphs. Order **31**(1), 35–43 (2014).

[2] Blair, C.: The lattice structure of the set of stable matchings with multiple partners. Math. Oper. Res. **13**(4), 619–628 (1988)

[3] Gale, D., Shapley, L.: College admissions and stability of marriage. Amer. Math. Monthly **69**(1), 9–15 (1962)

[4] Sands, B., Sauer, N., Woodrow, R.: On monochromatic paths in edge-coloured digraphs. J. Combin. Theory Ser. B **33**(3), 271–275 (1982)

[5] Tarski, A.: A lattice-theoretical fixpoint theorem and its applications. Pacific J. of Math **5**, 285–310 (1955)

*Supported by OTKA K108383 and OTKA K109240

†Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Magyar tudósok körútja 2., Budapest, Hungary, H-1117 and MTA-ELTE Egerváry Research Group.

‡Eötvös Loránd University, Faculty of Science, Institute of Mathematics, Pázmány Péter sétány 1/C, Budapest, Hungary, H-1117.

Market Design under Distributional Constraints: Diversity in School Choice and Other Applications*

Daniel Fragiadakis[1] and Peter Troyan[2]

[1]*Department of Economics, Texas A&M University, 4228 TAMU, College Station, Texas 77843*
dfragiadakis@tamu.edu
[2]*Department of Economics, University of Virginia, PO Box 400182, Charlottesville, Virginia 22904*
troyan@virginia.edu

Distributional constraints are important in many market design settings. Prominent examples include the minimum manning requirements at each Army branch in military cadet matching and diversity considerations in school choice, whereby school districts impose constraints on the demographic distribution of students at each school. Standard assignment mechanisms implemented in practice are unable to accommodate these constraints. This leads policymakers to resort to ad-hoc solutions that eliminate blocks of seats ex-ante (before agents submit their preferences) to ensure that all constraints are satisfied ex-post (after the mechanism is run). We show that these current solutions ignore important information contained in the submitted preferences, resulting in avoidable inefficiency. We then introduce new dynamic quotas mechanisms that result in Pareto superior allocations while at the same time respecting all distributional constraints and satisfying important fairness and incentive properties. We expect the use of our mechanisms to improve the performance of matching markets with distributional constraints in the field.

*We are extremely greatful to Fuhito Kojima, Muriel Niederle, and Al Roth for numerous conversations regarding this project. We would like to thank Brian Baisa, Scott Baker, Andrey Fradkin, Guillaume Haeringer, Matt Jackson, Paul Milgrom, Krishna Rao, Itay Saporta-Eksten, Ilya Segal, Pablo Villanueva, Alex Westkamp, Alex Wolitzky, and Edison Yu for helpful comments and suggestions, as well as seminar participants at Virginia, Iowa, Cal State East Bay, Western Ontario, and Illinois. Generous support from the B.F. Haley and E.S. Shaw Fellowship (Troyan) and the Leonard W. Ely and Shirley R. Ely Fellowship (Fragiadakis), both through the Stanford Institute for Economic Policy Research, is gratefully acknowledged

Robust models for the Kidney Exchange Problem

Kristiaan Glorie[1], Margarida Carvalho[2], Miguel Constantino[3], Paul Bouman[4], and Ana Viana[5]

[1]Department of Mathematics, VU University Amsterdam, km.glorie@vu.nl
[2]INESC TEC and Faculdade de Ciências da Universidade do Porto
[3]Departamento de Estatística e Investigação Operacional, Faculdade de Ciências da Universidade de Lisboa
[4]Rotterdam School of Management, Erasmus University Rotterdam
[5]INESC TEC and School of Engineering, Politechnic of Porto

Abstract

We consider the clearing of barter exchange markets in which proposed transactions must be verified before they can proceed. Proposed transactions may fail to go forward if verification fails or if a participant withdraws. The clearing problem for these markets is a combinatorial optimization problem that can be modelled as a vertex-disjoint cycle packing problem in an unreliable digraph. The arcs and nodes of this graph are subject to failure.

Our research finds a natural application in kidney exchange markets, which aim to enable transplants between incompatible donor-patient pairs. A set of pairs must be chosen in such a way that each selected patient can receive a kidney from a compatible donor from another pair in the set. The pairs are then notified and crossmatch tests must be performed to ensure the success of the transplants. We study the case in which if incompatibilities are discovered, a partaker has to withdraw and a new set of pairs may be selected. The new set should be as close as possible to the initial set in order to minimize the material and emotional costs of the alteration. Various recourse policies that determine the allowed post-matching actions are proposed. For each recourse policy, a robust model is developed. Besides the development of a novel adjustable robust optimization model, our contribution includes techniques to solve exactly the optimization problems in hand.

College Admissions with Entrance Exams: Centralized versus Decentralized[1]

Isa Hafalir Rustamdjan Hakimov Dorothea Kubler Morimitsu Kurino[2]

Extended Abstract

We theoretically and experimentally study a college admissions problem in which colleges accept students by ranking students' efforts in entrance exams. Students' ability levels affect the cost of their efforts. We solve and compare equilibria of ``centralized college admissions'' (CCA) where students apply to all colleges, and ``decentralized college admissions'' (DCA) where students only apply to one college.

After solving for the equilibrium of CCA and DCA, we compare the equilibria in terms of students' interim expected utilities. We show that students with lower abilities prefer DCA to CCA when the number of seats is smaller than the number of students. The main intuition for this result is that students with very low abilities have almost no chance of getting a seat in CCA, whereas their probability of getting a seat in DCA is bounded away from zero due to the fewer number of applications than the capacity. Moreover, we show that students with higher abilities prefer CCA to DCA. The main intuition for this result is that high-ability students (i) can only get a seat in the good school in DCA, whereas they can get seats in both the good and the bad school in CCA, and (ii) their equilibrium probability of getting a seat in the good school is the same across the two mechanisms.

