College admissions with stable score-limits

Péter Biró[1] and Sofya Kiselgof[2]

Abstract

A common feature of the Hungarian, Irish, Spanish and Turkish higher education admission systems is that the students apply for programmes and they are ranked according to their scores. Students who apply for a programme with the same score are in a tie. Ties are broken by lottery in Ireland, by objective factors in Turkey (such as date of birth) and other precisely defined rules in Spain. In Hungary, however, an equal treatment policy is used, students applying for a programme with the same score are all accepted or rejected together. In such a situation there is only one question to decide, whether or not to admit the last group of applicants with the same score who are at the boundary of the quota. Both concepts can be described in terms of *stable score-limits*. The strict rejection of the last group with whom a quota would be violated corresponds to the concept of H-stable (i.e. higher-stable) score-limits that is currently used in Hungary. We call the other solutions based on the less strict admission policy as L-stable (i.e. lower-stable) score-limits. We show that the natural extensions of the Gale-Shapley algorithms produce stable score-limits, moreover, the applicant-oriented versions result in the lowest score-limits (thus optimal for students) and the college-oriented versions result in the highest score-limits with regard to each concept. When comparing the applicant-optimal H-stable and L-stable score-limits we prove that the former limits are always higher for every college. Furthermore, these two solutions provide upper and lower bounds for any solution arising from a tie-breaking strategy. Finally we show that both the H-stable and the L-stable applicant-proposing score-limit algorithms are manipulable.

Keywords: stable matchings, college admissions, ties, manipulation
JEL classification: C78, I21

1 Introduction

Gale and Shapley [12] introduced a model and solution concept to solve the college admissions problem fifty years ago. In their model they suppose that the students submit preference lists containing the colleges they apply to, and each college ranks their applicants in a strict order and also provides an upper quota. Based on the submitted preferences a central body computes a fair solution. The fairness criterion they proposed is *stability*, which essentially means that if an application is rejected then it must be the case that the college must have filled its quota with applicants better than the our applicant's concerned. They gave an efficient method to find a stable matching and they proved that is actually optimal for the students in that sense that no student can be admitted to a better college in another stable matching. The Gale-Shapley algorithm has linear time implementation (see e.g. Knuth), which means that the running time of the algorithm is proportional to the number of applications. Another attractive property of this matching mechanism, proved by Roth , that it is strategy-proof for the students, i.e., no student can be admitted to any better college by submitting false preferences.

Later, it turned out (Roth [16]) that the algorithm proposed by Gale and Shapley had already been implemented in 1952 in the National Resident Matching Program and has

---

been used since to coordinate junior doctor recruitment in the US. Moreover, the very same method has been implemented recently in the Boston [4] and New York [3] high school matching programs. However, college admissions are still organized in a completely decentralized way in the US, with all its flaws, that is unraveling through early admissions and the coordination problems caused by too many or not enough students admitted. See some representative stories on American college admissions practices in the blog of Al Roth [26].

There are many other countries where higher education admissions are more regulated, but yet not centralized. In Russia, the common timetable of the admissions prevent the unraveling and the use of 'original documents' provide better coordination regarding the number of students admitted, but yet the solution is far from being optimal.[3] In the UK, there is a common platform to manage the admissions by UCAS [27] but there is no centralized matching mechanism, the decisions and actions of the users (students and higher education institutions) are still decentralized.

Finally, there are some countries which do have centralized matching schemes for higher education admissions. In particular, there are scientific papers on the Chinese [19, 20], German [9, 18, 22], Hungarian [6, 7], Spanish [15], Turkish [5] schemes.[4]

The Chinese higher education admissions system is certainly the largest in the world, with more that 20 million students enrolled in 2009 [20]. The system is based on a centralized exam, called National College Entrance Examinations, which provides a score assigned to each students and this induce a ranking of the students by universities. The matching process (see [19]) is a kind of Boston-mechanism with some extra tweaks that makes the system manipulable and controversial. The German clearinghouse for higher education admissions deals only with a small segment of subjects (about 13,000 student from the total 500,000, see [22]). The clearinghouse is a mixed system, in the first phase the Boston-mechanism is used and in the second phase the college-proposing Gale-Shapley, so the process is not incentive compatible [9, 18].

The Hungarian, Irish, Spanish and Turkish higher education matching schemes are all based on a centralized scoring system. The Irish system has not been described yet in a scientific paper to the best of our knowledge.[5] In the other three countries students are assigned a score with regard to each programme they applied to, these scores are coming mainly from their grades and entrance exams. The scores of a student may differ at two programmes, since when calculating the score of a student for a particular programme only those subjects are considered which are relevant for that programme. The solution of the admission processes are represented by the so-called *score-limits*, which are referred to as 'base scores' in Turkey [5] and 'cutoff marks' in Spain. The score-limit of a programme means the lowest score that allows a student to be admitted to that programme. The score-

---

[3]Each applicant applies to at most five universities, but does not inform universities about her preferences among them. Universities rank students using results of Unified State Exams. Two 'admission rounds' are organized that are similar to the first two steps of a deferred acceptance procedure. After the second step, universities that still have empty seats are allowed to organize additional admissions.

