P. J. Šafárik University
Faculty of Science
Institute of mathematics
Jesenná 5, 040 01 Košice, Slovakia

**K. Cechlárová, P. Eirinakis, T. Fleiner,
D. Magos, I. Mourtos and E. Potpinková**

Pareto optimality in many-to-many matching problems

IM Preprint, series A, No. 4/2013
December 2013

Pareto optimality in many-to-many matching problems

Katarína Cechlárová[1], Pavlos Eirinakis[2], Tamás Fleiner[3], Dimitrios Magos[4], Ioannis Mourtos[2] and Eva Potpinková[1]

[1]Institute of Mathematics, Faculty of Science, P.J. Šafárik University, Jesenná 5, 040 01 Košice, Slovakia

email: katarina.cechlarova@upjs.sk, eva.potpinkova@student.upjs.sk

[2]Department of Management Science & Technology, Athens University of Economics and Business, 76 Patission Ave., 104 34 Athens, Greece

email:{peir,mourtos}@aueb.gr

[3]Budapest University of Technology and Economics Magyar tudósok körútja 2, H-1117, Budapest, and MTA-ELTE Egerváry Research Group, Hungary

email: fleiner@cs.bme.hu

[4]Department of Informatics, Technological Educational Institute of Athens, Ag. Spyridonos Str., 12210 Egaleo, Greece

email:dmagos@teiath.gr

Abstract. Consider a many-to-many matching market that involves two finite disjoint sets, a set of applicants A and a set of courses C. Each applicant has preferences on the different sets of courses she can attend, while each course has a quota of applicants that it can admit. In this paper, we examine Pareto optimal matchings (briefly POM) in the context of such markets, that can also incorporate additional constraints, e.g., each course bearing some cost and each applicant having an available budget. We provide necessary and sufficient conditions for a many-to-many matching to be Pareto optimal and show that checking whether a given matching is Pareto optimal requires $O(|A|^2 \cdot |C|^2)$ time. Moreover, we provide a generalized version of serial dictatorship, which can be used to obtain any many-to-many POM. We also study the problems of finding a minimum cardinality and a maximum cardinality POM. We show that the former is NP-complete even in one-to-one markets with the preference list of each applicant containing at most two entries. For the latter problem we show that, although it is polynomially solvable in the special one-to-one case, it is NP-complete for many-to-many markets.

Keywords: Matching, Pareto optimality, Serial dictatorship, NP-completeness

AMS classification: 91A12, 91A06, 68Q25

1 Introduction

A university runs a leisure centre that offers a variety of activities, e.g., sports, language courses, etc. (we shall call all of them *courses*), to students and employees (we shall call all of them *applicants*). Each applicant can attend one or more courses and, because of various technical constraints, each course can only accept a restricted number of applicants. Furthermore, certain additional rules may apply. For example, each applicant, if accepted, may have to pay some fee as a contribution to cover the running costs of the course. On the other hand, each applicant may have a budget that she is able to allocate to these courses that she cannot exceed (hereafter, we refer to applicants as females). Two simple problems that naturally arise in this many-to-many matching context are those of assigning each applicant to all the courses she desires and of assigning each applicant to at least one course. Both cases reduce to well-known combinatorial optimization problems, namely the maximum flow problem and the maximum cardinality bipartite matching problem respectively, see e.g. [6].

In real life however, it is usually the case that the applicants do not desire equally all the courses they apply for; rather they have certain preferences over them. The problem that arises when taking these preferences into account will be called the *Course Allocation problem* (CAP). In the CAP setting, various optimization criteria for the obtained assignments can be formulated. Here, we shall concentrate on *Pareto optimality.*

Pareto optimality, sometimes called *Pareto efficiency,* is a well established notion in economic science. It is the primary welfare goal in many real matching markets, especially educational markets assigning pupils to schools (see [3, 4] for assigning students to public schools in several US school districts, [7] for college admission in Turkey) or students to campus housing [8], [10]. A detailed account of recent developments regarding Pareto optimality in the context of matching problems under preferences has appeared in [9].

The special case of one-to-one CAP is often called the House Allocation problem, as it arises in the context of assigning tenants to houses [1], [2]. A detailed study of computational aspects of the House Allocation problem was provided in [5]. The authors gave necessary and sufficient conditions for a matching to be Pareto optimal and showed that these conditions can be checked in polynomial time. They also established that any Pareto optimal matching (POM) can be obtained by the well-known *serial dictatorship mechanism* [1] and proposed an efficient algorithm to find a POM of maximum cardinality. Analogous results have been established in [13] for the *many-to-one* (*capacitated*) House Allocation problem, i.e., the variant where each house can accommodate more than one tenant.

Regarding intractability results, it has been established that finding a Pareto optimal one-to-one matching of minimum cardinality is NP-complete even in the one-to-one case [5]. A related recent paper [12] deals with the computational complexity of serial dictatorship. The authors prove that in this mechanism, the problem of deciding whether there exists an order of proposals such that a given agent receives a given object is NP-complete, while the problem asking whether in each order of proposals a given agent receives a given object can be decided in polynomial time.

In this paper, we completely characterize Pareto optimal matchings in the many-to-many setting and show that deciding whether a given matching is Pareto optimal requires polynomial time. Consequently, our work concludes the research of [5, 13] on these issues for two-sided markets as it treats the general case. We also generalize the serial dictatorship mechanism, thus providing a procedure that can be used to obtain any many-to-many POM. This result is important also because, unlike in the one-to-one case, serial dictatorship alone cannot guarantee that all Pareto optimal matchings will be generated [5]. Further, we show that finding a minimum cardinality POM is NP-complete, even when considering the simplified one-to-one case with the preference list of each applicant containing at most two entries, thus strengthening the result given in [5]. Moreover, we prove that the maximum cardinality POM problem is also NP-complete, although it is polynomially solvable in the (capacitated) House Allocation case [5], [9], [13].

2 Definitions

An instance of the Course Allocation problem involves a set A of n applicants and a set C of m courses. Each course $c \in C$ has a quota $q(c)$. A subset $A' \subseteq A$ of applicants is *feasible* for a course c if $|A'| \leq q(c)$. Each applicant a has a preference list $P(a)$, a strictly ordered list of a subset of courses. These courses are *acceptable* for a and we shall write $c \succ_a c'$ if applicant a prefers course c to course c'. Moreover, there is a family F_a of subsets of courses associated with each applicant a. We say that the sets of courses belonging to F_a are *feasible* for a, while all other sets are *infeasible*. For each applicant a, we suppose that F_a is *downward closed*, i.e. if $C'' \subseteq C'$ and $C' \in F_a$, then $C'' \in F_a$ too. Note that the House Allocation problem is obtained if $q(c) = 1$ for each $c \in C$ and all the feasible sets are just singletons, containing the acceptable courses for each applicant.

