University of Glasgow

Harrenstein, P., Manlove, D., and Wooldridge, M. (2013) *The joy of matching.* IEEE Intelligent Systems, 28 (2). pp. 81-85. ISSN 1541-1672

http://eprints.gla.ac.uk/92963/

Deposited on: 4 April 2014

The Joy of Matching

Paul Harrenstein

University of Oxford, Oxford OX1 3QD, UK   paulh@cs.ox.ac.uk

David Manlove

University of Glasgow, Glasgow G12 8QQ, UK   david.manlove@glasgow.ac.uk

Michael Wooldridge

University of Oxford, Oxford OX1 3QD, UK   mjw@cs.ox.ac.uk

Introduction

Many problems in economics can be understood as *matchmaking problems*, in the following sense. We have two sets, and we need to pair up elements of the first set with elements of the second set. For example, the first set could be people looking for a job, while the second set might be employers looking for employees. Elements of both sets are assumed to have *preferences* over those with whom they could be paired. The problem is then, how do we find the *best matching* between individuals from both sets? This problem was first studied in the early 1960s, by David Gale and Lloyd S. Shapley [Gale-Shapley-1962].   They proposed a simple and elegant algorithm to solve the problem, and showed that this algorithm was, in a precise sense, optimal. For this foundational work, Shapley was a joint recipient, with Alvin Roth, of the 2012 Nobel Prize in Economics (see Sidebar One: Lloyd S. Shapley). Sadly, Gale passed away before the award was made (see Sidebar Two: David Gale). For computer scientists, one intriguing aspect of Shapley's award is that it was made, at least in part, for an *algorithmic* research contribution. Our aim in this article is to give you a feel for what matching problems are, how the Gale-Shapley algorithm solves them, and some possible applications of matching techniques.

The Stable Marriage Problem

Let us consider the basic problem that is solved by the Gale-Shapley algorithm. The problem is one of finding *stable matchings*. Following the original paper on this subject, the problem is usually described in terms of matching potential husbands with potential wives. Although this "marriage market" scenario may not seem at first sight to be terribly plausible, we will later see that in fact it reflects the structure of many important real-world economic problems. Here is how the problem was originally described:

A certain community consists of *n* men and *n* women. Each person ranks those of the opposite sex in accordance with his or her preferences for a marriage partner. We seek a satisfactory way of marrying off all members of the community. […] We call a set of marriages *unstable* if there are a man and a woman who are not married to each other but prefer each other to their actual partner. [Gale-Shapley-1962]

A little more formally, we have two disjoint sets, the set of men *M* and the set of women *W*, each of size *n*. Each man has a personal ranking of the women, and each woman has a personal ranking of the men. To keep things simple, we can assume the rankings are simply linear lists of all members of the opposite sex, indicating most preferred down to least preferred. Consider the following example (taken from [Gale-Shapley-1962]). Suppose *n* = 3, and that the men *M* = {*X, Y, Z*} and women *W* = {*A, B, C*} have rankings as follows:

EXAMPLE 1.	
Men	*X: A B C*
	Y: B C A
	Z: C A B
Women	*A: Y Z X*
	B: Z X Y
	C: X Y Z

Thus, *X* ranks *A* first, then *B*, then *C*, while *Y* ranks *B* first, then *C*, then *A*. The goal is to come up with a *matching*. A matching is just a mapping from men to women such that every man is mapped to a unique woman, and, conversely, every woman is mapped to a unique man. In mathematical terms, a mapping is a bijective function $\mu : M \to W$, and so $\mu(m) = w$ means that under the matching μ, man *m* is matched to woman *w*. The following is a matching for Example 1:

$$\mu(X) = A; \qquad \mu(Y) = B; \qquad \mu(Z) = C.$$

This matching pairs man *X* with woman *A*, man *Y* with woman *B*, and man *Z* with woman *C*. Since the mapping μ is a bijection, we can safely cheat a little with our notation, and write $\mu(w) = m$ to mean that the woman *w* is matched up with the man *m*; thus $\mu(B) = Y$ in the above matching, for example.