We test the theory with the help of lab experiments. We implement five markets for the college admissions game that are designed to capture different levels of competition (in terms of the supply of seats, the demand ratio, and the quality difference between the two colleges). We compare the two college admission mechanisms and find that in most (but not all) markets, the comparisons of the students' ex-ante expected utilities, their effort levels, and the students' preferences regarding the two college admission mechanisms are well organized by the theory. However, the experimental subjects exert a higher effort than predicted. The overexertion of effort is particularly pronounced in DCA, which makes it relatively less attractive for the applicants compared to CCA.

---

[1] We would like to thank Ken Binmore, Francis Bloch, David Danz, Youngwoo Koh, Fuhito Kojima, Kai Konrad, Vijay Krishna, Ariel Rubinstein, Aner Sela, Ron Siegel, Alistair Wilson and participants at the 14th SAET conference at Waseda and the workshop ``Designing Matching Markets'' at WZB Berlin for helpful discussions as well as Nina Bonge for programming and helping us to run the experiments. Hafalir acknowledges financial support from National ScienceFoundation grant SES-1326584. All remaining errors are our own.

[2] E-mails and affiliations: isaemin@cmu.edu , Tepper School of Business (Hafalir), rustamdjan.hakimov@wzb.eu, WZB, (Hakimov), kuebler@wzb.eu , WZB (Kubler), kurino@sk.tsukuba.ac.jp , University of Tsukuba (Kurino)

Full Substitutability in Trading Networks*

John William Hatfield† Scott Duke Kominers‡ Alexandru Nichifor§
Michael Ostrovsky¶ Alexander Westkamp‖

Abstract

The trading network framework generalizes and unifies models of matching with bilateral contracts and indivisible goods exchange. We extend earlier models' canonical definitions of substitutability to that framework and show that all these definitions are equivalent. We also show that substitutability corresponds to submodularity of the indirect utility function, the single improvement property, and a no complementarities condition. We prove that substitutability is preserved under economically important transformations such as trade endowments, mergers, and limited liability. Finally, we show that substitutability implies monotonicity conditions called the Laws of Aggregate Supply and Demand.

JEL classification: C78; C71; D47; D85; L14

Keywords: Matching; Exchange Economies; Auctions; Trading Networks; Substitutes; Submodularity; Law of Aggregate Demand

---

*We thank Susan Athey, Peter Cramton, Vincent P. Crawford, Drew Fudenberg, Albert Gallegos, Çağatay Kayı, Paul Klemperer, Paul R. Milgrom, Benny Moldovanu, Philip Reny, Alvin E. Roth, Michael Schwarz, William Thomson, Utku Ünver, E. Glen Weyl, and many seminar and conference participants for helpful comments and suggestions. Part of this work was conducted while Hatfield was at Stanford University, Kominers was at the Becker Friedman Institute for Research in Economics at the University of Chicago, and Westkamp was at the University of Bonn. Hatfield and Nichifor appreciate the hospitality of Harvard Business School, Kominers appreciates the hospitality of Microsoft Research New England, and Nichifor and Westkamp appreciate the hospitality of the Stanford University. Kominers thanks the National Science Foundation (grant CCF-1216095 and a graduate research fellowship), the Harvard Milton Fund, the Yahoo! Key Scientific Challenges Program, the John M. Olin Center (a Terence M. Considine Fellowship), the American Mathematical Society, and the Simons Foundation for support. Nichifor received funding from the People Programme (Marie Curie Outgoing International Fellowship) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement 625718. Ostrovsky thanks the Alfred P. Sloan Foundation for support. Westkamp thanks the German Science Foundation for support.

†McCombs School of Business, University of Texas at Austin; 2110 Speedway, Austin, TX 78705, USA; `john.hatfield@utexas.edu`.

‡Society of Fellows, Department of Economics, Program for Evolutionary Dynamics, and Center for Research on Computation and Society, Harvard University, and Harvard Business School; Rock Center, Harvard Business School, Soldiers Field, Boston, MA 02163, USA; `kominers@fas.harvard.edu`.

§School of Economics and Finance, University of St Andrews, and Economics Department, Stanford University; Room 344, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305, USA; `nichifor@stanford.edu`.

¶Graduate School of Business, Stanford University; 655 Knight Way, Stanford, CA 94305, USA; `ostrovsky@stanford.edu`.

‖Economics Department AE1, Maastricht University; P.O. Box 616, 6200MD Maastricht, The Netherlands; `acwestk@gmail.com`.

Strategy-Proofness and Stability for Matching with Contracts*

John William Hatfield† Scott Duke Kominers‡ Alexander Westkamp§

Abstract

We consider the setting of many-to-one matching with contracts, where firms may demand multiple contracts but each worker desires at most one contract. We introduce three novel conditions—observable substitutability, observable size monotonicity, and non-manipulatability—and show that when these conditions are satisfied, a stable and strategy-proof (for workers) mechanism exists. Moreover, we show that when any of our three conditions fails, one may construct preferences for the doctors and unit-demand choice functions for the other firms such that no stable and strategy-proof mechanism exists. Finally, we show that, whenever our three conditions are satisfied, the outcome of any stable and strategy-proof mechanism coincides with the cumulative offer process.