As discussed in the Introduction, additional rules may apply in this setting. Consider, for instance, the case in which each course has an attendance cost and each applicant may have a budget. The corresponding framework can be obtained as follows. Suppose that each course c has a nonnegative price $p(c)$ and each applicant a has a budget $b(a)$. Let $p(C')$ denote the total price of all courses in the subset C', i.e. $p(C') = \sum_{c \in C'} p(c)$. Then $F_a = \{C' \subseteq P(a); p(C') \leq b(a)\}$. It is easy to see that F_a defined in this way is indeed downward closed. In this paper, we shall call this special case the *price-budget* CAP.

Other more complicated situations also fit in our model. For example, besides the price, each course might also have some time requirements and applicants might be restricted not only in the available budget, but also in time they are able to allocate to the courses. In another case, courses may be of different types (sports, languages, music, etc) and the applicants may wish to take at most one courses of each type. Any such case can be handled by appropriately defining F_a, i.e., the feasible set of courses of any applicant a, as long as F_a remains downward closed.

An *assignment* M is a subset of $A \times C$. The set of applicants assigned to a course c will be denoted by $M(c) = \{a \in A; (a, c) \in M\}$ and similarly, the set of courses assigned to an applicant a is $M(a) = \{c \in C; (a, c) \in M\}$. An assignment M is a *matching* if $M(a)$ is feasible for each applicant a and $M(c)$ is feasible for each course $c \in C$.

An applicant $a \in A$ is *assigned* if $M(a) \neq \emptyset$, otherwise she is *unassigned.* A course $c \in C$ is *open* if $M(c) \neq \emptyset$, otherwise it is *closed.* An applicant a and a course c are *undersubscribed* if $M(a)$ is not an inclusionwise maximal element of F_a and $|M(c)| < q(c)$, respectively. If $|M(c)| = q(c)$, we say that c is *full.*

Applicant a prefers matching M to matching M' if she prefers $M(a)$ to $M'(a)$. We suppose that applicants compare the sets of courses lexicographically. This means that applicant a orders the acceptable courses according to her preference list from the most preferred to the least preferred one and compares the characteristic vectors χ_a of feasible sets with the entries arranged in this order. More precisely, if C' and C'' are two feasible sets of courses, then $C' \succ_a C''$ if $\chi_a(C') >_{lex} \chi_a(C'')$, that is, the most preferred element of symmetric difference $C' \triangle C'' = (C' \setminus C'') \cup (C'' \setminus C')$ belongs to C'. We write $C' \succeq_a C''$ if either $C' \succ_a C''$ or $C' = C''$. Notice that the ordering of sets of courses generated by a strict preference order $P(a)$ is also strict.

Example 1 *Let us consider the price-budget* CAP *instance given in Table 1.*

applicant	preference list	budget	course	price	quota
a_1	c_1, c_2, c_3	2	c_1	2	2
a_2	c_2, c_1	3	c_2	1	1
a_3	c_3, c_1	2	c_3	1	1

Table 1: Price-budget CAP instance for Example 1.

The feasible sets of applicant a_1 in the order of her preference are

$$\{c_1\} \succ_{a_1} \{c_2, c_3\} \succ_{a_1} \{c_2\} \succ_{a_1} \{c_3\} \succ_{a_1} \emptyset,$$

for applicants a_2 and a_3, their orderings are

$$\{c_2, c_1\} \succ_{a_2} \{c_2\} \succ_{a_2} \{c_1\} \succ_{a_2} \emptyset \text{ and } \{c_3\} \succ_{a_3} \{c_1\} \succ_{a_3} \emptyset, \text{ respectively.}$$

Notice that this ordering is compatible with the assumption of the following *greedy* behaviour of applicants: if an applicant can freely pick her favourite assignment from some set of available courses, then she goes down her preference list and adds the next course if and only if the set of courses comprised of those chosen so far plus the new course is still feasible.

We say that a matching M' *dominates* a matching M ($M' \succ M$) if at least one applicant prefers M' to M and no applicant prefers M to M'.

Proposition 1 *The relation '$\succ$' forms a partial order over the set of matchings.*

A Pareto optimal matching is a matching that is not dominated by any other matching. Thanks to Proposition 1 and finiteness of the set of all matchings, a Pareto optimal matching exists for each instance of CAP.

3 Characterization of POM

Abraham et al. [5] characterized Pareto optimal one-to-one matchings as those that are maximal, trade-in-free and coalition-free. We generalize their result below.

Let M be any matching and $c \in P(a)\backslash M(a)$. Let us denote $D_M(a,c) = \{c' \in M(a); c \succ_a c'\}$. Intuitively, $D_M(a,c)$ is the set of courses applicant a is happy to drop from her current assignment $M(a)$ in exchange for getting course c, as, according to the lexicographic ordering, agent a prefers the set $(M(a)\backslash D_M(a,c)) \cup \{c\}$ to $M(a)$.

Let us say that a sequence of applicant-course pairs

$$K = ((a_0, c_0), (a_1, c_1), \ldots, (a_{r-1}, c_{r-1}))$$

is a *coalition* with respect to a matching M if for $i = 0, 1, \ldots, r-1$ we have (here and later, when dealing with coalitions, we always understand the indices modulo r):

a) $c_i \in M(a_i)$, $c_{i+1} \notin M(a_i)$;

b) a_i prefers c_{i+1} to c_i;

c) $(M(a_i)\backslash D_M(a_i, c_{i+1})) \cup \{c_{i+1}\} \in F_{a_i}$.

The matching

$$M' = M/K = (M\backslash\{(a_i, c); c \in D_M(a_i, c_{i+1}); 0 \le i \le r-1\}) \cup \{(a_i, c_{i+1}); 0 \le i \le r-1\}$$

is said to be obtained from M by *satisfying coalition K*.