[4]However, we shall note that regrettably these scientific papers deal only with some special features of these systems (as we also do in this paper) so not all the aspects of these schemes are described. But luckily, there is a new European research network, called Matching in Practice [25], one of whose aim is to collect and describe current matching practices in Europe. So hopefully we will have a better picture and understanding on the current practices, at least in Europe.

[5]From the information published at the website of the Central Applications Office [21] it seems that the college-proposing Gale-Shapley algorithm is used in Ireland with some special features. One is that students can apply for 'level 8' and 'level 7/6' courses simultaneously, and these applications are processed separately, so a student may receive more than one offer at a time. There are deadlines for accepting offers and if offers are rejected then further offers are made by the higher education institutions, so the mechanism is somewhat decentralized. The tie-breaking is based on 'random-numbers' assigned to students with regard to each programme they applied for, so the ties are broken differently for different programmes involving perhaps the same applicants.

limits together with the preferences of the students naturally induce a matching, where each student is admitted to the first place on her list where she achieved the score-limit.

In Turkey [5] the ties are broken according to the date of birth of the students and the college-proposing Gale-Shapley algorithm is used. In Spain the scoring method is fine enough (the admission marks are from 5 to 14 with 3 decimal fractions, and some further priority rules are also used), so ties are very unlikely. They use the applicant-proposing Gale-Shapley algorithm with the special feature of limiting the length of the preference lists, a setting that creates strategic issues that were studied in detail by Romero-Medina [15] and Calsamiglia et al. [10].

In fact, in most applications where ties may occur, the programme coordinators break these ties. In the high school matching schemes in New York [3] and Boston [4] lottery is used for breaking ties. However, this may lead to suboptimal solutions as Erdil and Erkin [11] pointed out, but according to the study by Abdulkadiroglu et al [1] this is the only way to keep the mechanism strategy-proof. In the Scottish Foundation Allocation Scheme [24], where the junior doctors are matched to hospitals, the organizers attempt to break the ties in such a way that in the resulted matching as many doctors are allocated as possible (see Irving and Manlove [14]).[6]

In contrast, in the Hungarian higher education admission scheme [23] the ties are not broken, therefore the students applying for a particular programme with equal scores are either all accepted or all rejected. We call this an *equal treatment* policy.

In particular, the ties are handled in the following way in Hungary. No quota may be violated, so the last group of students with the same score, with whom the quota would be exceeded, are all rejected. There is however an alternative policy that could be followed where the quotas may be exceeded by the admission of the last group of students with the same score, but only if there were unfilled places left otherwise.

As we will show in Section 3, both concepts can lead to matchings that satisfy special stability conditions based on score-limits that we formalize in Section 2. We refer to the first, more restrictive solution as *H-stable* (i.e., higher-stable) *score-limits* and we call the second, more permissive solution *L-stable* (i.e., lower-stable) *score-limits*. Note that these stable score-limit concepts generalize the original notion of stability by Gale and Shapley, since they are equivalent to that if no tie occurs. In Section 4, we show how one can extend the Gale-Shapley algorithm to find H-stable and L-stable score-limits. Moreover, in Section 5 we prove that the applicant-oriented versions provide the minimal stable score-limits (therefore they are the best possible solutions for the applicants), whilst the college-oriented versions provide maximal stable score-limits (therefore, they are the worst possible solutions for the applicants).

Furthermore, we show in Section 5 that comparing the H-stable and L-stable score-limits, the L-stable score-limits are more favorable for the applicants as they are lower. In particular, we show that no college can have a higher score-limit in the applicant-optimal L-stable solution than in the applicant-optimal H-stable solution (and the same applies for the applicant-pessimal solutions produced by the college-oriented versions). Interestingly, we also show that the applicant-optimal solution produced after a tie-breaking is always between these two kinds of solutions. Therefore the matchings corresponding to the H-stable and L-stable score-limits may provide upper and lower bounds for every applicant regarding her match in a scheme which uses any kind of tie-breaking strategy. Finally, in Section 6 we give examples showing that neither the H-stable nor the L-stable version of

---

[6]In SFAS [24], applicants are ranked by NHS Education for Scotland in a so-called master list, in order of score each applicant has a numerical score allocated partly on the basis of academic performance and partly as a result of the assessment of their application form. The range of possible scores (approximately 40 100) is much smaller than the number of applicants (around 750 each year), so there are ties of substantial length in the master list.

the applicant-oriented score-limit algorithm is strategy-proof. We conclude in Section 7.