JEL Classification: C62; C78; D44; D47

Keywords: Matching with contracts, Stability, Strategy-proofness, Substitutability, Size monotonicity

---

*The authors thank Fuhito Kojima, Alexey Kushnir, and Paul Milgrom for helpful comments. Kominers gratefully acknowledges the hospitality of Microsoft Research New England, as well as the support of National Science Foundation grant CCF-1216095 and the Harvard Milton Fund. Westkamp thanks the European Union for support via a Marie-Curie Intra-European fellowship for career development under REA grant agreement 628276.

†McCombs School of Business, University of Texas at Austin; 2110 Speedway, Austin, TX 78705, USA; `john.hatfield@utexas.edu`.

‡Society of Fellows, Department of Economics, Program for Evolutionary Dynamics, and Center for Research on Computation and Society, Harvard University, and Harvard Business School; Rock Center, Harvard Business School, Soldiers Field, Boston, MA 02163, USA; `kominers@fas.harvard.edu`.

§Economics Department AE1, Maastricht University; P.O. Box 616, 6200MD Maastricht, The Netherlands; `acwestk@gmail.com`.

On The Structure of Popular Matchings in The Stable Marriage Problem —Who Can Join a Popular Matching?

Mizuki Hirakawa   Yukiko Yamauchi   Shuji Kijima   Masafumi Yamashita

Graduate School of Information Science and Electrical Engineering
Kyushu University
744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
{mizuki.hirakawa, yamauchi, kijima, mak}@inf.kyushu-u.ac.jp

Gale and Sotomayor (1985) gave a remark on the *stable marriage problem* that the set of people who are matched with themselves is the same for all stable matchings. Motivated by a larger cardinality matching, Huang and Kavitha (2013) investigated the structure of *popular matchings*, an extended notion of stable matching: Given an instance of the stable marriage problem, a matching M is *popular* if there is no matching M' such that more vertices are better off in M' than in M. They established the Gale-Sotomayor's type theorem for *minimum* cardinality popular matchings, and showed that any stable matching is a minimum cardinality popular matching.

We establish the same type of the theorem for *maximum* cardinality popular matchings. To be precise, we show the following.

Theorem 1. *Let M be an arbitrary max popular matching. Then, $V(M') \subseteq V(M)$ holds for any popular matching M', where $V(M)$ denotes the set of end vertices of a matching M.*

Theorem 1 implies that the family $\mathcal{V}$ of sets of endpoints of popular matchings has the (unique) maximum and minimum with respect to the inclusion relation, combining with Huang-Kavitha's result. As a consequence, one may naturally presume that $\mathcal{V}$ is closed under intersection and union, and then $\mathcal{V}$ forms a (distributive) lattice. We disprove the former presumption, as follows:

Proposition 2. *There exists an instance of the stable marriage problem which has a pair of popular matchings M_1 and M_2 such that no popular matching M satisfies $V(M) = V(M_1) \cap V(M_2)$.*

Proposition 3. *There exists an instance of the stable marriage problem which has a pair of popular matchings M_1 and M_2 such that no popular matching M satisfies $V(M) = V(M_1) \cup V(M_2)$.*

References

[1] D. Gale, M. Sotomayor: Some remarks on the stable matching problem, Discrete Appl. Math., **11** (1985), 223–232.

[2] C.C. Huang, T. Kavitha: Popular matching in the stable marriage problem, Inf. Comput., 222 (2013), 180–194.

Time Horizons, Lattice Structures, and Welfare in Multi-period Matching Markets*

Sangram V. Kadam† Maciej H. Kotowski‡

Abstract

Consider a T-period, bilateral matching economy without monetary transfers. Under natural restrictions on agents' preferences, which accommodate switching costs, status-quo bias, and other forms of inter-temporal complementarity, dynamically-stable matchings exist. Generally, "optimal" dynamically-stable matchings may not exist, but under a suitable partial order the stable set forms a lattice. The welfare properties of different stable outcomes is ascertained and the implications for normative market-design are discussed. The robustness of dynamically-stable matchings with respect to the market's time horizon is examined.

*Kadam gratefully acknowledges the support of the Danielian Travel & Research Grant.

†Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge MA 02138. E-mail: <svkadam@fas.harvard.edu>

‡John F. Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge MA 02138. E-mail: <maciej_kotowski@hks.harvard.edu>

Matroid Generalizations of the Popular Matching and Condensation Problems with Strict Preferences

Naoyuki Kamiyama

Institute of Mathematics for Industry, Kyushu University
kamiyama@imi.kyushu-u.ac.jp

Abstract. The popular matching problem introduced by Abraham, Irving, Kavitha, and Mehlhorn [1] is a matching problem in which there exist applicants and posts, and applicants have preference lists over posts. A matching M is said to be popular, if there exists no other matching N such that the number of applicants that prefer N to M is larger than the number of applicants that prefer M to N. The concept of popularity was introduced by Gärdenfors [2] in the context of matching problems with two-sided preference lists. The goal of the popular matching problem is to decide whether there exists a popular matching in a given instance, and find a popular matching if one exists. Abraham, Irving, Kavitha, and Mehlhorn [1] presented polynomial-time algorithms for this problem. Since their seminal paper, several extensions of the popular matching problem have been investigated [6, 7, 8]. In this talk, we first consider a matroid generalization of the popular matching problem with strict preferences, and present a polynomial-time algorithm for this problem.

Unfortunately, it is known [1] that a given instance of the popular matching problem may admit no popular matching. For coping with such instances, several alternative solutions were presented. Kavitha and Nasre [4] considered the problem of deciding capacities of posts so that a given instance has a popular matching. Kavitha, Nasre, and Nimbhorkar [5] considered the problem of augmenting capacities of posts with minimum costs. These problems are hard in general. Wu, Lin, Wang, and Chao [9] considered the popular condensation problem whose goal is to transform a given instance by deleting a minimum number of applicants so that it has a popular matching, and gave a polynomial-time algorithm for this problem. In the second half of this talk, we consider a matroid generalization of the popular condensation problem with strict preferences, and give a polynomial-time algorithm for this problem.