Definition 1 *Let M be a matching. We say that M is*

(i) maximal, *if there exists no applicant $a \in A$ and course $c \notin M(a)$ such that*

$$M(a) \cup \{c\} \in F_a \text{ and } |M(c)| + 1 \le q(c);$$

(ii) trade-in-free, *if there exists no applicant $a \in A$ and course $c \notin M(a)$ such that*

$$(M(a)\backslash D_M(a,c)) \cup \{c\} \in F_a \qquad \text{and} \qquad |M(c)| + 1 \le q(c);$$

(iii) coalition-free, *if there exists no coalition with respect to M.*

Example 2 *Consider again the price-budget* CAP *instance given in Table 1 and a matching*

$$M_1 = \{(a_1, c_2), (a_1, c_3), (a_2, c_1), (a_3, c_1)\}.$$

As all the courses are full, M_1 is maximal as well as trade-in-free. However, it is not coalition-free, as it admits two coalitions, namely $K_1 = ((a_1, c_2), (a_2, c_1))$ and $K_2 = ((a_1, c_3), (a_3, c_1))$. The matching

$$M_2 = M_1/K_1 = \{(a_1, c_1), (a_2, c_2), (a_3, c_1)\}$$

is not trade-in-free, as applicant a_3 will be happy to drop course c_1 and get course c_3 while this course has a free quota. The obtained matching

$$M_3 = \{(a_1, c_1), (a_2, c_2), (a_3, c_3)\}$$

is not maximal, as the pair (a_2, c_1) can be added and we finally arrive at a POM

$$M_4 = \{(a_1, c_1), (a_2, c_2), (a_2, c_1), (a_3, c_3)\}.$$

Theorem 1 *A matching in an instance of* CAP *is Pareto optimal if and only if it is maximal, trade-in-free and coalition-free.*

Proof. It is easy to see that if M is a POM, then it is maximal, trade-in-free and coalition-free.

To show the opposite direction, assume that a matching M is not Pareto optimal but it is maximal, trade-in-free and coalition-free. Since M is not Pareto optimal, there is some matching M' and an applicant a_0 such that $M'(a_0) \succ_{a_0} M(a_0)$ and $M'(a) \succeq_a M(a)$ for all $a \in A$. Let c_1 be a_0's most preferred course in the symmetric difference $M'(a_0) \Delta M(a_0)$. As, $M'(a_0) \succ_{a_0} M(a_0)$, we have $c_1 \in M'(a_0)$. Further, since M is maximal,

$$(a1)\ |M(c_1)| + 1 > q(c_1) \text{ or } (a2)\ M(a_0) \cup \{c_1\} \notin F_{a_0}.$$

M being trade-in-free implies that

$$(b1)\ |M(c_1)| + 1 > q(c_1) \text{ or } (b2)\ (M(a_0) \cup \{c_1\}) \backslash D_M(a_0, c_1) \notin F_{a_0}.$$

Now distinguish two cases.
Case 1. $|M(c_1)| + 1 \leq q(c_1)$. In this case, both $(a2)$ and $(b2)$ must hold. The assumption that the families of feasible sets are downward closed implies that $(b2)$ is stronger, so it must be that $(M(a_0) \cup \{c_1\}) \backslash D_M(a_0, c_1) \notin F_{a_0}$. However, $(M(a_0) \cup \{c_1\}) \backslash D_M(a_0, c_1) \subset M'(a_0)$, a contradiction.
Case 2. $|M(c_1)| + 1 > q(c_1)$. In this case, since $a_0 \in M'(c_1) \backslash M(c_1)$, there exists $a_1 \in M(c_1) \backslash M'(c_1)$. Moreover, for a_1 we have $M'(a_1) \succ_{a_1} M(a_1)$: the reason is that we have assumed that M' dominates M, hence $M'(a_1) \succeq_{a_1} M(a_1)$, while $c_1 \in M(a_1) \backslash M'(a_1)$ means that $M(a_1) \neq M'(a_1)$.

We can now proceed as above with a_1 in the role of a_0 thus either yielding a contradiction (**Case 1**) or eventually revealing a coalition (since the applicants' set is finite) (**Case 2**), which in turn contradicts our assumption that M is coalition-free. ■

Testing Pareto optimality of a given matching can be performed by testing each of the three properties separately. However, we construct a special digraph, called the *extended envy graph*, that enables testing all the three conditions simultaneously. (Compare [5], where a simpler *envy graph* was used only to test the existence of coalitions.) Moreover, this digraph will be helpful in a generalization of serial dictatorship.

For convenience, let us define, for a matching M, an applicant a and courses $c' \in P(a) \backslash M(a)$ and $c \in M(a)$:

$$F_M(a, c, c') = (M(a) \cup \{c'\}) \backslash \{c'' \in M(a); c \succeq_a c''\}$$

to be the set of courses assigned to a if she acquires course c' and drops from $M(a)$ course c and all the courses she prefers less than c.

Definition 2 *The* extended envy graph $G(M) = (V_{G(M)}, E_{G(M)})$ *associated with a matching* M *is a digraph with* $V_{G(M)} = A \cup C \cup \{ac : (a,c) \in M\}$ *and* $E_{G(M)} = E^1_{G(M)} \cup E^2_{G(M)} \cup E^3_{G(M)}$ *where*

- $E^1_{G(M)} = \bigcup_{a \in A} \{(c,a) : c \in P(a) \backslash M(a), |M(c)| < q(c)\}$,
- $E^2_{G(M)} = \bigcup_{a \in A} \{(a, ac) : c \in M(a)$ *and no* $c' \in M(a) \backslash \{c\}$ *satisfies* $c \succ_a c'\} \bigcup \bigcup_{a \in A} \{(ac, ac')) : \{c, c'\} \subseteq M(a), c' \succ_a c$ *and no* $c'' \in M(a) \backslash \{c, c'\}$ *satisfies* $c' \succ_a c'' \succ_a c\}$,
- $E^3_{G(M)} = \bigcup_{a \in A} \{(a,c) : c \in P(a) \backslash M(a), M(a) \cup \{c\} \in F_a, |M(c)| < q(c)\} \bigcup \bigcup_{a \in A} \{(ac, c') : c' \in P(a) \backslash M(a), c' \succ_a c, F_M(a,c,c') \in F_a, |M(c')| < q(c')\} \bigcup \bigcup_{a \in A} \{(ac, a'c') : c' \in (P(a) \backslash M(a)) \cap M(a'), c' \succ_a c, F_M(a,c,c') \in F_a\}$.

