Discrete Optimization 29 (2018) 174–195

Contents lists available at ScienceDirect

Discrete Optimization

www.elsevier.com/locate/disopt

Pareto optimal matchings of students to courses in the presence of prerequisites[☆]

Katarína Cechlárová[a], Bettina Klaus[b], David F. Manlove[c,*]

[a] Institute of Mathematics, Faculty of Science, P.J. Šafárik University, Jesenná 5, 040 01 Košice, Slovakia
[b] Faculty of Business and Economics, University of Lausanne, Internef 538, CH-1015 Lausanne, Switzerland
[c] School of Computing Science, Sir Alwyn Williams Building, University of Glasgow, Glasgow, G12 8QQ, UK

ARTICLE INFO

Article history:
Received 8 July 2017
Received in revised form 13 February 2018
Accepted 22 April 2018
Available online 27 July 2018

Keywords:
Course allocation
Prerequisites
Pareto optimality
Sequential mechanism
NP-hardness

ABSTRACT

We consider the problem of allocating applicants to courses, where each applicant has a subset of acceptable courses that she ranks in strict order of preference. Each applicant and course has a *capacity*, indicating the maximum number of courses and applicants they can be assigned to, respectively. We thus essentially have a many-to-many bipartite matching problem with one-sided preferences, which has applications to the assignment of students to optional courses at a university.

We consider additive preferences and lexicographic preferences as two means of extending preferences over individual courses to preferences over bundles of courses. We additionally focus on the case that courses have prerequisite constraints: we will mainly treat these constraints as compulsory, but we also allow alternative prerequisites. We further study the case where courses may be corequisites.

For these extensions to the basic problem, we present the following algorithmic results, which are mainly concerned with the computation of Pareto optimal matchings (POMs). Firstly, we consider compulsory prerequisites. For additive preferences, we show that the problem of finding a POM is NP-hard. On the other hand, in the case of lexicographic preferences we give a polynomial-time algorithm for finding a POM, based on the well-known sequential mechanism. However we show that the problem of deciding whether a given matching is Pareto optimal is co-NP-complete. We further prove that finding a maximum cardinality (Pareto optimal) matching is NP-hard. Under alternative prerequisites, we show that finding a POM is NP-hard for either additive or lexicographic preferences. Finally we consider corequisites. We prove that, as in the case of compulsory prerequisites, finding a POM is NP-hard for additive preferences, though solvable in polynomial

☆ Supported by grant APVV-15-0091 from the Slovak Research and Development agency (Cechlárová), by grant 100018_150086 from the Swiss National Science Foundation SNFS (Klaus), by grants EP/P028306/1 and EP/K010042/1 from the Engineering and Physical Sciences Research Council (Manlove), and by a Short-Term Scientific Mission from COST Action IC1205 on Computational Social Choice (Manlove). Part of this work was carried out whilst the authors visited Corvinus University of Budapest, and whilst the third author visited P.J. Šafárik University. This research was initiated during Dagstuhl Seminar 15241 on Computational Social Choice: Theory and Applications [1]. We would like to thank the following Dagstuhl participants who were involved in initial discussions regarding the results in this paper: Haris Aziz, Péter Biró, Jiehua Chen and Nicholas Mattei.

* Corresponding author.
E-mail addresses: katarina.cechlarova@upjs.sk (K. Cechlárová), Bettina.Klaus@unil.ch (B. Klaus), david.manlove@glasgow.ac.uk (D.F. Manlove).

https://doi.org/10.1016/j.disopt.2018.04.004
1572-5286/

time for lexicographic preferences. In the latter case, the problem of finding a maximum cardinality POM is NP-hard and very difficult to approximate.

1. Introduction

Problems involving the allocation of indivisible goods to agents have gained a lot of attention in the literature, since they model many real scenarios, including the allocation of pupils to study places [2], workers to positions [3], researchers to projects [4], tenants to houses [5] and students to courses [6], etc. We assume that agents on one side of the market (pupils, workers, researchers, tenants, students) have preferences over objects on the other side of the market (study places, positions, projects, courses, etc.) but not vice versa. In such a setting where the choices of agents are in general conflicting, economists regard Pareto optimality (or Pareto efficiency) as a basic, fundamental criterion to be satisfied by an allocation. This concept guarantees that no agent can be made better off without another agent becoming worse off. A popular and very intuitive approach to finding Pareto optimal matchings is represented by the class of *sequential allocation mechanisms* [7–10].

In the one-to-one case (each agent receives at most one object, and each object can be assigned to at most one agent) this mechanism has been given several different names in the literature, including serial dictatorship [5,11], queue allocation [12], Greedy-POM [13] and sequential mechanism [9,10], etc. Several authors independently proved that a matching is Pareto optimal if and only if it can be obtained by the serial dictatorship mechanism (Svensson in 1994 [12], Abdulkadiroğlu and Sönmez in 1998 [5], Abraham et al. in 2004 [13], and Brams and King in 2005 [8]).

In general many-to-many matching problems (agents can receive more than one object, and objects can be assigned to more than one agent), the sequential allocation mechanism works as follows: a central authority decides on an ordering of agents (often called a *policy*) that can contain multiple copies of an agent (up to her capacity). According to the selected policy, an agent who has her turn chooses her most preferred object among those that still have a free slot. This approach was used in [9,10], where the properties of the obtained allocation with respect to the chosen policy and strategic issues are studied.

The serial dictatorship mechanism is a special case of the sequential allocation mechanism where the policy contains each agent exactly once, and when agent a is dealt with, she chooses her entire most-preferred bundle of objects. The difficulty with serial dictatorship is that it can output a matching that is highly unfair. For example, it is easy to see that if there are two agents, each of whom finds acceptable all objects and has capacity equal to the number of objects, and each object has capacity 1, then the serial dictatorship mechanism will assign all objects to the first agent specified by the policy and no object to the other agent.

In this paper we shall concentrate on one real-life application of this allocation problem that arises in education, and so our terminology will involve *applicants* (students) for the agents and *courses* for the objects. In most universities students have some freedom in their choice of courses, but at the same time they are bound by the rules of the particular university. A detailed description of the rules of the allocation process and the analysis of the behaviour of students at Harvard Business School, based on real data, is provided by Budish and Cantillon [6]. They assume that students have a linear ordering of individual courses and their preferences over bundles of courses are responsive to these orderings. The emphasis in [6] was on strategic questions. The empirical results confirmed the theoretical findings that, loosely speaking, dictatorships (where students choose one at a time their entire most preferred available bundle) are the only mechanisms that are strategy-proof and ex-post Pareto optimal.

Another field experiment in course allocation is described by Diebold et al. [14]. The authors compared the properties of allocations obtained by the sequential allocation mechanism where the policy is determined