The main results of this talk have appeared in [3].

References

[1] D. J. Abraham, R. W. Irving, T. Kavitha, and K. Mehlhorn. Popular matchings. *SIAM Journal on Computing*, 37(4):1030–1045, 2007.

[2] P. Gärdenfors. Match making: Assignments based on bilateral preferences. *Behavioral Science*, 20(3):166–173, 1975.

[3] N. Kamiyama. The popular matching and condensation problems under matroid constraints. In *COCOA*, volume 8881 of *LNCS*, pages 713–728, 2014.

[4] T. Kavitha and M. Nasre. Popular matchings with variable item copies. *Theoretical Computer Science*, 412(12):1263–1274, 2011.

[5] T. Kavitha, M. Nasre, and P. Nimbhorkar. Popularity at minimum cost. *Journal of Combinatorial Optimization*, 27(3):574–596, 2014.

[6] D. F. Manlove and C. T. S. Sng. Popular matchings in the capacitated house allocation problem. In *ESA*, volume 4168 of *LNCS*, pages 492–503, 2006.

[7] J. Mestre. Weighted popular matchings. *ACM Transactions on Algorithms*, 10(1):Article 2, 2014.

[8] C. T. S. Sng and D. F. Manlove. Popular matchings in the weighted capacitated house allocation problem. *Journal of Discrete Algorithms*, 8(2):102–116, 2010.

[9] Y.-W. Wu, W.-Y. Lin, H.-L. Wang, and K.-M. Chao. An optimal algorithm for the popular condensation problem. In *IWOCA*, volume 8288 of *LNCS*, pages 412–422, 2013.

The size of the core in assignment markets

Yash Kanoria Daniela Saban Jay Sethuraman*

Abstract

Assignment markets involve matching with transfers, as in labor markets and housing markets. We consider a two-sided assignment market with agent types and stochastic structure similar to models used in empirical studies, and characterize the size of the core in such markets. Each agent has a randomly drawn productivity with respect to each type of agent on the other side. The value generated from a match between a pair of agents is the sum of the two productivity terms, each of which depends only on the type but not the identity of one of the agents, and a third deterministic term driven by the pair of types. We allow the number of agents to grow, keeping the number of agent types fixed. Let n be the number of agents and K be the number of types on the side of the market with more types. We find, under reasonable assumptions, that the relative variation in utility per agent over core outcomes is bounded as $O^*(1/n^{1/K})$, where polylogarithmic factors have been suppressed. Further, we show that this bound is tight in worst case. We also provide a tighter bound under more restrictive assumptions.

Keywords: Assignment markets, matching, transferable utility, core, uniqueness of equilibrium, random market.

*All the authors are at Columbia University. Emails: {ykanoria,dhs2131,js1353}@columbia.edu

Overlapping Multiple Assignments

Jörgen Kratz

Department of Economics

Lund University, P.O. Box 7082, 222 07 Lund, Sweden

E-mail: jorgen.kratz@nek.lu.se

Abstract

This paper studies an allocation problem with multiple assignments, indivisible objects, no endowments and no monetary transfers, where a single object may be assigned to several agents as long as the set of agents assigned the object satisfy a compatibility constraint. The assignments of two agents are said to overlap if they have non-empty intersection. Due to the restrictions imposed on the compatibility structure, the set of agents can be partitioned into groups of compatible agents, such that each agent is incompatible with every agent belonging to a different group. An object may be assigned to any number of compatible agents, but it may never be assigned to a set containing incompatible agents. Only direct mechanisms are considered in this paper. Agents report their preferences over bundles of objects and a rule selects an allocation of objects to agents. On the domain of complete, transitive and strict preferences, it is shown that group-sorting sequential dictatorships are the only rules that are coalitionally strategyproof, Pareto efficient and group-monotonic. This characterization still holds if coalitional strategyproofness is replaced by strategyproofness and nonbossiness or if group-monotonicity is replaced by group-invariance. A sequential dictatorship is group-sorting if the priority structure associated with the rule is sorted by groups of compatible agents until every object has been assigned to at least one agent. When assignments are not allowed to overlap, it has been demonstrated by Pápai (2001) that a rule is strategyproof, Pareto efficient and non-bossy if and only if it is a sequential dictatorship. This result is contained as a special case of the characterization of group-sorting sequential dictatorships above. Finally, some different properties featured in various characterizations of serial dictatorships for similar allocation problems without overlapping assignments are considered. It is shown that neither serial dictatorships nor group-sorting sequential dictatorships are consistent or population-monotonic when assignments are allowed to overlap. Furthermore, on the domain of complete, transitive and strict preferences, there exists no rule that satisfies both Pareto efficiency and resource-monotonicity. This result holds regardless of whether assignments are allowed to overlap.

References

Pápai, S. (2001), "Strategyproof and nonbossy multiple assignments." *Journal of Public Economic Theory*, 3, 258–271.

Two School Systems, One District: What to do when a unified admissions process is impossible*

Vikram Manjunath† and Bertan Turhan‡

When groups of schools within a single district run their admission processes independently of one another, the resulting match is often inefficient: many children are left unmatched and seats are left unfilled.

In a context where school priorities are to be respected, we study the problem of re-matching students to take advantage of these empty seats in a context where there are priorities to respect. We propose an iterative way in which each group may independently match and re-match students to its schools.

The advantages of this process are that every iteration leads to a Pareto improvement and a reduction in waste while maintaining respect of the priorities. Furthermore, it reaches a non-wasteful match within a finite number of iterations.

