

Extended Abstract

On Matchability When Preferences Are Restricted To Shortlists

by

Ahmet Alkan

Sabancı University

October 2008

Centralized university student matching institutions, in Turkey, China and many other countries, operate essentially as in the Gale Shapley *college admissions* model : Students submit their orderings for university seats they have in mind, universities order students based on their scores in national exams and secondary school grade point averages, a stable matching is computed and enforced.

A particular criticism directed at this practice is that, not having the opportunity for any close look at individual students beyond what their scores reveal, university orderings disregard much relevant information, leading in fact to deterioration in pre-university education. A remedial institution, that has been recommended by the author, is *two – stage shortlist matching*, whereby in the first stage students and university seats are each matched with a *shortlist* of, say  $k$ , candidates, utilizing a stable *multipartner* matching procedure but with the coarse university orderings based on scores. In the second stage, universities and students are allowed to take a closer look at their potential mates in the shortlists and submit orderings subsequently. A stable matching  $\mu$  that assigns each student to at most one university is then computed and enforced.

The present paper offers an exploration on the *matchability* of participants in this setting, namely the likelihood that  $\mu$  will assign a student to a university that was in her shortlist at the conclusion of the first stage. We restrict attention to marriage markets with  $n$  individuals on each side and, in a stylized manner, probe the effect of  $k$  on matchability in worst cases. That is, we look for *minimum maximal matchings*, i.e., those with minimum cardinality, in general and some special classes of  $k$ -regular bipartite graphs of size  $n$ .

Let  $B = B(n, k)$  be a (connected)  $k$ -regular bipartite graph  $(M, W, k)$  with  $n = \text{card}M = \text{card}W \geq k$ . Let  $\mu$  be any maximal matching in  $B$ .

Theorem  $\text{Card}\mu \geq kn/(2k - 1)$ .

Proof: Define  $M^+ = \mu(W)$ ,  $W^+ = \mu(M)$ ,  $M^- = M - M^+$ , and  $W^- = W - W^+$ . By maximality of  $\mu$ ,  $B$  has no edges in  $M^- \times W^-$ . So  $M^- \times W^+$  and  $M^+ \times W^-$  each has  $k(n - \text{Card}\mu)$  edges. So the number of edges in  $M^- \times W^-$ , say  $\alpha = kn - 2k(n - \text{Card}\mu) = k(2\text{Card}\mu - n)$ . Thus  $\text{Card}\mu \leq \alpha$  iff  $\text{Card}\mu \geq kn/(2k - 1)$ .

Note  $\text{Card}\mu \leq \alpha \leq k\text{Card}\mu \leq n$  and for  $\alpha$  'midway'  $\text{Card}\mu$  is approximately  $3n/4$ .

Call  $m$  and  $w$  *admissible* for each other if  $(m, w)$  is an edge of  $B$ .

Say  $B = B(n, k, s)$  has *span*  $s$  if it is possible to order  $M$  and  $W$  such that, for every  $i$ , the admissible  $w$  for  $m_i$  is an ordered set  $(w_i, \dots, w_j)$  with  $j - i = s$ . Note  $s \geq k$ . Consider  $B(n, k, s)$ .

Define the maximal matching  $\mu^*$ .  $Card\mu^*$  is approximately  $2n/3$  independent of  $k$ .

Theorem :  $\mu^*$  is a minimum maximal matching.

The proof follows from the Proposition below :

Call any set  $I = (m_i, m_{i+1}, \dots, m_{i+j})$  of consecutive  $m$  an *interval*, and any  $R = (m_i, w_i, m_{i+1}, w_{i+1}, \dots, m_{i+j}, w_{i+j})$  a *bi-interval*, of length  $j + 1$

*Proposition* Let  $\mu$  be a minimum maximal matching. Every bi-interval of length  $3(s - 1)$  contains at least  $4(s - 1)$  matched elements.

Proof : Let  $R$  be any bi-interval of length  $3(s - 1)$  and  $u_m$  (resp  $u_w$ ) be the number of unmatched  $m$  (resp  $w$ ) in  $R$ .

Consider the case  $u_m \geq u_w$ . The proof for the other case is similar and will be omitted.

If  $u_m \leq s - 1$ , then  $u_m + u_w \leq 2(s - 1)$  and the result follows. If otherwise  $u_m \geq s$  then

by Key Lemma, there are at least  $s + u_m - 1$  matched  $w$ , and there are  $3(s - 1) - u_m$  matched  $m$ , adding up to at least  $4(s - 1)$  elements, in  $R$ .

*Key Lemma* : Suppose  $m_1$  and  $m_h$  are unmatched and there are  $u \geq s \geq 2$  unmatched  $m$  in the interval  $I = (m_1, \dots, m_h)$ . Then there are at least  $s + u - 1$  matched  $w$  in  $(w_1, \dots, w_h)$ .

Proof :

(i)  $h > u \geq s$ .

Otherwise  $h = u$  and *all*  $m$  in  $I$  are unmatched, in particular all the admissible  $m$  of  $w_h$ , contradicting the fact that  $w_h$  is matched (by maximality of  $\mu$ ).

(ii)  $w_s$  and  $w_h$  are matched,

by (i) and maximality, since  $m_1$  and  $m_h$  are unmatched.

(iii)  $h > s + 1$ .

Otherwise  $h = u + 1 = s + 1$  and by (ii)  $w_s$  and  $w_{s+1}$  are matched, which is impossible since all their admissible  $m$  are in  $I$  but  $I$  has only one matched  $m$ .

(iv) There is an unmatched  $w$  in the interval  $S = (w_{s+1}, \dots, w_{h-1})$ .

Otherwise all  $w$  in the interval  $(w_s, \dots, w_h)$  are matched, and note all their admissible  $m$  are in  $I$ . So  $I$  has at least  $h - s + 1$  matched  $m$ , namely  $h - u \geq h - s + 1$ , so  $u \leq s - 1$ . Contradiction.

Now let  $w_r$  be an unmatched  $w$  in  $S$  and  $m_q$  any unmatched  $m$  in  $I$  other than  $m_1$ . Then  $w_r$  is not admissible for  $m_q$  so either  $q + s - 1 < r$  or  $r < q$  and note that in the first case  $w_{q+s-1}$  and in the second case  $w_q$  is matched (by maximality).

Thus,  $m_1$  implies the existence of  $k$  matched  $w$ , and every other unmatched  $m$  in  $S$  implies the existence of at least one *distinct* matched  $w$ , adding up to at least  $k + u - 1$  matched  $w$  in  $(w_1, \dots, w_h)$ .