2   The definition of stable score-limits

Let $A = \{a_1, a_2, \ldots, a_n\}$ be the set of applicants and $C = \{c_1, c_2, \ldots, c_m\}$ be the set of colleges, where q_u denotes the quota of college c_u. Let the ranking of the applicant a_i be given by a preference list P^i, where $c_v >_i c_u$ denotes that c_v precedes c_u in the list, i.e. the applicant a_i prefers c_v to c_u. Let s^i_u be a_i's final score at college c_u. Final scores are positive numbers, as in practice the students with scores below a common minimum threshold are rejected automatically (this minimum score is currently 200 in Hungary, and it applies for every study).

The *score-limits* of the colleges are represented with a non-negative integer mapping $l : C \rightarrow \mathbb{N}$. An applicant a_i is admitted to a college c_u if she achieves the score-limit at college c_u, and that is the first such place in her list, i.e. when $s^i_u \geq l(c_u)$, and $s^i_v < l(c_v)$ for every college c_v such that $c_v >_i c_u$.

If the score-limits l imply that applicant a_i is allocated to college c_u, then we set the Boolean variable $x^i_u(l) = 1$, and 0 otherwise. Let $x_u(l) = \sum_i x^i_u(l)$ be the number of applicants allocated to c_u under score-limits l.

Furthermore, let $l^{u,t}$ be defined as follows: $l^{u,t}(c_u) = l(c_u) + t$ and $l^{u,t}(c_v) = l(c_v)$ for every $v \neq u$. That is, we increase the score-limit of college c_u by t (or decrease it if t is negative), but we leave the other score-limits unchanged.

To introduce the H-stable and L-stable score-limits, first we define the corresponding feasibility notions. Score-limits l are *H-feasible* if $x_u(l) \leq q_u$ for every college $c_u \in C$. That is, the number of applicants may not exceed the quota at any college. This means that the last group of students with equal scores, with whom the quota would be exceeded, are all rejected. Score-limits l are *L-feasible* if for every college $c_u \in C$ such that $x_u(l) \geq q_u$ it must be the case that $x_u(l^{u,1}) < q$. So the quotas may be exceeded at any college, but only with the worst group of students who are admitted there with equal scores.

We say that score-limits l are *H-stable* (resp. *L-stable*) if l are H-feasible (L-feasible) and for each college c_u either $l(c_u) = 0$ or $l^{u,-1}$ are not H-feasible (resp. L-feasible). Thus H-stability means that we cannot decrease the score-limit of any college without violating its quota assuming that the others do not change their limits. L-stability means that no college c_u can admit a student if at least q_u of its current assignees have a higher score, but otherwise the score limits must be as small as possible. H-stability is the concept that is currently applied in the Hungarian higher education matching scheme.

We note that if no tie occurs (i.e. every pair of applicants have different scores at each college), then the two feasibility and stability conditions are the same and they are both equivalent to the original stability concept defined by Gale and Shapley. The correspondence between stable score-limits and stable matchings in case of strict preferences was first observed by Balinski and Sönmez [5] in relation with the Turkish college admissions scheme (where ties do not occur due to a tie-breaking strategy based on the age of the applicants). Furthermore Azevedo and Leshno [2] have also used this observation in a general college admissions model involving continuum number of students.

3   Stable score-limit algorithms

Both the H-stable and L-stable score-limit algorithms are natural extensions of the Gale–Shapley algorithm. The only difference is that now, the colleges cannot necessarily select exactly as many best applicants as their quotas allow, since the applicants may have equal scores. If the scores of the applicants are all different at each college then these algorithms

are equivalent to the original one. In this section we will present the applicant-proposing and the college-proposing score-limit algorithms. For simplicity we describe these algorithms with regard to the H-stability concepts only and we add some information about the L-stable versions in brackets whenever they differ from the H-stable versions.

College-oriented algorithms:

In the first stage of the algorithm, let us set the score-limit at each college independently to be the smallest value such that, when all applicants are considered, the number of applicants offered places does not exceed its quota (resp. may exceed the quota but only if without the last tie of these students the quota is unfilled). Let us denote these score-limits by l_1. Obviously, there can be some applicants who are offered places by several colleges. These applicants keep their best offer, and reject all the less preferred ones, moreover they also cancel their less preferred applications.

In the subsequent stages, the colleges check whether their score-limits can be further decreased, since some of their offers may have been rejected in the previous stage, hence they look for new students to fill the empty places. So each college sets its score-limit independently to be the least possible that keeps the solution H-feasible (resp. L-feasible) considering their actual applications. If an applicant get a proposal from some new, better college, then she accepts the best offer, at least temporarily, and rejects or cancels her other, less preferred applications.

Formally, let l_k be the score-limit after the k-th stage. In the subsequent stage, at each college c_u, the largest integer t_u is chosen, such that $t_u \leq l_k(c_u)$ and $x_u(l_k^{u,-t_u}) \leq q_u$ (resp. if $x_u(l_k^{u,-t_u}) \geq q_u$ then $x_u(l_k^{u,-t_u+1}) < q_u$). That is, by decreasing its score-limit by the largest score t_u that keeps the solution H-feasible, i.e., where the number of applicants offered a place by c_u does not exceed its quota (resp. may exceed the quota but only if without the last tie of these students the quota is unfilled), by supposing that all other score-limits remained the same. For each college c_u let $l_{k+1}(c_u) := l_k^{u,-t_u}(c_u)$ be the new score-limit. Again, some applicants can be offered a place by more than one college, so $x_u(l_{k+1}) \leq x_u(l_k^{u,-t_u})$. Obviously, the new score-limits remain feasible.