While iterating may be costly, as it involves asking for inputs from the children, there are significant gains from the first few iterations. We show this analytically for two stylized problems. Both involve a continuum of children but a finite number of schools. The priority of a child at each school is drawn randomly from the uniform distribution. The first stylized problem is where every child has the same preferences over this finite set of schools. The second is where each child's preference is randomly drawn from the uniform distribution. More general problems where a child's preferences are informed by a convex combination of a private value and a common value are not analytically tractable. Instead, we confirm this result through simulations.

*We thank M. Oğuz Afacan, Ahmet Alkan, Samson Alva, Orhan Aygun, Isa Hafalir, Parag Pathak, William Phan, Erin Richards, Tayfun Sönmez, Utku Ünver, and Rodrigo Velez for helpful comments and suggestions. Manjunath gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada.

†Texas A&M University, vmanjunath@tamu.edu

‡Boston College, turhan@bc.edu

Contracts versus Salaries in Matching: A General Result

Jan Christoph Schlegel*
Faculty of Business and Economics, University of Lausanne, Switzerland
jschlege@unil.ch

Abstract

It is shown that a matching market with contracts may be embedded into a matching market with salaries under weaker conditions than substitutability of contracts. In particular, the result applies to the recently studied problem of cadet-to-branch matching. As an application of the embedding result, a new class of mechanisms for matching markets with contracts is defined that generalize the firm-proposing deferred acceptance algorithm to the case where contracts are unilateral substitutes for firms. *JEL-classification:* C78
Keywords: Matching; Matching with contracts; Matching with salaries; Embedding; Substitutes; Unilateral substitutes; Bilateral substitutes

*I am very grateful to my supervisor Bettina Klaus for many helpful discussions and comments. I thank William Thomson for detailed comments that greatly improved the exposition of the results. I thank Battal Doğan, Federico Echenique, Lars Ehlers, Flip Klijn, Fuhito Kojima, Scott Kominers, Tayfun Sönmez, participants of the 2014 Meeting of the Social Choice and Welfare Society and workshop participants in Marseille and Montréal for comments on a previous version of this paper. I gratefully acknowledge financial support by the Swiss National Science Foundation (SNSF) and by the COST Action IC1205 on Computational Social Choice. The paper extends and replaces a previous comment that circulated under the title: "Contracts versus Salaries in Matching: Comment"

Trading networks with bilateral contracts

Alexander Teytelboym*

Abstract

We consider general networks of bilateral contracts that include supply chains (Ostrovsky, 2008; Westkamp, 2010; Hatfield and Kominers, 2012). We define a new stability concept called *path stability* and show that *any* network of bilateral contracts has a path-stable outcome whenever agents' preferences satisfy full substitutability (same-side substitutability and cross-side complementarity). In supply chains, path stability is equivalent to chain stability (Ostrovsky, 2008). However, in general contract networks, path-stable outcomes may not be immune to group deviations or efficient. We examine previous results on (group) strategy-proofness and the rural hospitals theorem. When contracts specify trades and prices (Hatfield et al. 2013), we also show that competitive equilibrium exists in networked markets even in the absence of transferrable utility. The competitive equilibrium outcome is path-stable.

*Institute for New Economic Thinking, University of Oxford, Walton Well Road, Oxford, OX2 6ED, United Kingdom. Email: alexander.teytelboym@inet.ox.ac.uk. I would like to thank Vincent Crawford, Jens Gudmundsson, Claudia Herrestahl, Paul Klemperer, Collin Raymond, and Zaifu Yang for their valuable comments on a much earlier draft that appeared in my D.Phil. thesis "Essays on Networks and Market Design" (2013). I also benefitted from comments, especially from Samson Alva, Umut Dur, Tamás Fleiner, Alex Westkamp and M. Bumin Yenmez, during talks at the Southern Methodist University, National University of Singapore, CIREQ Matching Conference (Montréal), Workshop on Coalitions and Networks (Montréal), and the 12th Meeting of the Society of Social Choice and Welfare (Chestnut Hill).

Matroidal Choice Functions

Yu YOKOI

Department of Mathematical Informatics
Graduate School of Information Science and Technology
University of Tokyo
yu_yokoi@mist.i.u-tokyo.ac.jp

Abstract

Choice functions are common tools in many-to-one or many-to-many matching models. They can represent more general preferences of agents than the classical form that consists of a list and a quota. Choice functions are usually assumed to satisfy the substitutability, which is an essential condition for the existence of stable matchings.

In this paper, we introduce "matroidal choice functions" as a class of choice functions which satisfy a kind of matroid constraints in addition to the substitutability. We show that matroidal choice functions admit succinct representations, with which one can find a stable matching efficiently utilizing a greedy algorithm for matroids.

Furthermore, we show that matroidal choice functions afford nice properties of stable matchings such as the strategy-proofness of the deferred acceptance algorithm, and the distributive lattice structure of the set of stable matchings.

The full version is available at http://www.keisu.t.u-tokyo.ac.jp/research/techrep/data/2014/METR14-32.pdf.

Poster Abstracts

On the Susceptibility of the Deferred Acceptance Algorithm

Haris Aziz
NICTA and UNSW
Australia
haris.aziz@nicta.com.au

Hans Georg Seedig
TU München
Germany
seedig@in.tum.de

Jana Karina von Wedel
TU München
Germany
wedel@ma.tum.de

ABSTRACT

The Deferred Acceptance Algorithm (DAA) is the most widely accepted and used algorithm to match students, workers, or residents to colleges, firms or hospitals respectively. In this paper, we consider for the first time, the complexity of manipulating DAA by agents such as colleges that have capacity more than one. For such agents, truncation is not an exhaustive strategy. We present efficient algorithms to compute a manipulation for the colleges when the colleges are proposing or being proposed to. We then conduct detailed experiments on the frequency of manipulable instances in order to get better insight into strategic aspects of two-sided matching markets. Our results bear somewhat negative news: assuming that agents have information other agents' preference, they not only often have an incentive to misreport but there exist efficient algorithms to find such a misreport.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems; J.4 [**Computer Applications**]: Social and Behavioral Sciences - Economics

Keywords

Stable matchings, Gale-Shapley algorithm, matching markets, college admission.