That is, $G(M)$ has one vertex per applicant, one per course and one per pair in M (matching vertices). The arc set $E_{G(M)}$ has three types of arcs, namely the *availability arcs* $E^1_{G(M)}$ that indicate that a course c is available but not matched to applicant a; the *matching arcs* $E^2_{G(M)}$ that form vertex-disjoint paths, each containing $|M(a)| + 1$ vertices in increasing order with respect to $P(a)$ (in this context, vertex a corresponds to applicant a being unassigned); and the *envy arcs* $E^3_{G(M)}$ that establish that applicant a envies some course not contained in $M(a)$. Envy arcs are, in turn, of three types, namely arcs (a,c) showing that a wishes and can (in terms of feasibility) to add c to her courses; arcs (ac, c') indicating that a wishes and can trade-in c'; and arcs $(ac, a'c')$ indicating that a wishes and can acquire c' matched to some other applicant a' by dropping from $M(a)$ the courses that are not better than c. Notice that, under the reasonable assumption that any course is acceptable and individually feasible by some applicant, $c \in V_{G(M)}$ has no incident arcs if and only if $|M(c)| = q(c)$.

Extended envy graphs for matchings from Example 2 are given in Figure 1.

Theorem 2 *M is a POM if and only if $G(M)$ is acyclic.*

Proof. If M is not a POM, Theorem 1 yields that M is not maximal or not trade-in-free or not coalition-free. If M is not maximal, Definition 2 implies that $G(M)$ contains arcs $(c,a) \in E^1_{G(M)}$ and $(a,c) \in E^3_{G(M)}$ for some $a \in A$ and $c \in C$, hence a cycle. If M is not trade-in-free, again Definition 2 yields that there is an applicant a and courses c, c' such that $c' \succ_a c$ and $G(M)$ contains arc $(c', a) \in E^1_{G(M)}$, a path from a to ac, arcs in $E^2_{G(M)}$ and the arc $(ac, c') \in E^3_{G(M)}$, hence a cycle. If there is a coalition $K = ((a_0, c_0), \dots, (a_{r-1}, c_{r-1}))$ in M, Definition 2 implies that arcs $(a_i c_i, a_{i+1} c_{i+1})\,, 0 \leq i \leq r-1$ are in $E^3_{G(M)}$, hence $G(M)$ contains a cycle.

To prove the converse, assume that $G(M)$ contains a cycle $\mathfrak{C}$.

Case 1. $\mathfrak{C}$ contains an availability arc $(c,a) \in E^1_{G(M)}$.

Case 1a. $(a,c) \in E^3_{G(M)}$ yields M is not maximal, since by Definition 2, $c \in P(a) \backslash M(a)$, $|M(c)| < q(c)$ and $M(a) \cup \{c\} \in F_a$.

Case 1b. If $(a', c) \in E^3_{G(M)}$ for some $a' \neq a$, then Definition 2 implies that $(c, a') \in E^1_{G(M)}$ too. So this case reduces to Case 1a.

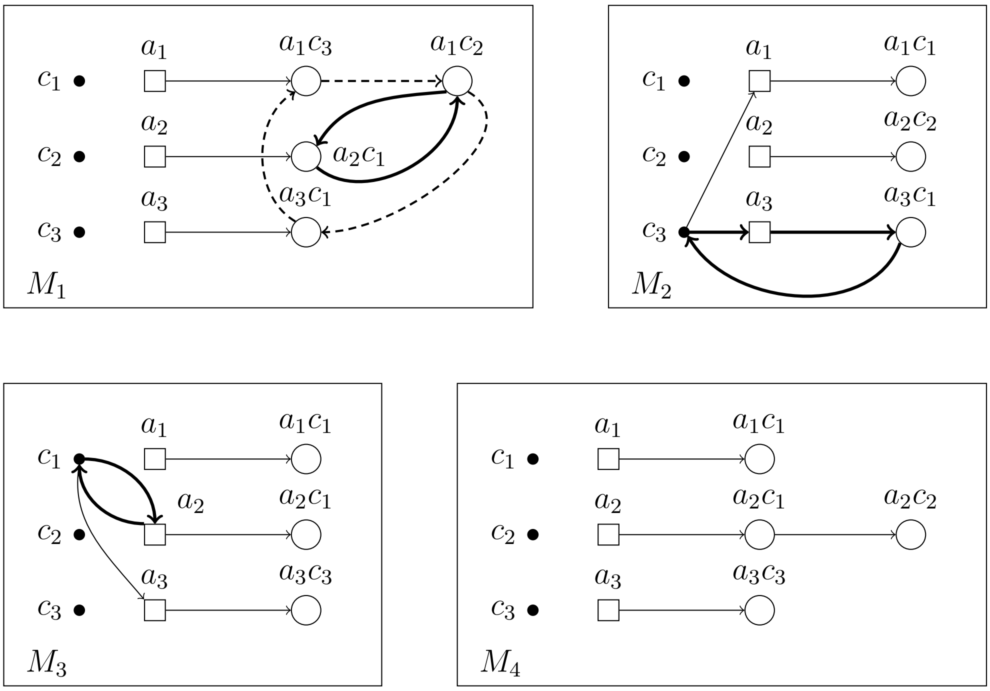

Figure 1: Extended envy graphs for matchings M_1, M_2, M_3, M_4 of Example 2

Case 1c. $E^3_{G(M)}$ contains no arc of the form (a', c). Then the only incoming arc to vertex c can be an envy arc from some vertex $a'c'$ such that $F_M(a', c', c) \in F_{a'}$ (also notice that, in that case, $(c, a') \in E^1_{G(M)}$). Thus applicant a' would trade-in c, i.e., M is not trade-in-free.

