Book Introduction by the Authors

Invited by

Kazuo Iwama

iwama@kuis.kyoto-u.ac.jp
Bulletin Editor
Kyoto University, Japan

ALGORITHMICS OF MATCHING UNDER PREFERENCES*

David F. Manlove
School of Computing Science,
University of Glasgow,
Glasgow G12 8QQ,
UK
david.manlove@glasgow.ac.uk

Introduction

Matching problems involving preferences occur in widespread applications such as the assignment of children to schools, school-leavers to universities, junior doctors to hospitals, students to campus housing, kidney transplant patients to donors and so on. Very often the common thread is that agents have *ordinal preferences* over the possible outcomes – that is, some notion of first, second, third choice, etc. The task is to find a *matching* (i.e., an assignment of the participants to one another) that is in some sense optimal with respect to these preferences.

These problems are growing in importance in an era in which more and more elements of society are embracing diverse forms of electronic communication, and individuals are increasingly used to making choices via the internet. The ease by which preference information can now be collected has contributed to the growing tendency for matching processes to be centralised. Due to the typical size of applications (for example, in China, over 10 million students apply for admission to higher education annually through a centralised process), trying to construct optimal allocations manually (given a suitable definition of "optimal") is simply not feasible.

Thus algorithms are required to automate the process of constructing optimal matchings. Again, due to the size of typical applications, the efficiency of the algorithms is of paramount importance. The notion of optimality is also a key

*World Scientific, 2013, 524pp, ISBN 978-981-4425-24-7, www.optimalmatching.com/AMUP.

consideration: many matching processes are conducted by publicly-funded organisations, and there is an increasing tendency for the decisions reached by these organisations to be scrutinised both in the media and by individuals through Freedom of Information requests, for example. Thus the algorithms need to construct matchings that are not just provably optimal, but also are seen to be "fair" by the agents involved.

This book focuses on algorithmic aspects of matching problems involving preferences — the aim is to describe polynomial-time algorithms that produce optimal matchings (under many different notions of optimality) or to highlight complexity results that imply the non-existence of such algorithms. Some of the many applications in which these algorithms feature are also described.

Prior work

The archetypal matching problem involving preferences is the celebrated Stable Marriage problem, first introduced by David Gale and Lloyd Shapley in 1962 [1]. The main contribution of their paper was an algorithm, known as the Gale–Shapley algorithm, to solve this problem. This algorithm has been put to practical use in a wide-range of large-scale applications in countries throughout the world.

Three research monographs have since focused on the Stable Marriage problem and its variants [4, 2, 5]. In particular, Gusfield and Irving's book [2] is the standard reference in the literature for structural and algorithmic aspects of the Stable Marriage problem, and indeed its non-bipartite generalisation, the Stable Roommates problem. Whilst in a sense *Algorithmics of Matching Under Preferences* could be regarded as a "sequel" to Gusfield and Irving's monograph [2], bringing the community up to date with the latest developments in this area since their book was published, it also broadens the range of matching problems that they considered and additionally includes alternative optimality criteria besides stability.

The importance of the study of underlying computational matching problems that arise in matching markets was recognised by the award, in 2012, of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel (commonly known as the Nobel Prize in Economic Sciences) to Alvin Roth and Lloyd Shapley, who are both leading figures in the research area.

Problem classification

The matching problems that are considered in this book can be fairly comprehensively classified as follows:

(1) *Bipartite matching problems with two-sided preferences*. Here the participating agents can be partitioned into two disjoint sets, and each member of

one set ranks a subset of the members of the other set in order of preference. Example applications include assigning junior doctors to hospitals, pupils to schools and school-leavers to universities.

(2) *Bipartite matching problems with one-sided preferences*. Again the participating agents can be partitioned into two disjoint sets, but this time each member of only one set ranks a subset of the members of the other set in order of preference. Example applications include campus housing allocation, DVD rental markets and assigning reviewers to conference papers.