In contrast, the following is not a matching, since it associates the same woman (*A*) with two men:

$$\mu(X) = A; \qquad \mu(Y) = B; \qquad \mu(Z) = A.$$

Now, the goal of the matching process is not just to find a matching, but to find a *good* matching. There are at least two properties that we look for in a good matching. The first is *stability*. A matching is said to be *stable* if there is no man-woman pair who would prefer each other to the partners that they are matched with. Expressed a little more formally, μ is stable if there is no man *m* and woman *w* such that:

- *m* prefers *w* to μ(*m*), and
- *w* prefers *m* to μ(*w*).

So, if a matching is stable, then there will be no man-woman pair who would prefer to elope. Returning to our example above, there are several stable matchings. For example, consider the matching in which all the men are matched with their first choice women: μ(X) = A, μ(Y) = B, and μ(Z) = C. Since every man gets their first choice, there can be no man/woman pair who would rather elope than be with the partner they are matched with. By the same argument, the matching in which every woman is paired with her most preferred man is also stable. However, not all matchings are stable for this scenario: consider μ(X) = C, μ(Y) = B, and μ(Z) = A. In this case, *X* and *B* would prefer to elope.

Let's see another example.

EXAMPLE 2.	
Men	*X:* A C B
	Y: A C B
	Z: C A B
Women	*A:* Y Z X
	B: Z X Y
	C: Z X Y

Here, for example, the matching with μ(X) = B, μ(Y) = A, and μ(Z) = C is stable.

Stability forms our basic criterion for a good matching, but it is not the only one. In addition, we might consider *optimality* requirements to be important. However, formulating optimality is slightly tricky, because whether or not we judge a matching to be optimal can depend on whether we adopt the perspective of the men or the women. We say a stable matching is *M*-optimal if it is at least as good as any other stable matching for all the men; and similarly, we say a stable matching is *W*-optimal if it is at least as good as any other stable matching for all the women. For Example 1, the matching where every man gets their most preferred woman is *M*-optimal but not *W*-optimal, while the matching in which every woman gets their first choice man is *W*-optimal but not *M*-optimal.

The Gale-Shapley Algorithm

We have now seen the basic problem that we want to address: finding stable matchings between sets of individuals. However, we haven't yet seen how to find such matchings, or even whether stable matchings always exist. Moreover, if there are *n* men and *n* women, there will be *n*! possible matchings in total, and so exhaustively searching through the whole set of matchings will not be practicable: this raises the question of whether we can find stable matchings in polynomial time. We now introduce the Gale-Shapley algorithm (also sometimes called the *deferred acceptance algorithm*), which has some remarkable properties. First, it demonstrates that a stable matching always exists, and guarantees to always find such a matching. Second, it guarantees to find an *M*-optimal matching (or, with a simple tweak, a *W*-optimal matching). Third, it is computationally efficient: it runs in linear time.

The algorithm is as follows:

1. Each man *m* proposes to his most-preferred woman. Each woman who receives one or more proposals accepts the proposal of the man she prefers the most among those who propose to her; she rejects all the others. Men whose proposals are rejected remain single for the time being. When a man has a proposal accepted, we say the man and corresponding woman are *engaged*. Engagement is not in any way binding at this stage.

2. Each man who remains single after the previous stage proposes to the woman he most prefers among those he has not previously proposed to (in other words, he goes to the next woman on his list). Each woman who is not currently engaged who receives proposals accepts the proposal of the most preferred man, and rejects the rest. Each woman who is currently engaged accepts the proposal of whoever she most prefers, from the set of men she receives proposals from and the man she is currently engaged to; she rejects the rest (this may mean rejecting the person she is currently engaged to). Rejected men are then single.

3. Repeat stage (2) until everybody is engaged, at which point the set of current engagements becomes the final matching.