Finally, if no college can decrease its score-limit then the algorithm stops. The H-stability (resp. L-stability) of the final score-limits is obvious by definition. Let us denote the corresponding solutions of the H-stable and L-stable versions by l_C^H and l_C^L, respectively.

Applicant-oriented algorithms:

Let each applicant propose to her first choice in her list. If a college receives more applications than its quota, then let its score-limit be the smallest value such that the number of provisionally accepted applicants does not exceed its quota (resp. may exceed the quota but only if without the last tie of these students the quota is unfilled). We set the other score-limits to be 0.

Let the score-limits after the k-th stage be l_k. If an applicant has been rejected in the k-th stage, then let her apply to the subsequent college in her list, say c_u, where she achieves the actual score-limit $l_k(c_u)$, if there remains such a college in her list. Some colleges may receive new proposals, so if the number of provisionally accepted applicants exceeds the quota at a college (resp. exceeds the quota and without the last tie of these students the quota is still filled), then it sets a new, higher score-limit $l_{k+1}(c_u)$.

Again, for each such college c_u, this is the smallest score-limit such that the number of applicants offered a place by c_u does not exceed its quota (resp. may exceed the quota but only if without the last tie of these students the quota is unfilled), by supposing that all

other score-limits remained the same. This means that c_u rejects all those applicants that do not achieve this new limit.

The algorithm stops if there is no new application. The final score-limits are obviously H-feasible (resp. L-feasible). The solution is also H-stable (resp. L-stable), because after a score-limit has increased for the last time at a college, the rejected applicants get less preferred offers during the algorithm. So if the score-limit in the final solution were decreased by one for this college, then these applicants would accept the offer, and the solution would not remain H-feasible (resp. L-feasible). Let us denote the corresponding solutions by the H-stable and L-stable applicant-oriented versions by l_A^H and l_A^L, respectively. The following result is therefore immediate.

Theorem 3.1. *The score-limits l_C^H and l_C^L obtained by the college-oriented score-limit algorithms are H-stable and L-stable, respectively. The score-limits l_A^H and l_A^L obtained by the applicant-oriented score-limit algorithms are H-stable and L-stable, respectively.*

4   Optimality of the outputs

It is easy to give an example to show that not only some applicants can be admitted by preferred places in l_A^H as compared to l_C^H, but the number of admitted applicants can also be larger in l_A^H (and the same applies for the L-stable setting). We say that score-limits l are *better* than l_* for the applicants if $l \leq l_*$, i.e., if $l(c_u) \leq l_*(c_u)$ for every college c_u. In this case every applicant is admitted to the same or to a preferred college under score-limits l than under l_*.

Theorem 4.1. *Given a college admission problem with scores, l_C^H are the worst possible and l_A^H are the best possible stable score-limits for the applicants, i.e. for any H-stable score-limits l, $l_A^H \leq l \leq l_C^H$ holds.*

Proof. Suppose first for a contradiction that there exists a H-stable score-limit l_* and a college c_u such that $l_*(c_u) > l_C^H(c_u)$. During the college-oriented algorithm there must be two consecutive stages with score-limits l_k and l_{k+1}, such that $l_* \leq l_k$ and $l_*(c_u) > l_{k+1}(c_u)$ for some college c_u.

Obviously, $l_k^{u,-t_u}(c_u) = l_{k+1}(c_u)$ by definition. Also, $x_u(l_k^{u,-t_u}) \leq q_u < x_u(l_*^{u,-1})$, where the first inequality holds by definition of t_u, as we choose the new limit for college c_u such a way that the number of temporarily admitted applicants does not exceed its quota. The second inequality holds by the H-stability of l_*. So there must be an applicant, say a_1, who is admitted to c_u at $l_*^{u,-1}$ but not admitted to c_u at $l_k^{u,-t_u}$.

On the other hand, the indirect assumption implies that $l_k^{u,-t_u}(c_u) = l_{k+1}(c_u) \leq l_*(c_u) - 1 = l_*^{u,-1}(c_u)$. Applicant a_1 has a score of at least $l_k^{u,-t_u}(c_u)$, which is enough to be accepted to c_u, so she must be admitted to some college c_v under $l_k^{u,-t_u}(c_u)$ which is preferred to c_u. Obviously a_1 must be also admitted to c_v under l_k. But the H-stability of l_* implies that $l_*(c_v) > l_k(c_v)$, a contradiction.