Case 2. $\mathfrak{C}$ contains no availability arc. Then it contains no vertices in $A \cup C$. Hence, Definition 2 yields that $\mathfrak{C}$ can only be a sequence of sub-paths, each containing some (possibly an empty set of) matching arcs plus a single envy arc of type 3. Formally, $\mathfrak{C}$ comprises sub-paths P_i, $i \in \{0, \dots, r-1\}$, where each P_i starts at vertex a_ic_i, then proceeds using only matching arcs to vertex $a_ic'_i$ and ends with an envy arc $(a_ic'_i, a_{i+1}c_{i+1})$; where by definition $c_{i+1} \succ_{a_i} c'_i$ (recall that indices are taken modulo r, hence $a_{i+1} = a_0$ for $i = r-1$). If applicants a_i are pairwise different, observe that $((a_0, c_0), (a_1, c_1), \dots, (a_{r-1}, c_{r-1}))$ is a coalition. Otherwise, there exists $j \neq i$ such that $a_i = a_j$. Let us denote $a_i = a_j$ by a and suppose w.l.o.g. that $c_j \succ_a c_i$. As the indices on $\mathfrak{C}$ are taken modulo r, we can suppose $i < j$. Then we replace the part of $\mathfrak{C}$ between the vertices a_ic_i and $a_jc'_j$ simply by the path consisting of the matching arcs leading from a_ic_i to $a_jc'_j$ (remember that a_i and a_j are both equal to a and $c'_j \succ_a c_j \succ_a c_i$, so this path exists). If necessary, a similar shortcut can be applied several times, eventually obtaining a cycle with all applicants mutually different, hence giving a coalition. ■

Corollary 1 *Checking Pareto optimality of a matching can be performed in $O(|A|^2 \cdot |C|^2)$ steps.*

Proof. Follows from Theorem 2 and the fact that a cycle can be found in $O(|V_{G(M)}| + |E_{G(M)}|)$ steps; $|V_{G(M)}|$ is $O(|A| \cdot |C|)$ and $|E_{G(M)}|$ is $O(|A|^2 \cdot |C|^2)$,

applicant	preference list	budget	course	price	quota
a_1	c_1, c_2	2	c_1	1	2
a_2	c_3, c_4	2	c_2	1	2
a_3	c_3, c_1	1	c_3	1	1
a_4	c_2, c_4	1	c_4	1	1
a_5	c_1, c_4	1			

Table 2: Price-budget CAP instance for Example 3.

i.e., $|E_{G(M)}|$ is determined by the fact that $O(|A| \cdot |C|)$ vertices of the 'ac' type have $O(|A| \cdot |C|)$ incident envy arcs. ■

Let us remark here that the generalized envy graph could also be used for finding a POM. Simply start by any matching M, construct $G(M)$ and if $G(M)$ is not acyclic, improve it according to any existing cycle. For the new matching construct again its generalized envy graph etc., until an acyclic graph is obtained. It is easy to see that this procedure will eventually lead to a POM, but it is difficult to derive a bound on the number of matchings constructed. In the next section, we propose a much faster approach.

4 Pareto optimality and serial dictatorship

Recall that serial dictatorship (henceforth abbreviated by SD) is the following algorithm: applicants are considered in a certain order. Each applicant on her turn chooses the most preferred set of courses among those that are still available. Serial dictatorship or its variants are used in real labour or educational markets, see Example 4.3 of [11] or [7].

It is easy to see that the following assertion is true.

Proposition 2 *Each matching obtained by SD is Pareto optimal.*

Abraham et al. [5] showed that in the one-to-one case the converse is also true: each Pareto optimal matching can be obtained by SD in a suitable order. We show that in the many-to-many case the situation changes.

Example 3 *Let us now consider the instance of the price-budget* CAP *given in Table 2. In this instance, at least four applicants are matched in each Pareto optimal matching M. First, note that $q(c_2) = 2$, while only a_1 and a_4 apply to c_2 and both have enough budget, so both are assigned to it and $M(a_4) = \{c_2\}$. Further, exactly one of a_2, a_3 is assigned to c_3, otherwise the trade-in-free property is violated. Hence, to achieve a POM with only three matched applicants, there are only two possibilities. It is easy to see that in both cases pair (a_5, c_1) violates the trade-in-free property (case 1: a_2 and a_5 unassigned; case 2: a_3 and a_5 unassigned).*

Further, it is easy to construct orders of proposals, producing Pareto optimal matchings with four and five matched applicants; see Table 3 for examples of such

	a_1, a_2, a_3, a_4, a_5	a_5, a_4, a_3, a_2, a_1
$M(a_1)$	$\{c_1, c_2\}$	$\{c_1, c_2\}$
$M(a_2)$	$\{c_3, c_4\}$	$\{c_4\}$
$M(a_3)$	$\{c_1\}$	$\{c_3\}$
$M(a_4)$	$\{c_2\}$	$\{c_2\}$
$M(a_5)$	$\emptyset$	$\{c_1\}$

Table 3: POMs with four and five matched applicants and the orders in SD to obtain them

POMs together with the orders in which the SD mechanism considers the choices of each applicant.

Now we show that in this example, not all Pareto optimal matchings can be obtained by serial dictatorship in the many-to-many case. Consider the matching M given by

$$M(a_1) = \{c_1, c_2\},\ M(a_2) = \{c_3\},\ M(a_3) = \{c_1\},\ M(a_4) = \{c_2\},\ M(a_5) = \{c_4\}.$$

It is easy to see that M is a POM. Suppose that M was obtained by SD. Then, since only applicants a_1 and a_4 are assigned to their first choices, one of them must have been the first one to make the choice.

Suppose that a_1 was the first. Then neither of the applicants a_2, a_3 and a_5 could be in the second position, since they would have chosen $\{c_3, c_4\}$, $\{c_3\}$ and $\{c_1\}$ respectively, as these are their first choices and they were still available after the move of a_1. Hence the second one to make the choice was a_4. After the choice of a_4, courses c_3 and c_4 are still available, and since they are among the first choices of a_3 and a_2, it can be the turn for neither of them. So the third one to choose must be a_5. But c_1, the first choice for a_5, is still available, so neither applicant a_5 can be in the third position.

Suppose that a_4 was the first one to make the choice. Again, a_1 must follow immediately and we arrive at exactly the same situation as before, when no player could make her choice as the third one.

Consider the following mechanism, called here *Generalized Serial Dictatorship* (GSD). Initially, all courses are closed and all applicants are labelled as *active*. At each round, let $S(a)$ denote the set of courses already assigned to applicant a and $A' \subseteq A$ be the subset of active applicants. Each GSD round amounts to arbitrarily selecting an applicant $a \in A'$ who receives her most preferable course c that is undersubscribed and satisfies $S(a) \cup \{c\} \in F_a$. If no such course exists, a is removed from A'. The GSD terminates once $A' = \emptyset$.

Intuitively, the GSD is a sequence $a^1, a^2, \ldots, a^r$ of applicants, in which repetitions may occur, such that applicant a^i selects a single course c^i.