If we run the algorithm with preferences as in Example 1, we see that the men each make proposals to their most desired women, which are all accepted; everybody is then engaged, and the algorithm then terminates.

If we run the algorithm with preferences as in Example 2, then on the first round, *X* proposes to *A*, *Y* proposes to *A*, and *Z* proposes to *C*. The proposals of *Y* and *Z* will be accepted, but the proposal of *X* will be rejected. Thus *X* and *B* remain single; on the next round, *X* proposes to *C* but is rejected because *C*'s current engagement to *Z* is preferred by *C*. In the next round, *X* proposes to B and is accepted. The matching we end up with is $\mu(X) = B$, $\mu(Y) = A$, and $\mu(Z) = C$.

Let us consider properties of the algorithm. First, it is easy to see that it is guaranteed to terminate – the worst case would require $O(n^2)$ rounds, and so the algorithm runs in

polynomial time. (With an appropriate choice of data structure, this can be improved to linear time.) Also, it should be obvious that the matching produced is stable. Otherwise, there would be some man/woman pair who would rather be with each other than with the partner they are paired with in the final set of engagements. But since the man prefers the woman over his final engagement, he must have previously proposed to her and been rejected. So the woman cannot prefer the man over her final engagement. By a similar argument, we can see that the matching produced is *M*-optimal. To find a *W*-optimal matching, we simply have to swap the roles of men and women, so that the women are the ones making proposals.

Matching Problems in Practice

In their original paper on stable matching, Gale and Shapley wrote: "In making the special assumptions needed in order to analyze our problem mathematically, we [eventually] abandoned reality altogether and entered the world of mathematical make-believe. The practical-minded reader may rightfully ask whether any contribution has been made toward an actual solution of the original problem."

Interestingly, not only has the Gale-Shapley algorithm since been put to use in a wide range of application domains in many countries worldwide, but also, when they wrote their paper, Gale and Shapley were unaware that their algorithm was in fact already in use on a large scale, and had been so utilised for some eleven years prior to the publication of their paper!

To describe this interesting chronology in more detail, we rewind to the first half of the 20th Century, where in the United States medical interns sought positions at hospitals as part of their postgraduate medical education. The hospitals had more open positions than there were interns, resulting in a competition for the latter. As a consequence, many problems ensued: for example prospective interns were being signed up for an internship years before they started. Several solutions were proposed and tried, all with very limited success. The breakthrough came at the beginning of the 1950s with a voluntary program called NIMP (National Intern Matching Program), designed by J. M. Stalnaker and F. J. Mullin. Medical students and hospitals could now submit their preferences over one another, and an algorithm was then used to produce the allocation of students to hospitals.

In the 1970s it was observed that this algorithm was essentially equivalent to the "hospital-proposing" version of the Gale-Shapley algorithm (see Sidebar 2: David Gale). Here, the algorithm is an extension of the one described above for the Stable Marriage problem, which Gale and Shapley presented in the context of assigning students to colleges. However the model involving the assignment of medical residents to hospitals is entirely equivalent to the problem solved by the Gale-Shapley algorithm.

The NIMP, now known as the National Resident Matching Program (NRMP), survives to this day. In 2012, no less than 38,777 aspiring medical residents applied for 26,772 available residency positions via the NRMP. Counterparts of the NRMP are in existence in many other countries, including Canada, Japan, and the UK. It is the stability of the

constructed matching that is seen as being crucial to the success of the NRMP and other similar centralised matching schemes.

The Gale-Shapley algorithm is also used in many other contexts, for example in public school admissions in Boston and New York. The inputs include the preferences of pupils' parents over schools, and school "preferences" are determined through priority levels for pupils based on factors such as geographical location, sibling already at the school, etc.

Previously in Boston and New York, so-called "priority-based mechanisms" were used which led to widespread strategic behaviour on the part of the applicants, since it was often unsafe for applicants to submit their true preferences. This led to much uncertainty and dissatisfaction with the mechanism on the part of parents (and other stakeholders). One major advantage of the applicant-proposing Gale-Shapley algorithm is that it is *strategy-proof* for that side of the market. This means that it is safe for an applicant to reveal their true preferences, because they cannot obtain a better outcome by falsifying their preferences.