To prove the other direction, we suppose for a contradiction that there exists H-stable score-limits l_* and a college c_u such that $l_*(c_u) < l_A^H(c_u)$. During the applicant-oriented algorithm there must be two consecutive stages with score-limits l_k and l_{k+1}, such that $l_* \geq l_k$ and $l_*(c_u) < l_{k+1}(c_u)$ for some college c_u. At this moment, the reason for the incrementation is that more than q_u students are applying for c_u with a score of at least $l_*(c_u)$. This implies that one of these students, say a_i, is not admitted to c_u under l_* (however she has a score of at least $l_*(c_u)$ there). So, by the H-stability of l_*, she must be admitted to a preferred college, say c_v under l_*. Consequently, a_i must have been rejected by c_v in a previous stage of the algorithm, and that is possible only if $l_*(c_v) < l_k(c_v)$, a contradiction. $\square$

Theorem 4.2. *Given a college admission problem with scores, l_C^L are the worst possible and l_A^L are the best possible L-stable score-limits for the applicants, i.e. for any L-stable score-limits l, $l_A^L \leq l \leq l_C^L$ holds.*

Proof. Suppose first for a contradiction that there exist stable score-limits l_* and a college c_u such that $l_*(c_u) > l_C^L(c_u)$. During the college-oriented algorithm there must be two consecutive stages with score-limits l_k and l_{k+1}, such that $l_* \leq l_k$ and $l_*(c_u) > l_{k+1}(c_u)$ for some college c_u.

This assumptions imply that $x_u(l_k^{u,-t_u+1}) < q_u \leq x_u(l_*)$. Here, the first inequality holds by the L-feasibility of l_{k+1}, and the second inequality by the L-stability of l_*. At the same time, by our assumption, $l_*(c_u) > l_{k+1}(c_u)$, so $l_*(c_u) \geq l_{k+1}(c_u) + 1 = l_k^{u,-t_u+1}(c_u)$.

From the two above statements it follows that there must be an applicant, say a_1, who has a score $s_u(a_1) \geq l_*(c_u)$ and is admitted to c_u under l_*, but is not admitted to c_u under $l_k^{u,-t_u+1}$. So a_1 must have a seat at some college c_v under $l_k^{u,-t_u+1}$ such that $c_v >_{a_1} c_u$. Obviously, a_1 is also admitted to c_v under l_k. But a_1 is not admitted to c_v under l_*, therefore $l_k(c_v) < l_*(c_v)$, a contradiction.

To prove the other direction, we suppose for a contradiction that there exist stable score-limits l_* and a college c_u such that $l_*(c_u) < l_A^L(c_u)$. During the applicant-oriented algorithm there must be two consecutive stages with score-limits l_k and l_{k+1}, such that $l_* \geq l_k$ and $l_*(c_u) < l_{k+1}(c_u)$ for some college c_u.

At this moment, the reason for the incrementation is that more than q_u students are applying for c_u with score at least l_*, and c_u can choose a new score-limit $l_{k+1}(c_u) = l_k^{u,-t_u}(c_u)$, where $t_u > l_*(c_u) - l_k(c_u)$.

This implies that one of those students, who are admitted by c_u under l_{k+1}, say a_1, is not admitted to c_u under l_*. However she has a score higher than score-limit $l_*(c_u)$ there. So, by the L-stability of l_*, she must be admitted to a preferred college, say c_v, under l_*. Consequently, in the applicant-proposing procedure a_1 must have been rejected by c_v at some previous stage, and that is possible only if $l_*(c_v) < l_k(c_v)$, a contradiction. □

5 Comparison of the H-stable and L-stable versions

Intuitively it seems that the L-stable version of the algorithm is more *applicant-friendly* than the H-stable version. It turns out that we can prove the following result.

Theorem 5.1. *The score-limits obtained in the L-stable version of the applicants-oriented procedure are always equal or lower than the score-limits obtained in the H-stable version of the applicant-oriented procedure: i.e. $l_A^L \leq l_A^H$.*

Proof. Part I. Some colleges may have number of admitted students less than or equal to their quota under l_A^H, i.e. $q_u - x_u(l_A^H) \geq 0$. Each college c_u has a "waiting" list of applicants, who would prefer to be admitted to c_u rather than to their currently assigned colleges.

Let us apply some random tie-breaking to the original preference relation of the colleges. Each applicant a_i will get a new score $p_u^i \geq s_u^i$ such that no two applicants will have the same score at any college. Moreover, the new scores satisfy the following condition: if $s_u^j < s_u^i$, then $p_u^j < s_u^i$. These p_u^i scores are positive real numbers. For example, if there are three applicants with scores $s_u^1 = s_u^2 = 1$, $s_u^3 = 2$, the new scores might be $p_u^1 = 1$, $p_u^2 = 1.5$, $p_u^3 = 2$.