Lemma 1 *The output of a GSD is a POM.*

Proof. The output of a GSD is a matching M, since quotas of courses and feasibility sets of applicants are checked at each GSD round. By Theorem 1, if M is not a POM, it is not maximal or not trade-in-free or not coalition-free. If M is not maximal, there is a course c that is undersubscribed and an applicant a such that $M(a) \cup \{c\} \in F_a$. Then, applicant a is active in GSD terms, a contradiction to the fact that GSD has terminated. If M is not trade-in-free, there is an undersubscribed course c and an applicant a such that $(M(a) \cup \{c\}) \setminus D_M(a,c) \in F_a$; but then, applicant a prefers c to any $c' \in D_M(a,c)$, a contradiction to the rule that at each GSD round an applicant selects her most preferable course that has an empty slot.

If there is a coalition $K = ((a_0,c_0),\ldots,(a_{r-1},c_{r-1}))$ in M, let $a_i, i \in \{0,\ldots,r-1\}$ be the applicant at the earliest among the GSD rounds in which the pairs in K were matched. This implies that a_i selected c_i instead of the more preferred c_{i+1}, which, by definition of a_i was selected by a_{i+1} at a subsequent round ; the only reason for that to occur would be that c_{i+1} had no empty slot, which in turns contradicts that c_{i+1} was available for a_{i+1} at a subsequent round. ■

The converse is also true.

Theorem 3 *Any POM is obtainable by the GSD in a suitable order.*

Proof. If M is a POM, then $G(M)$ is acyclic by Theorem 2 thus admitting a topological ordering τ. Let us denote by $(a^1,c^1),(a^2,c^2),\ldots,(a^{|M|},c^{|M|})$ the pairs matched in M ordered according to the inverse of τ restricted to matching vertices. For brevity, let us denote by $GSD(\tau)$ the realization of GSD with the order of applicants $a^1,a^2,\ldots,a^{|M|}$. We show that $GSD(\tau)$ outputs exactly M, i.e., in step i, applicant a^i will choose exactly c^i, for $i = 1,2,\ldots,|M|$.

To get a contradiction, suppose that $GSD(\tau)$ outputs some other POM, denoted by M'. Let j be the first step in $GSD(\tau)$, where applicant a^j chooses something different from c^j. Let us denote by $S^j(a)$ the set of courses that applicant a has obtained under $GSD(\tau)$ up to (but not including) step j. Notice that $S^j(a) \subseteq M(a)$ for each applicant $a \in A$. Let us consider two cases in turn.

Case 1. Applicant a^j did not choose c^j because this course was not available in step j. This means $|\{a \in A; c^j \in S^j(a)\}| = q(c^j)$. As $a^j \in M(c^j)$ too, this implies $|M(c^j)| \geq q(c^j)+1$, a contradiction.

Case 2. Applicant a^j chose a course $c^* \in P(a^j)\setminus S^j(a^j)$ such that $c^* \succ_{a^j} c^j$. Notice that in this case necessarily $c^* \in P(a^j)\setminus M(a^j)$. Let $M(a^j) = \{c_1,c_2,\ldots,c_k\}$ and suppose that the courses are written here in the order of decreasing preference of a^j and that $c^j = c_\ell$. Notice that $G(M)$ contains the path P joining the vertices $a^j, a^jc_k,\ldots,a^jc_\ell,\ldots,a^jc_1$ in this order.

Case 2a. Course c^* has a free slot in M. This means that $G(M)$ contains in addition to the arcs of path P arcs (c^*,a^j) and (a^jc_ℓ,c^*), thus a cycle. This is a contradiction to Theorem 2.

Case 2b. Course c^* is full in M. This means that there are applicants $a'_1,a'_2,\ldots,a'_{q(c^*)}$ who are matched to c^* in M. Due to the definition of $F_M(a^j,c^j,c^*)$, digraph $G(M)$ contains arcs (a^jc^j,a'_ic^*) for each $i = 1,2,\ldots,q(c^*)$. This means that in $GSD(\tau)$ all the applicants $a'_1,a'_2,\ldots,a'_{q(c^*)}$ had chosen c^* before step j and hence c^* was not available for a^j when making her choice. We arrived at a contradiction again. ■

Hence to obtain a POM, one can run the GSD with an arbitrary order of applicants. After an applicant has made her choice, she will either leave the market (if she cannot choose any available course to remain feasible) or enter the queue again. Let us suppose that each applicant can make her choice in constant time. Then, if a POM M is obtained, the number of rounds will be $O(|M|)$. In general, the computational complexity of GSD can be bounded by $O(|A| \cdot |C|)$, or more precisely, by $O(L)$, where L is the number of acceptable pairs.

5 Minimum and maximum Pareto optimal matchings

Let us denote by MIN-POM and MAX-POM the problems to decide, given an instance I of the many-to-many Pareto optimal matching problem and an integer k, whether I admits a Pareto optimal matching M of cardinality at most k and at least k, respectively. NP-completeness of the former problem has been proved in Theorem 2 of [5] and in Theorem 6.6. in [9] for the one-to-one case. We give here a slightly stronger result by proving the NP-completeness even for the case when the preference list of each applicant is restricted to contain at most two entries. By contrast, MAX-POM is polynomially solvable in the one-to-one case and we prove that it is also NP-complete here.

Theorem 4 MIN-POM *is NP-complete even in the one-to-one case and when the preference list of each applicant contains at most two entries.*

Proof. By Corollary 1, MIN-POM belongs to NP. We shall prove the NP-completeness by a polynomial transformation from VERTEX COVER. Let $G = (V, E)$ be a graph and k an integer. We define an instance I of POM as follows. For each vertex $v \in V$ we define a *vertex course* c_v and for each edge $e \in E$ an *edge course* c_e. Further, for each edge $e = \{u, v\} \in E$ there are two applicants e_u and e_v with preference lists

$$P(e_u) : c_e, c_u; \qquad P(e_v) : c_e, c_v.$$

Let the quota of each course be 1 and the feasible sets of courses of each applicant be singletons.

Now we show that G has a vertex cover of size at most k if and only if I admits a Pareto optimal matching M such that $|M| \leq |E| + k$.