Maximin Envy-Free Division of Indivisible Items

Steven J. Brams[1], D. Marc Kilgour[2], Christian Klamler[3]

Abstract

In this paper, we assume that two players strictly rank a set of indivisible items from best to worst. If there is an envy-free allocation of these items, two algorithms, AL (Brams, Kilgour, and Klamler, 2014) and SA (Brams, Kilgour, and Klamler, 2015), have been proposed for finding such an allocation. However, neither algorithm guarantees that an allocation will be *maximin*—one that maximizes the ranking of the players' lowest-ranked items.

We propose a new algorithm, SD, which provides this guarantee. The allocation it yields is envy-free, based on an item-wise definition of envy-freeness, if there exists an envy-free allocation. If there is no such allocation, an SD allocation will still be maximin.

In the paper we also define four properties of fair division—Pareto-optimality, envy-freeness, maximinality, and Borda maximinality—that we use to assess the fairness of maximin allocations. We prove two lemmas about their *maximin depth*, which is the lowest rank, of either player, of a maximin allocation. Furthermore, we give an algorithm for determining all maximin allocations, which we illustrate with an example.

The maximin algorithm is applied to several examples to determine which, if any, of the maximin allocations is envy-free. We provide two conditions for the existence of an envy-free allocation, the second of which simplifies the first. It is then proven that if there exists an envy-free allocation that is not maximin, there is always one that is maximin.

The first algorithm we propose is *single-stage SD*, which ensures that a maximin allocation is envy-free if there is an envy-free allocation. We show how SD can be revised by applying it in later stages to items not allocated earlier—unless both players rank one of the remaining items last—which we call *multi-stage SD*.

In addition we show that maximin, envy-free allocations may not satisfy other properties, such as Borda maximinality and Pareto-optimality. Although not strategyproof, the SD algorithms would be difficult to manipulate unless one player has complete information about the preference rankings of the other player.

Finally, we offer some thoughts on the relative merits of SD, SA, and AL. Although SD may require the application of AL to some, if not all, the items, SD, especially the multi-stage version, is generally simpler to compute than AL. It is preferable to SA if one wishes to ensure that the allocation is maximin. We conclude by suggesting SD's applicability to real-world problems, such as assigning people to committees and allocating the marital property in a divorce.

---

[1] Department of Politics, New York University, New York, NY 10012, USA, steven.brams@nyu.edu

[2] Department of Mathematics, Wilfrid Laurier University, Waterloo, Ontario N2L 3C5, CANADA, mkilgour@wlu.ca

[3] Institute of Public Economics, University of Graz, A-8010 Graz, AUSTRIA, christian.klamler@uni-graz.at

Dynamics of Swaps in House Markets

A. Damamme[1], A. Beynier[1], Y. Chevaleyre[2], and N. Maudet[1]

[1] Sorbonne Universités, UPMC Univ Paris 06
CNRS, UMR 7606, LIP6
F-75005, Paris, France
anastasia.damamme@lip6.fr aurelie.beynier@lip6.fr nicolas.maudet@lip6.fr
[2] Université Paris 13, Sorbonne Paris Cité
LIPN, CNRS, UMR 7030, F-93430, Villetaneuse, France
yann.chevaleyre@lipn.univ-paris13.fr

Abstract. A *house market* is a resource allocation setting consisting in assigning exactly one resource per agent, with each agent initially owning one such resource. In this setting, the *top trading cycle* procedure stands out as the uncontroversial method of choice, since it satisfies key desirable properties: *Pareto-efficiency*, *individual rationality*, and *strategy proofness*. It remains however a centralized procedure which may not well suited in the context of multiagent systems, where distributed coordination may be problematic. In this paper, we investigate the power of dynamics based on sequences of rational bilateral deals (*swaps*) in such settings. Agents randomly meet in a pairwise fashion, and contract a deal with their partner if exchanging their resources proves to be mutually beneficial. The process iterates until a stable state (an equilibrium) is reached. The same resource can thus successively be held by several agents over the sequence. While it is clear that they may induce a high efficiency loss (in particular, Pareto-efficiency is not guaranteed any longer), we provide several new elements that temper this fact:

1. we show that when preferences of agents are *single-peaked*, convergence to a Pareto-optimal allocation can still be guaranteed,
2. under a Borda count interpretation of preferences, we show that while the worst-case loss of utilitarian welfare —*i.e.* the Price of Anarchy— is 2, it is as good as it can be under the assumption of individual rationality (in particular, *top-trading cycle* does not perform better in this respect),
3. we provide a number of experimental results, under different preferences cultures, showing that such dynamics often provide good outcomes, especially in light of their simplicity, and
4. we prove the NP-hardness of deciding whether an allocation maximizing utilitarian or egalitarian welfare is reachable from a given initial allocation.

Acknowledgments. This work is partially supported by the ANR project 14-CE24-0007-01 - CoCoRICo-CoDec.