After that the following procedure is organized. If the number of applicants on c_u college's waiting list is more than the number of empty seats then college c_u sets it's new score-limit $m_A^H(c_u) \leq l_A^H(c_u)$ equal to the score p_u^i of the last admitted applicant in its

waiting list. Otherwise let $m_A^H(c_u) = 0$. Note that the new score-limits m_A^H are non-negative real numbers. This means that each college make offers to applicants from its waiting list who fit the new score-limit.

Some applicants may receive more than one proposal. Each applicant accepts one, from the most preferred college, and rejects the others. If there remain any empty seat in colleges then the second step is organized in the same manner and so on. Thus essentially we run a college-proposing deferred-acceptance procedure with regard to the new scores. At the end of this procedure some new score-limits m^R are achieved such that $m^R \leq l_A^H$ by construction. These new score-limits m^R and the corresponding matching μ^R are stable (in the Gale-Shapley sense) according to new strict preferences of colleges, also by construction.

Part II. For the strict preference profile and corresponding scores p_u^i from Part I we can organize applicant-proposing deferred acceptance procedure (which is, in case of strict preferences, equivalent to both the H-stable and L-stable applicant-oriented algorithms). The resulting matching μ_A^R is, of course, stable under strict preferences. Furthermore, we can define score-limits m_A^R that are equal to the score of the last accepted applicant if college has no empty seats and to 0 otherwise. These score-limits m_A^R must be the lowest among all stable score-limits by the optimality theorem of Gale and Shapley. Therefore $m_A^R \leq m^R$ in particular.

Part III. Now we deal with m_A^R score-limits. Let us get back to the original weak order preferences of the colleges and corresponding applicants' scores s_u^i. For each college with $x_u(l_A^R) = q_u$ we can construct a "waiting" list of applicants, who prefer college c_u to their current matches under m_A^R.

Let us now apply the L-feasibility concept. At the first stage each college sets it's new score-limit $l_A^R(c_u) \leq m_A^R(c_u)$, that is the largest value, which allows to admit equal or more than the quota under weak order preferences as L-feasibility prescribes. For example, if there are two applicants with the same score s_u^i, such that one of them is admitted to c_u under m_A^R and the other is on the waiting list then we have to 'treat them equally', so we should lower the score-limit. Each college makes offers to these additional applicants.

Some applicants may receive more than one offer from colleges; in this case each applicant chooses the most preferred college. After that if there is any college with number of admitted applicants less than its quota then a new round starts. Each college chooses new, lower, L-feasible limit, and so on. That is we run the college-proposing score-limit procedure under L-stability. At the end, some new score-limits l^L are achieved such that $l^L \leq m_A^R$ by construction. These new score limits are L-feasible and L-stable, obviously.

Part IV. For each L-stable score-limit l^L we know that $l_A^L \leq l^L$ from Theorem 4.2, where l_A^L are stable score-limits obtained by the L-stable applicant-oriented algorithm.

Now we can construct the following inequalities: $l_A^L \leq l^L \leq m_A^R \leq m^R \leq l_A^H$. So we can conclude that for any college admissions problem with score-limits the outcome by the L-stable applicant-oriented algorithm is better for the applicants (i.e. yields lower score-limits) than the outcome of the H-stable applicant-oriented algorithm.

$\square$

Theorem 5.2. *The score-limits obtained in the L-stable version of the college-oriented procedure are always equal or lower than the score-limits obtained in the H-stable version of the college-oriented procedure: i.e. $l_C^L \leq l_C^H$.*

Proof. Part I. Let us consider the l_C^L score-limits. Some colleges may have number of admitted students more than or equal to their quota, $x_u(l_C^H) \geq q_u$.

Let us apply a random tie-breaking to the original preference relation of the colleges. Each applicant a_i gets a new score $p_u^i \geq s_u^i$ such that no two applicants have the same score at any college, and these new scores do not contradict with the original ordering. Moreover, if $s_u^j < s_u^i$, then $p_u^j < s_u^i$). These p_u^i scores are positive real numbers.

After that the following procedure is organized. At the first stage each college sets its new score-limit $m_C^L(c_u) \geq l_C^L(c_u)$ such that according to the new scores p_u^i the number of applicants who fit this score-limit would be exactly q_u. The new score-limits m_C^L are non-negative real numbers. Let $m_C^L(c_u)$ be equal to 0 if the number of students admitted to c_u is less than c_u's quota and otherwise let $m_C^L(c_u)$ be equal to the lowest score p_u^i of any admitted student.

Some applicants are rejected from colleges they were assigned under l_C^L. Each rejected applicant then applies to the subsequent college in her list. Colleges receive new applications and, if necessary, raise the limits so that number of accepted applicants are equal to their quota. Some new applicants may be rejected, so a second round is organized in the same manner and so on. Thus we run an applicant-proposing deferred-acceptance procedure with respect to the perturbed strict preferences. At the end, some new score-limits m^R are obtained such that $m^R \geq l_C^L$ by construction. These new score-limits are stable (in the Gale-Shapley sense) according to the new strict preferences of colleges by construction.