Let $W \subseteq V$ be a vertex cover of size $\ell \leq k$. We construct a matching M of cardinality at most $|E| + \ell$ as follows. For each edge $e = \{u, v\}$ pick a vertex in W incident to e, say $u \in W$. Assign e_v to c_e in M and denote the set of these applicants by A_1. This means that the applicants not assigned so far belong to $\{e_u, e \in E; u \in W\}$. These applicants cannot be matched with edge courses (since all courses have quota 1 and are already matched with some applicant in A_1) and hence desire only vertex courses associated with vertices in W. Moreover, to each such course we can assign at most one acceptable applicant. Denote these applicants by A_2. Clearly, $|M| \leq |E| + \ell$ and it remains to show that M is Pareto optimal. This is easy, as M can be obtained by a serial dictatorship if we first let applicants in A_1 make their choices (in an arbitrary order), then applicants in A_2 (in an arbitrary order) and then the rest of applicants.

Conversely, let M be a POM such that $|M| \leq |E| + k$. Clearly, each edge course is full, otherwise M would not be maximal or would not be trade-in-free. Moreover, at most k further applicants are assigned to vertex courses. Let us denote the set of full vertex courses by W. To show that W corresponds to a vertex cover in G, let us suppose that for some edge $e = \{u, v\}$ both c_u and c_v are closed. As one of the applicants e_u and e_v is unassigned, this is a contradiction with maximality of M and this concludes the proof. ■

The following results is in a sharp contrast with with the polynomial solvability of MAX-POM in the one-to-one case. Abraham et al. [5] gave an $O(\sqrt{|A|}L)$ algorithm that was extended to the one-to-many case (Capacitated House Allocation problem) by Sng [13], see also [9], Chapter 6.

Theorem 5 MAX-POM *is NP-complete.*

Proof. By Corollary 1, MAX-POM belongs to NP. To show NP-completeness, we again give a polynomial transformation from VERTEX COVER. Let $G = (V, E)$ be a graph and k be an integer. Denote $|V|$ by n and $|E|$ by m. Define an instance I of POM as follows. For each vertex $v \in V$ there is an applicant a_v. For each edge $e \in E$ there is a course c_e with price 1. Moreover, there are $n - k$ special courses $z_1, z_2, \ldots, z_{n-k}$, each of price n. The quota of each course is 1 and the budget of each applicant is n. The preference list of applicant a_v contains first all the special courses ordered $z_1, z_2, \ldots, z_{n-k}$ and then all the courses corresponding to the edges incident to v in any strict order.

We show that G has a vertex cover of size $\ell \leq k$ if and only if I has a POM with cardinality $n + m - k$.

Let $W = \{w_1, w_2, \ldots, w_\ell\}$, $\ell \leq k$ be a vertex cover in G. For each edge $e \in E$ take a vertex $w \in W$ that is incident with e (if e is covered by two vertices in W, choose the one with the smaller index in W) and assign in M applicant a_w to c_e. The number of filled places is so far equal to m. Of the remaining $n - \ell \geq n - k$ applicants assign exactly $n - k$ of them to special courses arbitrarily, one applicant to each course. As all the courses are full, the size of the obtained matching $|M|$ is $n + m - k$. To see that M is Pareto optimal it suffices to realize that it can be obtained by serial dictatorship in the order $M(z_1), M(z_2), \ldots, M(z_{n-k}), w_1, \ldots, w_\ell$ and the rest of the applicants in an arbitrary order.

Conversely, let M be any POM in I of size $n + m - k$. As M is maximal and trade-in-free, all special courses are full. Applicants assigned to them cannot be assigned to any other course because of the budget constraints. This means that the remaining $n - (n - k) = k$ applicants are assigned all the m courses $c_e, e \in E$, and so these applicants correspond to a vertex cover of cardinality k. ■

Notice that the hardness results apply also when one wants to maximize or minimize the number of open courses. The latter might also have an economic interpretation: if each open course bears some fixed cost, minimizing their number means minimizing these costs.

By contrast, a POM that maximizes the number of assigned applicants can be found in polynomial time by the following procedure. First, given a many-to-many instance I derive an associated one-to-one instance J by leaving for each applicant only feasible sets of cardinality one and making as many clones of each course as is its quota. Find a maximum cardinality POM in J by the algorithm described in [5]. We know that in the one-to-one case each POM can be obtained by a serial dictatorship in a suitable order, again [5] provides a polynomial algorithm for finding this order. Then, after merging the clones of individual courses back, continue in the GSD with the used ordering of applicants as the starting point of the GSD. Since the number of assigned applicants will never be decreased, we can find a POM that maximizes the number of assigned applicants.

6 Acknowledgement

This work was supported by VEGA grants 1/0410/11 and 1/0479/12 and by the bilateral Slovak-Greek grant of The Slovak Research and Development Agency SK-GR-0005-11 (Cechlárová, Potpinková), by OTKA K108383 and the ELTE-MTA Egerváry Research Group (Fleiner). The authors also gratefully acknowledge the support of the Operational Program Education and Research funded by the European Social Fund, grant Education at UPJŠ – Heading towards Excellent European Universities, ITMS project code: 26110230056. The project is implemented under the Act "Bilateral S&T cooperation Greece - Slovak Republic 2011 - 2012". It is also co-financed by the European Union (European Regional Development Fund -ERDF) and the Greek Ministry of Education (GSRT) and is a part of the Operational Programme "Competitiveness & Entrepreneurship" (OPCE). (Eirinakis, Magos, Mourtos).

References

[1] A. Abdulkadiroglu, T. Sönmez, *Random serial dictatorship and the core from random endowments in house allocation problems*, Econometrica 66(3) (1998), 689–701.

[2] A. Abdulkadiroglu, T. Sönmez, *House allocation with existing tenants*, J. Econ. Theory 88 (1999), 233–260.

[3] A. Abdulkadiroglu, T. Sönmez, *School choice: A mechanism design approach*, Amer. Econ. Rev. 93 (2003), 729–747.

[4] A. Abdulkadiroglu, P. A. Pathak, A. E. Roth, *Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match*, American Economic Review 99(5) (2009), 1954–1978.

[5] D. Abraham, K. Cechlárová, D. Manlove and K. Mehlhorn, *Pareto optimality in house allocation problems*, Lecture Notes in Comp. Sci. **3341**, Algorithms and Computation, ISAAC 2004, Hong Kong, December 2004, Eds. R. Fleischer, G. Trippen, 3-15 (2004).