A Continuum of Stable Matching Relaxations

Brian Dean
School of Computing,
Clemson University
bcdean@cs.clemson.edu

Chris Gropp
School of Computing,
Clemson University
cgropp@clemson.edu

Rommel Jalasutram
School of Computing,
Clemson University
rjalasu@clemson.edu

Abstract

In this note we study properties of a general parametrized form of fractional relaxation that naturally captures the weighted bipartite matching, optimal stable matching, and optimal "uniform" stable allocation problems, all by varying the norm used in a single constraint. Among our results, we show equivalence, in terms of polynomial-time approximability, of the optimal stable matching and optimal uniform stable allocation problems with ties and incomplete preference lists, giving a richer understanding of recently-proposed methods by Huang and Kavitha [1] and Radnai [2], and broadening the design space for future work on these problems. We also discuss potential heuristic applications.

References

[1] C.-C. Huang and T. Kavitha. An improved approximation algorithm for the stable marriage problem with one-sided ties. In *IPCO*, pages 297–308, 2014.

[2] A. Radnai. Approximation algorithms for the stable matching problem. Master's thesis, Eötvös Loránd University, 2014.

Preference Elicitation, Approximate Stability, and Interview Minimization in Stable Matchings

Joanna Drummond and Craig Boutilier
Department of Computer Science
University of Toronto
{jdrummond,cebly}@cs.toronto.edu

Algorithms for stable marriage and related matching problems typically assume that full preference information is available. While the Gale-Shapley algorithm can be viewed as a means of eliciting preferences incrementally, it does not prescribe a general means for matching with incomplete information, nor is it designed to minimize elicitation. Furthermore, little work has investigated schemes for effectively eliciting agent preferences using either preference (e.g., comparison) queries or interviews (to form such comparisons); and no work has addressed how to combine both.

We describe the use of *maximum regret* to measure the (inverse) degree of stability of a matching with partial preferences; *minimax regret* to find matchings that are maximally stable given partial preferences; *minimax regret* to find matchings that are maximally stable in the presence of partial preferences; and heuristic elicitation schemes that use max regret to determine relevant preference queries. We show that several of our schemes find stable matchings while eliciting considerably less preference information than Gale-Shapley.

We also develop a new model for representing and assessing agent preferences that accommodates both eliciting known preference information and (heuristically) minimizing the number of queries and interviews required to determine a stable matching. Our *Refine-then-Interview (RtI)* scheme uses coarse preference queries to refine knowledge of agent preferences and relies on interviews only to assess comparisons of relatively "close" options. Empirical results show that RtI compares favorably to a recent pure interview minimization algorithm, and that the number of interviews it requires is generally independent of the size of the market.

Acknowledgments. We acknowledge the support of NSERC. Drummond was supported by OGS and a Microsoft Research Graduate Women's Scholarship. Thanks to Ettore Damiano for helpful discussions and the reviewers for their suggestions.

Linear and Integer Programming formulations for Stable Allocations and House Allocations*

P. Eirinakis[1], D. Magos[2], I. Mourtos[1,]
[1]Department of Management Science & Technology,
Athens University of Economics and Business, 76 Patission Ave., 104 34 Athens, Greece
email: {peir,mourtos}@aueb.gr
[2]Department of Informatics, Technological Educational Institute of Athens
Ag. Spyridonos Str., 12210 Egaleo, Greece
email: dmagos@teiath.gr

For several variants of matching under preferences, it is customary to reveal some combinatorial properties of the set of solutions before obtaining algorithms finding a feasible or an optimal solution. The approach of examining the problem structure before embarking on algorithmics is common also in the polyhedral combinatorics literature, i.e., the literature that examines a problem via studying the convex hull of vectors that represent its feasible solutions. This approach has been applied to the fundamental variants of two-sided matchings under preferences, namely Stable Marriage [4], Stable Admissions [1] and Stable b-matchings [2].

Here, we proceed in the same direction by introducing the first, to the best of our knowledge, formulations of Stable Allocations and House Allocations (see definitions in [3]). The former problem is a generalization of all two-sided stable matching problems, while the latter is the simplest matching problem with one-sided preferences. Specifically, we provide a linear program for extreme Stable Allocations and establish the minimality of that formulation by employing known polyhedral results on partially ordered sets. Then, we provide an integer program for House Allocations, which has an exponential number of constraints hence being accompanied by a polytime separation algorithm. Last, we discuss outstanding issues regarding the polyhedral structure of these problems.

References

[1] M. Baïou, M.L. Balinski, The stable admissions polytope, Math Program 87 (2000), 427-439.

[2] T. Fleiner, On the stable b-matching polytope, Math Social Sci 46 (2003), 149-158.

[3] D. Manlove, Algorithmics Of Matching Under Preferences, World Scientific Publishing (2013).

[4] A.E. Roth, U.G. Rothblum, J.H. Vande Vate, Stable matchings, optimal assignments, and linear programming. Math Oper Res 18 (1993), 803-828.

*This research has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARCHIMEDES III. Investing in knowledge society through the European Social Fund.

DYNAMIC VCG MECHANISMS IN QUEUEING

SAMBUDDHA GHOSH, YAN LONG, AND MANIPUSHPAK MITRA

ABSTRACT. In a dynamic queueing problem, agents arrive at discrete times to use a rival resource for one period each, and exit permanently thereafter. Each agent privately knows his own per-period waiting cost, and does not observe any other information. The mechanism designer knows neither costs nor future arrivals, and can charge agents present in the system.

We identify the complete class of outcome-efficient and dynamically strategy-proof mechanisms for queueing that use only the reported waiting costs of past and current cohorts to determine an agent's transfer. Finally, from within this class we characterise a canonical one that also achieves dynamic budget balance under equal treatment of equals and a weak constraint on the sequence of arrivals.