Part II. For strict preference profile and corresponding scores p_u^i from Part I we can organize a college-oriented deferred-acceptance procedure. The resulting score-limits m_C^R are, of course, stable according to these strict preferences. Furthermore, the corresponding score-limits must be the lowest among all stable score-limits [12]. So, $m_C^R \geq m^R$.

Part III. Now we deal with m_C^R score-limits. For each college c_u, $x_u(l_A^R) \leq q_u$ holds under m_C^R. Each college c_u with number of assigned students lower than its quota has score-limit $l_A^R(c_u) = 0$. Now we get back to the original weak order preferences of the colleges and original applicants' scores s_u^i.

Let us now apply the H-feasibility concept. For each college we can construct a list of applicants, who prefer college c_u to their current matches under m_C^R. After that the following deferred acceptance procedure is organized. At the first stage each college sets it's new score-limit $l_C^R(c_u) \geq m_C^R(c_u)$ that is the smallest value, which allows to admit equal or less than the quota under weak order preferences as H-feasibility prescribes. Therefore some colleges may reject applicants. Each rejected applicant applies to the next college in her list. Colleges receive new applications and, if necessary, raise their score-limits in such a way that the number of accepted applicants is less than or equal to their quota. Some applicants may be rejected and a second round is organized in the same manner and so on. Thus we run an applicant-proposing deferred-acceptance procedure with regard to H-stability. At the end, each applicant is either accepted to some college or rejected by all acceptable colleges. Some new score-limits l^H are achieved such that $l^H \geq m_C^R$ by construction. These new score-limits are H-feasible and H-stable, obviously.

Part IV. For each H-stable score-limit l^H we know that $l_C^H \geq l^H$ from theorem 4.1, where l_C^H is a H-stable score-limit obtained by the applicant-oriented score-limit algorithm.

Now we can construct the following inequalities: $l_C^L \leq m^R \leq m_C^R \leq l^H \leq l_C^H$. So we can conclude that for any college admissions problem with score-limits the outcome by the L-stable college-oriented algorithm is better for the applicants (i.e. yields lower score-limits) than the outcome of the H-stable college-oriented algorithm.

$\square$

Corollary 1. *Applicant-optimal H-stable and L-stable scorelimits (l_A^H and l_A^L) are upper and lower bounds (respectively) for scorelimits under any Pareto-optimal stable matching with random tie-breaking.*

6 Strategic issues

Here we give two examples showing that neither of the above described score-limit algorithms is strategy-proof. The manipulability from the applicants' side is only interesting in the case

of applicant-oriented algorithms, as the applicants may successfully manipulate the college-oriented versions even for strict preferences (i.e., for scores with no ties). Therefore we only consider the applicant-oriented versions in the examples below.

Example 1. *Suppose that we have two colleges, c_u and c_v with one seat in each of them, and two applicants s_1 and s_2 applying to both c_u and c_v with a preference towards c_u and with equal scores at both places. So the preference list of the colleges and students are as follows.*

$$a_1 : c_u, c_v \qquad c_u : (a_1, a_2)$$
$$a_2 : c_u, c_v \qquad c_v : (a_1, a_2)$$

Figure 1: An example for the manipulability of the H-stable applicant-proposing algorithm

Here the only stable solution is the empty matching (i.e., score-limits higher than the scores of a_1 and a_2 at both colleges). However, if either of the students, say a_1 withdraws her application at c_u then the unique H-stable solution (under falsified preferences) is matching where a_1 is allocated to c_v and a_2 is allocated to c_u. So the manipulator (and actually the other student also) would improve.

The following example is essentially the same as the one that Hatfield and Milgrom [13] constructed in a different setting but for a similar purpose.

Example 2. *Suppose that we have two colleges, c_u and c_v with one seat in each of them, and three applicants a_1, a_2 and a_3 applying to both c_u and c_v with the following scores, $s_u^1 = 1$, $s_u^2 = 1$, $s_u^3 = 2$, $s_v^1 = 3$, $s_v^2 = 2$ and $s_v^3 = 1$. These can be described equivalently with the preference lists below.*

$$a_1 : c_u, c_v \qquad c_u : a_3, (a_1, a_2)$$
$$a_2 : c_v, c_u \qquad c_v : a_1, a_2, a_3$$
$$a_3 : c_v, c_u$$

Figure 2: An example for the manipulability of the L-stable applicant-proposing algorithm

Here the only L-stable solution is the matching $\{(a_1, c_v), (a_3, c_u)\}$ (i.e., with score-limits $l(c_u) = 2$ and $l(c_v) = 3$). However, if a_2 were to reverse her preferences with regard to the two colleges then the L-stable applicant-oriented algorithm would produce the matching $\{(a_1, c_u), (a_2, c_u), (a_3, c_v)\}$, where the manipulator (and actually both the two other applicants) would improve.

7 Further notes

In this paper we studied the concept of stable score-limits for higher education admissions. In particular we introduced and analyzed the notions of H-stability and L-stability when ties occur, a situation currently present in the Hungarian scheme.