[6] R. K. Ahuja, T. L. Magnanti, J. B. Orlin, *Network flows: theory, algorithms, and applications*, Prentice Hall (1993).

[7] M. Balinski, T. Sönmez, *A tale of two mechanisms: student placement*, J. of Economic Theory 84 (1999), 73–94.

[8] Y. Chen, T. Sönmez, em Improving efficiency of on-campus housing: An experimental study, American Economic Review 92, 5, pp. 1669–1686 (2002).

[9] D. Manlove, *Algorithmics Of Matching Under Preferences*, World Scientific Publishing (2013).

[10] N. Perach, J. Polak, U. G. Rothblum, *A stable matching model with an entrance criterion applied to the assignment of students to dormitories at the Technion*, International Journal of Game Theory 36, 3-4, 519–535 (2008).

[11] A.E. Roth, M.A.O. Sotomayor, *Two-sided matching. A study in game-theoretic modeling and analysis*, Econometric Society Monographs, 18. Cambridge University Press, Cambridge, 1990.

[12] D. Saban, J. Sethuraman, *Complexity of Computing the Random Priority Allocation Matrix*, preprint (2013).

[13] C.T.S. Sng, *Efficient algorithms for bipartite matching problems with preferences,* Ph. D. thesis, University of Glasgow, Department of Computer Science (2008). Available at http://theses.gla.ac.uk/301/1/2008sngphd.pdf.

Recent IM Preprints, series A

2009

1/2009 Zlámalová J.: *On cyclic chromatic number of plane graphs*
2/2009 Havet F., Jendroľ S., Soták R. and Škrabuľáková E.: *Facial non-repetitive edge-colouring of plane graphs*
3/2009 Czap J., Jendroľ S., Kardoš F. and Miškuf J.: *Looseness of plane graphs*
4/2009 Hutník O.: *On vector-valued Dobrakov submeasures*
5/2009 Haluška J. and Hutník O.: *On domination and bornological product measures*
6/2009 Kolková M. and Pócsová J.: *Metóda Monte Carlo na hodine matematiky*
7/2009 Borbeľová V. and Cechlárová K.: *Rotations in the stable b-matching problem*
8/2009 Mojsej I. and Tartaľová A.: *On bounded nonoscillatory solutions of third-order nonlinear differential equations*
9/2009 Jendroľ S. and Škrabuľáková E.: *Facial non-repetitive edge-colouring of semiregular polyhedra*
10/2009 Krajčiová J. and Pócsová J.: *Galtonova doska na hodine matematiky, kvalitatívne určenie veľkosti pravdepodobnosti udalostí*
11/2009 Fabrici I., Horňák M. and Jendroľ S., ed.: *Workshop Cycles and Colourings 2009*
12/2009 Hudák D. and Madaras T.: *On local propeties of* 1*-planar graphs with high minimum degree*
13/2009 Czap J., Jendroľ S. and Kardoš F.: *Facial parity edge colouring*
14/2009 Czap J., Jendroľ S. and Kardoš F.: *On the strong parity chromatic number*

2010

1/2010 Cechlárová K. and Pillárová E.: *A near equitable* 2*-person cake cutting algorithm*
2/2010 Cechlárová K. and Jelínková E.: *An efficient implementation of the equilibrium algorithm for housing markets with duplicate houses*
3/2010 Hutník O. and Hutníková M.: *An alternative description of Gabor spaces and Gabor-Toeplitz operators*
4/2010 Žežula I. and Klein D.: *Orthogonal decompositions in growth curve models*
5/2010 Czap J., Jendroľ S., Kardoš F. and Soták R.: *Facial parity edge colouring of plane pseudographs*
6/2010 Czap J., Jendroľ S. and Voigt M.: *Parity vertex colouring of plane graphs*
7/2010 Jakubíková-Studenovská D. and Petrejčíková M.: *Complementary quasiorder lattices of monounary algebras*
8/2010 Cechlárová K. and Fleiner T.: *Optimization of an SMD placement machine and flows in parametric networks*
9/2010 Skřivánková V. and Juhás M.: *Records in non-life insurance*
10/2010 Cechlárová K. and Schlotter I.: *Computing the deficiency of housing markets with duplicate houses*
11/2010 Skřivánková V. and Juhás M.: *Characterization of standard extreme value distributions using records*
12/2010 Fabrici I., Horňák M. and Jendroľ S., ed.: *Workshop Cycles and Colourings 2010*

2011

1/2011 Cechlárová K. and Repiský M.: *On the structure of the core of housing markets*
2/2011 Hudák D. and Šugerek P.: *Light edges in* 1-*planar graphs with prescribed minimum degree*
3/2011 Cechlárová K. and Jelínková E.: *Approximability of economic equilibrium for housing markets with duplicate houses*
4/2011 Cechlárová K., Doboš J. and Pillárová E.: *On the existence of equitable cake divisions*
5/2011 Karafová G.: *Generalized fractional total coloring of complete graphs*
6/2011 Karafová G and Soták R.: *Generalized fractional total coloring of complete graphs for sparse edge properties*
7/2011 Cechlárová K. and Pillárová E.: *On the computability of equitable divisions*
8/2011 Fabrici I., Horňák M., Jendroľ S. and Kardoš F., eds.: *Workshop Cycles and Colourings 2011*
9/2011 Horňák M.: *On neighbour-distinguishing index of planar graphs*

2012

1/2012 Fabrici I. and Soták R., eds.: *Workshop Mikro Graph Theory*
2/2012 Juhász M. and Skřivánková V.: *Characterization of general classes of distributions based on independent property of transformed record values*
3/2012 Hutník O. and Hutníková M.: *Toeplitz operators on poly-analytic spaces via time-scale analysis*
4/2012 Hutník O. and Molnárová J.: *On Flett's mean value theorem*
5/2012 Hutník O.: *A few remarks on weighted strong-type inequalities for the generalized weighted mean operator*

2013

1/2013 Cechlárová K., Fleiner T. and Potpinková E.: *Assigning experts to grant proposals and flows in networks*
2/2013 Cechlárová K., Fleiner T. and Potpinková E.: *Practical placement of trainee teachers to schools*
3/2013 Halčinová L., Huntík O. and Molnárová J.: *Probabilistic-valued decomposable set functions with respect to triangle functions*

Preprints can be found in: http://umv.science.upjs.sk/index.php/veda-a-vyskum/preprinty