Extra Notes 5 (23/5 and 24/5)

These notes are based on the following paper:

D. Gale and L.S. Shapley, *College Admissions and the Stability of Marriage*, Amer. Math. Monthly **69**, No. 1 (1962), 9–15.

The general setup is that we are given the following data:

Dataset E.21. Given are

- two disjoint, finite sets X, Y.
- for each $x \in X$ a permutation π_x of the elements of Y. We shall write $y >_x y'$ if y appears before y' in the permutation π_x. We shall say that x *ranks* y *above* y'.
- analogously, for each $y \in Y$, a permutation π_y of the elements of X.

A *matching* will refer to a set of pairs (x_i, y_i) such that no element of $X \cup Y$ appears in two or more pairs. The elements x_i and y_i are said to be *matched*. The number of pairs is called the *size* of the matching and is denoted $|M|$. Obviously, $|M| \leq \min\{|X|, |Y|\}$.

Definition E.22. A matching M will be called *stable* if there does not exist any pair $(x, y) \notin M$ such that
- x is either unmatched or is matched to some y' such that $y >_x y'$, and similarly
- y is either unmatched or is matched to some x' such that $x >_y x'$.

Note that, for a stable matching M, one must have $|M| = \min\{|X|, |Y|\}$. For if $|M|$ were strictly smaller, then there would be at least one pair (x, y) such that neither x nor y was matched to anyone, and any such pair satisfies the conditions of Definition E.22.

For an arbitrary matching M, a pair (x, y) satisfying the conditions of Definition E.22 is said to *cause instability*.

Theorem E.23. (Gale-Shapley theorem) Given the dataset E.21, there always exists a stable matching.

The obvious application of this result is that X is a set of men and Y a set of women (or vice versa) and we want to matchmake in such a way that (i) as few people as possible are left single (ii) there will be no divorces. The theorem says that both conditions can be satisfied simoultaneously. For the proof, we will refer to the elements of X as *boys* and to those of Y as *girls*. A reader who finds this sexist is free to invent their own terms.

The proof involves the description of an explicit procedure for finding a stable matching, which is now known as the *Gale-Shapley algorithm*. The procedure involves a sequence of rounds, which we can describe inductively:

ROUND 1: Each boy x *proposes* to his favorite girl, i.e.: to the element $y = y(x)$

appearing first in the permutation π_x. Each girl y who receives more than one proposal *rejects* all but one of them, namely that from the boy who, amongst the ones who have actually proposed to her, she ranks highest in her permutation π_y. This boy she places on her *string*.

ROUND k: Each boy who has received a rejection in round $k-1$ proposes to his favorite girl amongst those from whom he has not yet received a rejection (equivalently, those to whom he has not yet proposed), provided at least one such girl remains. Each girl who receives at least one proposal in round k compares these proposals with the boy currently on her string and issues a rejection to everyone except the one she ranks highest. This boy is now placed on her string instead.

The procedure terminates at a round in which no rejections are issued, since then no boy will make any further proposals and so nobody does anything more. Since the boys will keep proposing as long as there is a girl left to propose to and since every girl has at most one boy on her string at any point, it is easy to see that the procedure will indeed terminate and, at that point, each girl who received at least one proposal will have exactly one boy on her string and no two girls will have the same boy on their string. Thus, if every girl now accepts the boy on her string, then exactly $\min\{|X|, |Y|\}$ pairs will be matched. It remains to show that this matching is stable.

So suppose, by way of contradiction, that there is a pair (x, y) which causes instability. Then x must either be unmatched or matched with some y' such that $y >_x y'$, and y must either be unmatched or matched with some x' such that $x >_y x'$. Since x ended up either single or with y', he must have at some point proposed to y' (and, if single, been rejected by her). Since he prefers y to y', this means he must at some earlier point have been rejected by y. But if y rejected x, she must have received at least one proposal from an x'' whom she ranked higher than x. In that case, there is no way x' could have ended up on her string, since she prefers x'' to x'. And there is no way she could have ended up single, since any girl who receives at least one proposal will be matched at the end. Contradiction !