In New York, around 90,000 students enter high school each year. Prior to the mechanism being changed to the Gale-Shapley algorithm in 2003, around 30,000 students were allocated to schools that they had not ranked on their preference list. In the first year of operation of the new scheme, this number dropped to around 3,000.

The Gale-Shapley algorithm has been applied in many other settings, including:

• assigning children to daycare places in Denmark;
• higher education admission in Hungary and Spain;
• university faculty recruitment in France;
• placing military cadets in branches in the USA;
• placement of graduating rabbis in the USA and Israel;
• online dating in the USA and online matrimony in India;
• auction mechanisms for sponsored search in internet search engines.

It is likely that, with the increasing tendency for administrative processes to be handled online, and given the relative ease with which participants' preferences can be collected by this route, the potential for deploying the Gale-Shapley algorithm within centralised clearinghouses will only grow further.

Conclusions

At first sight, matchmaking problems of the kind discussed in this article may seem almost too simple to be of interest to theoreticians, and too unrealistic to be applicable in real-world settings. And yet, such problems turn out not only to be fascinating to work with from the point of view of mathematics and computing, but also have many practical and important applications.

Further Reading

[Gale-Shapley-1962] D. Gale and L. S. Shapley. *College Admissions and the Stability of Marriage.* In *The American Mathematical Monthly*, 69(1):9-15, January 1962. (The original paper on the Gale-Shapley algorithm: very readable and superbly clear. No complex mathematics involved!)

[Gusfield-Irving-1989] D. Gusfield and R. W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989. (An algorithmic study of the stable marriage problem and many related problems.)

[Poundstone-1992] W. Poundstone. *Prisoner's Dilemma.* Oxford University Press, 1992. (A fascinating history of the early days of game theory, and the role that game theory and the RAND Corporation played in cold-war military thinking.)

[Roth-Sotomayor-1990] A. Roth and M. A. O. Sotomayor. *Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis.* Cambridge University Press, 1990. (A standard text on the economics of matching.)

Sidebar One: Lloyd S. Shapley

Few academics can have made such a sustained and high-level contribution to their discipline as Lloyd S. Shapley. Born on June 2, 1923 to an academic family in Cambridge, MA, Shapley gained a Bronze Star for work on breaking enemy codes for the Army Air Corps during World War II. After the war, Shapley studied mathematics at Princeton, gaining his first degree in 1948. His PhD was supervised by Albert W. Tucker and he was awarded his research degree in 1953 – just three years after game theory legend John Forbes Nash gained his PhD under the same supervisor. He then worked at the RAND Corporation (*see Sidebar Three: The RAND Corporation*) until 1981, when he joined the faculty of UCLA. He spent the remainder of his working life at UCLA, and remains a Professor Emeritus to the present day.

Shapley's contributions to game theory and economics have been wide-ranging and fundamentally important. Apart from his work on game-theoretic matching, which we describe in the main body of the article, Shapley's best-known contribution is a "solution" for cooperative game theory – a solution that we now call the Shapley value. Roughly speaking, the problem that Shapley considered was the following. Suppose a group of agents work together, and cooperatively obtain some payoff, which can be shared amongst members of the team in any way that the team choose. Then, what is a *fair* way to divide the value a team obtains amongst members of that team? Shapley's approach was ingenious and hugely influential. First, he proposed a set of axioms, which he suggested that any *fair* payoff division scheme should satisfy. For example, he argued that players who are identical in terms of their contributions should receive the same payoff, and that players who make no contribution should receive nothing. Then, he defined a particular payoff distribution scheme, and showed that in fact this scheme is the *only* payoff scheme that satisfies the axioms. The basic idea behind what we now call the Shapley value is, crudely, that a team participant should be paid his *expected marginal contribution*: the amount that the player could be expected to contribute, on average, to a team selected at random. By only paying players according to the amount they contribute, we ensure fairness. Shapley's analysis remains at the heart of cooperative game theory, and has led to a huge body of subsequent research.