As future research, we would like to investigate the college admissions practices of other countries, in particular those which have centralized systems based on score-limits. Regarding the Hungarian application, we would like to conduct an experiment with real data and compute the four possible extreme stable score-limits, namely the applicant-optimal vs applicant-pessimal score-limits under H-stability and L-stability. Finally, it would be also

interesting to see how these concepts can be used in other settings, e.g. what could be the corresponding solutions for the Boston and New York high school matching programs.

Regarding the theoretical problems, we would like to investigate whether there is any structure behind the H-stable and L-stable score-limits. It would be also worth to study further the relation of solutions satisfying equal treatment policy and those produced by tie-breaking strategies. For instance, one may can show some intuitive statements such that finer scoring methods lead to solutions 'closer' to the stable matchings obtained by tie-breaking strategies, and finer scoring methods are 'harder' to manipulate.

Acknowledgment

References

[1] A. Abdulkadiroğlu, P.A. Pathak, and A.E. Roth. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *American Economic Review*, 99(5):1954–1978, 2009.

[2] E. Azevedo and J. Leshno. A supply and demand framework for two-sided matching markets. *mimeo*, 2011.

[3] A. Abdulkadiroğlu, P.A. Pathak, and A.E. Roth. The New York City high school match. *American Economic Review, Papers and Proceedings*, 95(2):364–367, 2005.

[4] A. Abdulkadiroğlu, P.A. Pathak, A.E. Roth, and T.Sönmez. The Boston public school match. *American Economic Review, Papers and Proceedings*, 95(2):368–371, 2005.

[5] M. Balinski and T. Sönmez. A tale of two mechanisms: Student placement. *J. Econom. Theory*, 84(1):73–94, 1999.

[6] P. Biró. Student admissions in Hungary as Gale and Shapley envisaged. *Technical Report no. TR-2008-291 of the Computing Science Department of Glasgow University*, 2008.

[7] P. Biró, T. Fleiner, R.W. Irving, and D.F. Manlove. College admissions with lower and common quotas. *Theoretical Computer Science*, 411:3136–3153, 2010.

[8] P. Biró, F. Klijn. Matching with Couples: a Multidisciplinary Survey. To appear in *International Game Theory Review*, 2012.

[9] S. Braun, N. Dwenger, and D. Kübler. Telling the truth may not pay off: An empirical study of centralised university admission in Germany. 2010. The B.E. Journal of Economic Analysis and Policy, 10(1): Article 22.

[10] C. Calsamiglia, G. Haeringer, and F. Klijn. 'Constrained School Choice: An Experimental Study. *American Economic Review*, 100(4):1860–1874, 2010.

[11] A. Erdil and H. Erkin. What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, 98:669–689, 2008.

[12] D. Gale and L.S. Shapley. College Admissions and the Stability of Marriage. *Amer. Math. Monthly*, 69(1):9–15, 1962.

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

[14] R.W. Irving and D.F. Manlove. Finding large stable matchings. *ACM Journal of Experimental Algorithmics*, 14, section 1, article 2, 30 page.

[15] A. Romero-Medina. Implementation of stable solutions in a restricted matching market. *Review of Economic Design*, 3(2):137–147, 1998.

[16] A.E. Roth. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, 6(4):991–1016, 1984.

[17] C.-P. Teo, Y. Sethuraman, and W.-P. Tan. Gale-Shapley stable marriage problem revisited: Strategic issues and applications. *Management Science*, 47:1252–1267, 2001.

[18] A. Westkamp. An analysis of the German university admissions system. *mimeo, University of Bonn*, 2011.

[19] H. Zhang. An Analysis of The Chinese College Admission System. *PhD Thesis, University of Edinburgh*, 2009.

[20] Y. Zhang. The determinants of National College Entrance Exam performance in China - with an analysis of private tutoring. *PhD Thesis, Columbia University*, 2011.

[21] Central Applications Office, Ireland website. www.cao.ie

[22] D. Kübler. University admission practices - Germany. matching-in-practice.eu, accessed on 23 May 2012

[23] P. Biró. University admission practices - Hungary. matching-in-practice.eu, accessed on 23 May 2012

[24] R. Irving. Matching practices for entry-labor markets - Scotland. matching-in-practice.eu, accessed on 23 May 2012

[25] Matching in Practice website. matching-in-practice.eu

[26] Al Roth's market designer blog. http://marketdesigner.blogspot.com

[27] UCAS website. http://www.ucas.com

Péter Biró
Institute of Economics, Research Centre for Economic and Regional Studies
Hungarian Academy of Sciences
H-1112, Budaörsi út 45, Budapest, Hungary
Email: birop@econ.core.hu

Sofya Kiselgof
Laboratory of Decision Choice and Analysis (DecAn)
NRU Higher School of Economics, Moscow
Moscow, Pokrovskiy Boulevard, 11, Russia
Email: skiselgof@hse.ru