Examples E.24. As our first example, we consider an extreme setting where there is universal agreement amongst the boys about how to rank the girls and vice versa. So let $|X| = m$, $|Y| = n$. We can set $X = \{1, \dots, m\}$ and $Y = \{1', \dots, n'\}$ such that $i < j$ means i is ranked above j by every girl and $i' < j'$ means i' is ranked above j' by every boy. Here we claim that there is a unique stable matching M, namely that which matches (i, i'), for $i = 1, \dots, \min\{m, n\}$. For suppose there was some other stable matching M^*. Since $M^* \neq M$, there must be a smallest i such that $(i, i') \notin M^*$. Thus $(j, j') \in M^*$ for all $1 \leq j \leq i-1$. Hence, either i is unmatched in M^* or matched to some k' where $k > i$. Similarly, either i' is unmatched in M^* or matched to some l where $l > i$. In all four cases, it is easy to see that the pair (i, i') causes instability in M^*, contradiction !

For our second example, consider the "opposite extreme" where the rankings are as different as they possibly could be. In this case, there can be many different stable matchings - see Example 1 in the Gale-Shapley paper. In particular, when looking for a stable matching one has three a priori strategies: (i) try to get the best possible outcome for the boys (ii) try to get the best-possible outcome for the girls (iii) make everyone compromise.

Our third example is Example 2 in the G-S paper. This example illustrates a setting in which there is a unique stable matching in which nobody gets their top choice. Welcome to real life !!

Definition E.25. Given a dataset E.21, an element $x \in X$ and an element $y \in Y$, we say that y is a *possible stable match* for x, and vice versa, if there exists at least one stable matching M such that $(x, y) \in M$.

Given a dataset E.21, an element $x \in X$ and a stable matching M, we say that M is *optimal for* x if there is no stable matching M' in which x is matched to someone he ranks higher than his match in M. If x is unmatched in M, this means that there is no stable matching *at all* in which x is matched. Optimality for an element $y \in Y$ is defined analogously.

Theorem E.26. *The matching produced by the Gale-Shapley algorithm is always optimal for every element of X, i.e.: for everyone in the group that does the proposing.*

Proof: It suffices to prove, by induction on k, that in Round k of the algorithm, no boy will be rejected by a possible stable match.

ROUND 1: Suppose, by way of contradiction, that some element $x \in X$ is rejected by a possible stable match $y \in Y$. Let M be a stable matching in which x is matched to y. Since y rejected x, she must have received another proposal from some x' such that $x' >_y x$. In Round 1, the boys propose to their favorites, so y must be the favorite of x'. Since $(x, y) \in M$, thus $(x', y) \notin M$. So either x' is unmatched in M or matched to some y' such that $y' <_{x'} y$. In either case, the unmatched pair (x', y) causes instability in M, since y prefers x' to x and y is the favorite of x'. Contradiction !

ROUND k: Assume that no rejections have been issued by possible stable matches in the first $k-1$ rounds and suppose, by way of contradiction, that some element $x \in X$ is rejected by a possible stable match $y \in Y$ in Round k. Let M be a stable matching in which x is matched to y. Since y rejected x in Round k, she must have received at least one further proposal, at or before Round k, from some x' such that $x' >_y x$. We consider two cases:

Case 1: $(x', y') \in M$ for some $y' >_{x'} y$. The fact that x' proposed to y at or before Round k means he must have been rejected by everyone he ranked higher, in particular by y', at or before Round $k-1$. But this contradicts our inductive asumption, since y' is a possible stable match for x'.

Case 2: Either x' is unmatched in M or matched to some y' such that $y' <_{x'} y$. But then, for the same reason as in Round 1, the unmatched pair (x', y) would cause instability in M, contradiction !

Stable assignments. One can extend the setting of the previous discussion by introducing an asymmetry between X and Y, namely, for each $y \in Y$, there is a *quota* $q(y) \in \mathbb{N}$. An *assignment* is a collection of pairs (x, y) such that each $y \in Y$ appears in at most $q(y)$ pairs. The previous discussion is the special case where $q(y) = 1$ for every y. Stability of an assignment can be defined in exactly the same way as in Definition E.22. It is then an exercise to check that the Gale-Shapley algorithm can be extended to find a stable assignment, for any values of the quotas, with the only modification being that, for $y \in Y$, "her" string has room for up to $q(y)$ elements of X. One natural application of this generalisation is to assigning students to colleges. You may read the G-S paper yourself for further details.