In October 2012, it was announced that the 89-year old Shapley would receive the 2012 Nobel Prize in Economics, jointly with Stanford economist Alvin E. Roth. The award citation was "for the theory of stable allocations and the practice of market design." From a computer science perspective, this award was particularly interesting because it was arguably the first time the citation for a Nobel prize referred to an *algorithmic* contribution: the algorithm in question being, of course, the Gale-Shapley algorithm.

Sidebar Two: David Gale

David Gale (13 December 1921 – 7 March 2008) was a professor of mathematics at the University of California, Berkeley, and during his career he made fundamental contributions to mathematics, game theory, and convex analysis.

He was born in Manhattan and received his PhD from Princeton in 1949. He was, of course, the joint author with Lloyd Shapley of the seminal paper "College Admissions and the Stability of Marriage," appearing in American Mathematical Monthly in 1962. Marilda Sotomayor, a long-standing colleague and friend of Gale, wrote that Gale realised in the 1970s that the algorithm he first published with Shapley had in fact been used since 1951 to allocate medical interns to residency positions in the United States.

Another important contribution of Gale was his "Top trading cycles algorithm" which appeared in 1974 (in fact in a paper authored by Shapley and Scarf). This algorithm has been of fundamental importance in economics and has been applied in the context of school choice and in the allocation of kidney donors to patients.

In 1980, David Gale shared the von Neumann Theory Prize with Harold Kuhn and Albert Tucker for their seminal role in laying the foundations of game theory, linear and non-linear programming.

According to an obituary by the University of California at Berkeley, in a note to Gale's family following his death, Al Roth wrote that Gale "has had a giant influence in economics as well as in mathematics" and mentioned that he had nominated Gale and Shapley to the Nobel Committee for Economics. David Gale would surely have been a joint recipient in 2012 had he lived longer.

Sidebar Three: The RAND Corporation

The RAND ("Research and Development") Corporation was founded in the aftermath of World War II, with the laudable and seemingly innocuous goal of promoting what we would now call technology transfer – taking theoretical ideas from mathematics, economics, logistics, and operations research, and applying them in real-world settings. An offshoot of the Douglas Aircraft Company, RAND enjoyed a close relationship with the US military from its foundation. In the 1950s and 60s, the work of RAND became very closely linked with cold-war military doctrine in the public perception. Many of the leading early proponents of game theory worked at RAND at some stage in their careers, including Lloyd Shapley, John Nash, John von Neumann, Kenneth Arrow, Thomas Schelling, Melvin Dresher, and Merill Flood, as well as hugely influential mathematicians such as Richard Bellman, the father of dynamic programming.

However, the researcher who did probably most to define the public perception of RAND was Herman Kahn. Born in 1922, Kahn was a futurist, who gained notoriety for trying to think rationally about something that most people regard as unthinkable: nuclear war. His 1960 book *On Thermonuclear War* set out his views on the subject, and it was perhaps the suggestion in this book that such a war was *winnable* that subsequently made it notorious. Many people believed it was only the idea that nuclear war was unwinnable that prevented the military from using nuclear weapons; the idea that RAND were apparently undertaking serious studies into how a nuclear war might be fought was for many an appallingly cold-hearted prospect. Kahn's views became widely associated with RAND, even to the point where they were parodied in Hollywood movies and made the subject of popular songs. In Stanley Kubrick's blistering cold-war satire *Dr Strangelove*, a mad scientist enthusiastically refers to a study by "The BLAND Corporation." Folk singer Pete Seeger's 1961 protest song *RAND Hymn* claimed "they think all day long for a fee/they sit and play games about going up in flames." So strong was the popular association between RAND, game theory, and nuclear war that this association ultimately contorted public perceptions of the discipline of game theory itself [Poundstone-1992].