(3) *Non-bipartite matching problems with preferences*. Here the participating agents form a single homogeneous set, and each agent ranks a subset of the others in order of preference. Example applications include finding kidney exchanges involving incompatible patient–donor pairs, creating partnerships in P2P networks and forming pairs of agents for chess tournaments.

In the following sections we describe informally some of the key matching problems that belong to each part of the above classification.

Class (1): Bipartite matching problems with two-sided preferences

The classical *Stable Marriage problem* [1, 2] is the central matching problem in this class. An instance of this problem comprises a set of men and women, and each person ranks each member of the opposite sex in strict order of preference.

A many–one generalisation of SM is the *Hospitals / Residents problem* [1, 2], where each man corresponds to a resident and each woman corresponds to a hospital which potentially can be assigned multiple residents up to some fixed capacity. HR models the assignment of junior doctors to hospitals and many other related applications.

Other generalisations of HR that belong to this class are the *Workers / Firms problem* (WF) and the *Student–Project Allocation problem* (SPA).

In each of the problems in this class, the task is to find a *stable matching*. Informally, a *matching* is a set pairs, each of which represents the assignment of an agent from one set to an agent from the other, such that no agent is assigned more agents than its capacity. A matching is *stable* if no two agents prefer one another to one of their current assignees. Were such a pair of agents to exist, they could undermine the matching by forming a private arrangement outside of it.

Roth and his co-authors stressed the importance of stability as a solution concept for matching problems in this class. Most of our treatment of bipartite matching problems with two-sided preferences involves stability as the solution concept, but we also consider alternative optimality criteria in such settings.

Class (2): Bipartite matching problems with one-sided preferences

The *House Allocation problem* (HA) is the variant of SM in which the women do not have preference lists over the men. The men are now referred to as *applicants* and the women are referred to as *houses*. The problem name stems from the application where students are assigned to campus housing, based on their preferences over the available accommodation. This is accomplished using a centralised matching scheme in a number of universities in the US, for example.

A many–one extension of HA, called the *Capacitated House Allocation problem* (CHA) arises when each house can accommodate multiple applicants up to some fixed capacity. CHA can also be regarded as the variant of HR in which hospitals do not have preference lists over residents.

In the context of HA and CHA, only applicants have preferences over houses, so the notion of stability is not relevant. Other optimality criteria have been formulated in the literature, including *Pareto optimality*, *popularity* and *profile-based optimality*. Informally, a matching is *Pareto optimal* if there is no other matching in which some applicant is better off, whilst no applicant is worse off. A matching is *popular* if there is no other matching that is preferred by the majority of the applicants. Finally, the *profile* of a matching M is a vector indicating the number of applicants with their first, second and third choice, etc., in M. Optimising the profile of M might, for example, involve maximising the number of applicants with their first choice, and subject to this, maximising the number with their second choice, etc.

Class (3): Non-bipartite matching problems with preferences

The *Stable Roommates problem* (SR) [1, 3, 2] is the non-bipartite generalisation of SM in which each agent ranks all of the others in strict order of preference. Stability is once again relevant in this context, and the definition of a stable matching is a straightforward extension of the definition in the SM case.

Many–many generalisations SR have been considered in the literature, such as the *Stable Fixtures problem*, the *Stable Multiple Activities problem* and the *Stable Allocation problem*. Variants of SR have also been considered in which agents can form partnerships into sets of size > 2 — this is the *Coalition Formation Game*.

Most of our analysis of non-bipartite matching problems with preferences involves stability as the solution concept, but there are also occasions when we consider optimality criteria other than stability in this context.

Contribution of the book

Following the introductory chapter which gives background, definitions and describes motivating applications, the book is divided into two main parts. Part 1, spanning Chapters 2-5, deals with stable matching problems in Classes (1) and (3). Part 2, spanning Chapters 6-8, focuses on other forms of optimality criteria mainly applied to matching problems in Class (2), but also in instances of problems in Classes (1) and (3).