GHOSH: DEPARTMENT OF ECONOMICS, BOSTON UNIVERSITY, 270 BAY STATE ROAD, BOSTON, MA 02215, USA, *email*: SGHOSH@BU.EDU;

LONG: ADAM SMITH BUSINESS SCHOOL, UNIVERSITY OF GLASGOW, GLASGOW G12 8QQ, UK, *email*: Y.LONG.1@RESEARCH.GLA.AC.UK;

MITRA: ECONOMIC RESEARCH UNIT, INDIAN STATISTICAL INSTITUTE, 203, B.T. ROAD, KOLKATA-700108, INDIA, *email*: MMITRA@ISICAL.AC.IN.

Redesigning the Entry-Level of the Italian Academic Job Market : a Two-Sided Matching with the Agency Problem

Eliana Quintilii
PhD Student - DREAMT
Dep. of Economics and Management
University of Pavia
Via S. Felice 5, Pavia (Italy)
eliana.quintilii@unipv.it

Abstract

Since its first application in 1984 (Roth 1984), the practice of market design has been gaining a certain relevance in the field of labor markets, especially at their entry-level. The purpose of this paper is to use what learned from the literature for presenting a proposal of redesign of the entry-level of the Academic Job Market in Italy. The two-sided matching researchers/Universities, currently obtained through local contests, suffer different inefficiencies firstly due to its decentralized structure. I deeply analyzed the selection structure and process in order to highlight problems as thickness, congestion and safety, to officially categorize the market as failing (Roth 2007), and meritocracy (Perotti 2008; Perotti et al. 2009) in addition. The first step of the study was to create a unique centralized procedure of matching based on a researchers-proposing deferred-acceptance algorithm that cleverly solve the first three issues. On the other hand, the question of merit – established as the possibility that non-deserving candidates will be hired instead of more qualified ones - represents a new challenge in the practice of market design. It has been treated as a misrepresentation of the Universities' preferences due to an agency problem between the institutions – that express propensities in regulations – and committees – who practically have to realize the list of candidates using their judgments in line with the institutions' guidelines – into the decision process. A misalignment of interests between the two agents causes the false declaration of Universities' preferences and the outcomes of the matching procedure result to be not stable (Roth and Sotomayor 1992; Roth 2008). I formalize the agency problem in the study of the Universities' decision process in order to analyze the players' behaviors and relations for working out a set of rules to control it. Focusing on how agents build up their preferences' lists, a new categorization of two-sided matching markets is offered in order to classify the kind of markets where the agency problem could be noticed and, mostly, can cause the failure of the matching system.

Keywords: *matching, market design, two-sided, preferences, academic job market, researchers, decision process, agency problem*

Estimation of two-sided choice models: an application to public school choice

Matthew Weldon and Andrew Titman

Department of Mathematics and Statistics, Lancaster University, Lancaster, UK

Email address for correspondence: `m.weldon@lancaster.ac.uk`

Consider the problem of studying the choices that people make over which school to send their children to, or which university to attend, or which employment to take. In these contexts it is often difficult to find useful data about agents' preferences to inform research. The classic revealed preference approach would be to use the observed choices themselves as revealing agents' preferences . However, in the settings described above, choice is usually constrained by scarcity of places at institutions, and the fact that the institutions themselves have preferences (or act as if they have preferences) over whom to admit. These institutional preferences may themselves be of substantive interest.

The problem of empirically modelling preferences in two-sided matching markets has received increasing attention recently. It has been shown that, under an assumption of stability, parameters of underlying random utility models for both agents and institutions are identifiable (cf. $[1, 2, 3, 4]$). So far, however, methods to estimate models for unaggregated data have suffered from computational intractability or partial identification that has restricted their use in real-world applications.

This paper introduces a flexible, tractable partial-likelihood for estimating the parameters of a two-sided random utility model, which uses the information within a stable many-to-one matching, such as a school choice setting. This stability–likelihood can be incorporated into maximum–likelihood or Bayesian approaches, and a number of methods are possible. In this paper we focus on a computationally convenient maximum–likelihood method and show that it can be fitted quickly even for large matchings.

We study a context in which the data at our disposal is a many-to-one matching $\mathbf{m}$ between agents and institutions, that we can assume is stable, and a set of attributes of the agents and institutions. We wish to estimate θ, the model parameters that govern the relationship between observable attributes and latent utilities. Given $\mathbf{m}$, a likelihood of the form $\mathrm{P}(\mathbf{m}|\theta)$ is intractable, as it depends on the unknown matching mechanism used. In the absence of a generative model for the matching itself, we base inference on a model for the stability of the observed matching: $\mathrm{P}(\mathbf{m} \in \mathcal{M}^*|\mathbf{m}, \theta)$. We present monte carlo simulation results for identifiability and consistency, and discuss the theoretical properties of this class of model.

Finally, we present the results of a pilot study, estimating parental preferences for school attributes in a UK public school setting. In the pilot study, conducted using the admissions data for seven schools and 838 pupils in a single school district, we present evidence that parents attach high importance to proximity to the school. There is also some tentative evidence of ethnic in-group preferences.

References

[1] N. Agarwal and W. Diamond. Identification and Estimation in two-sided matching markets. Cowles Foundation discussion paper no. 1905. Yale University, 2013.

[2] F. Echenique, S. Lee and M. Shum. Partial Identification in two-sided matching models. In *Advances in Econometrics, Volume 31*, pp.117–139, 2013.

[3] J. Logan, P. Hoff and M. Newton. Two-sided estimation of mate-preferences for similarities in age, education and religion. *Journal of the American Statistical Association,* **103**(482) pp.559–569, 2008.

[4] K. Menzel. Large matching markets as two-sided demand systems. Technical report, New York University, 2014.

Author Index