In Part 1, Chapter 2 deals with the central stable matching problem, namely SM, presenting key developments that have appeared in the literature following the publication of Gusfield and Irving's monograph. We provide updates to lists of open problems from Refs. [4, 2], review two important papers by Subramanian and Feder and describe linear and constraint programming approaches to SM, decentralised algorithms for SM and some beautiful results concerning *median* stable matchings. Among the many other results presented, we show how stable matching theory led to a very elegant solution to the Dinitz conjecture.

The extensions of SM and HR in which preference lists can include ties and other forms of indifference led to a substantial revival in the study of stable matching problems in the late 1990s and early 2000s. In Chapter 3 we describe algorithmic results for problems involving computing stable matchings in the presence of indifference. One particular problem, namely that of finding a stable matching that matches as many people as possible, given an instance of SM where the preference lists may involve ties and may be incomplete, has led to an interesting "race" to find the tightest, fastest and simplest approximation algorithm.

SR, the non-bipartite version of SM, has traditionally been studied less extensively than SM. However following the publication of [2], some important structural and algorithmic results due to Tan were published, guaranteeing the existence of a so-called *stable partition*, a generalisation of a stable matching, even in instances of SR that admit no stable matching. We describe Tan's results, and many other more recent results for SR and its variants, in Chapter 4.

Further results for stable matching problems are presented in Chapter 5. We describe extensions of HR in which hospitals can have lower and/or common quotas, which present additional constraints on the numbers of assignees that they can/must obtain in a stable matching. We also consider the variant of HR in which couples can provide joint preference lists in order to be matched to hospitals that are geographically close to one another. Other problems considered include SPA, WF and three-dimensional stable matching problems.

Part 2 is concerned with optimality criteria that can be defined for matching problems in Class (2). These include Pareto optimality, popularity and profile-based optimality. These criteria are mainly applied to HA, CHA and their variants, but are also studied in the context of SM, HR and SR. Issues of interest in each case include the existence of an optimal matching, the algorithmic complexity of finding an optimal matching, and the structure of the set of optimal matchings in a given problem instance. Generally speaking, Pareto-optimal and profile-based optimal matchings are bound to exist, but there is no such guarantee in the case of popular matchings. Results for Pareto optimal, popular and profile-based optimal matchings are given in Chapters 6, 7 and 8 respectively.

One of the purposes of this book is to stimulate further research in the area of matching under preferences, and to this end we identify a range of open problems for future investigation. These are presented in the concluding section of each chapter.

Final remarks

To summarise, this book is largely a comprehensive, classified and guided survey through the literature on matching problems with preferences from an algorithmic perspective. The intended readership includes PhD students, postdoctoral researchers and academic staff engaged in research on matching under preferences, senior undergraduate and taught postgraduate students engaged in project work relating to matching under preferences or taking an advanced course on matching theory, and indeed administrators of centralised matching schemes who are interested in the algorithms that underpin these programmes. It contains a Foreword by Kurt Mehlhorn, who wrote:

> “This book covers the research area in its full breadth and beauty. Written by one of the foremost experts in the area, it is a timely update to âĂIJThe Stable Marriage Problem: Structure and AlgorithmsâĂİ (D. Gusfield and R.W. Irving, 1989). This book will be required reading for anybody working on the subject; it has a good chance of becoming a classic.”

References

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

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

[3] R.W. Irving. An efficient algorithm for the “stable roommates” problem. *Journal of Algorithms*, 6:577–595, 1985.

[4] D.E. Knuth. *Mariages Stables*. Les Presses de L’Université de Montréal, 1976. English translation in *Stable Marriage and its Relation to Other Combinatorial Problems*, volume 10 of CRM Proceedings and Lecture Notes, American Mathematical Society, 1997.

[5] A.E. Roth and M.A.O. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, volume 18 of *Econometric Society Monographs*. Cambridge University Press, 1990.