Ethics Mini-Project 1: Stable Matchings

Version 2: We corrected a typo in the opening text of Subsection 2.3, "Proposing" (Gale-Shapley advantages proposers). We clarified the requirements in Find A Theorem to clarify what to do if the theorem is not numbered.

Due Date: This assignment is due at 11:59 PM Friday January 29 (Seattle time, i.e. GMT-8).
You will submit as a PDF to gradescope.

Algorithms have effects in the real-world. Stable Matchings in particular have been so widely-applied that research into them resulted in the awarding of the 2012 Nobel Prize in Economics. But no application in the real world is perfect. In this project you'll think about impacts of stable matching algorithms in the real world.

The goal of this exercise is for you to consider the effects of running algorithms in the real-world. This assignment is a mix of technical tasks (finding and applying theorems) and non-technical ones (considering tradeoffs between various real-world effects and groups). The technical aspects can be "right" or "wrong", but the non-technical aspects are unlikely to be simply "right" or "wrong" – we won't have to **agree** with the non-technical aspects of your analysis to consider them a good analysis. Our evaluation will be based on how well they connect to the technical aspects, as well as the depth of reasoning demonstrated.[1]

Collaboration Policy

You are to conduct your own search and analysis for this assignment. While you may get feedback form other students on your writing, you cannot just use the results of another student's search.

1. Application Review

Choose one of the following real-world uses of stable matchings:

- Medical Resident Matching (NRMP or programs in other countries)
- high school matchings (New York City, Boston, or other cities)
- Any other real-world application of stable matchings you can find

1.1. Find a Theorem

We've covered a few theorems about stable matchings in lecture (e.g. proposer-optimality). In a reliable source, find a theorem about stable matchings we **haven't** covered in class.

- Copy-paste the theorem statement, the theorem number (or name, or some other unique identifier in that text), and cite your source.
- restate the theorem (in your own words) applied to horses and riders
- then state it as applied to your real-world application (e.g. 'doctors' and 'hospitals' or 'students' and 'schools').

Some places to look:

- Two-Sided Matching by Roth and Sotomayor
- Algorithmics of Matching Under Preferences by Manlove
- The Stable Marriage Problem: Structure and Algorithms by Gusfield and Irving

[1] For example, if you say "Riders should propose to horses, because Gale-Shapley gives an advantage to people choosing between proposals" that's technically incorrect, because Gale-Shapley does the opposite.
If you say "Riders should propose to horses because Gale-Shapley gives an advantage to the proposing side, and I like riders more than I like horses" that's technically correct, but not well-thought out or justified.
If you say "Horses should propose to riders because Gale-Shapley gives an advantage to the proposing side, and riders are able to make an informed choice about whether to be a rider at this stable or another, while horses have no choice but to give rides at the stable they currently work at" that's correct technically, and well-justified (even if the staff-member grading believes horses aren't sentient and we should prioritize sentient species).

- Any other textbook or peer-reviewed paper

The first two books are available online through UW libraries (click the links). The third is available physically through UW libraries, but not online. For papers, you can usually find PDFs via google scholar (if they aren't available there, see if a librarian can help, or ask a staff member. Through UW library agreements, you should have access to just about every peer-reviewed paper written in the last 30 years).

Note that wikipedia/blog posts/etc. are **not** valid sources (though you may search through these places and then trace citations to find a reliable source).

1.2. Consider the implications

Identify three groups of people (or individuals) that are affected by this theorem. State the implications of the theorem for each of the groups. For each group also state whether you think it would be better for them to use a stable matching algorithm or a free-for-all market (like job markets in industry).

2. A New Application

2.1. Choose a new scenario

Identify a different real-world scenario where stable matchings **could** be used (but aren't currently used). Scenarios could be close-to-home/personal (say, something for your family or an RSO you are involved with) or much larger (graduate school admissions, a job market you're about to apply for, etc.), but don't use one of the examples we listed in the last section.

Choose a scenario, and clearly state: which agents are on each side, what a preference list looks like, and what a matching looks like (e.g. can agents on one side accept multiple applicants?).

2.2. Implications

Repeat the directions of Section **??** ("Consider the implications") for your new (possibly hypothetical) matching market. You may either reuse the theorem from the last section (if it still applies) or choose a new theorem (either from class or from a trusted source). If you choose a different theorem, copy-paste the theorem statement from a relaible source, and cite it (for this part, our course materials are considered "reliable")

2.3. Proposing

We learned in class that Gale-Shapley can disadvantage the choosing side, but there is a nice property we haven't discussed. Gale-Shapley is a "truthful" algorithm for the proposing side. That is, it is if you know you will be on the proposing side, it will never be to your benefit to lie about your preference list (this statement only applies to the proposing side, not the choosing side).

When you're choosing whether to implement the stable matching algorithm in your new context. After some experimentation on old preference lists, you realize that getting a matching "in the middle" isn't going to be feasible (there are too many matchings in the middle to look through them and pick one fairly). Your supervisor gives you these options for disclosing your methodology to participants:

1. Announce before you receive preference lists that you will run Gale-Shapley with one side proposing.
2. Announce before you receive prerence lists that you will run Gale-Shapley with the other side proposing.
3. Announce that you will flip a coin **after** receiving the preference lists. If its heads one side proposes, if it is tails the other side proposes.

Note that option 3 won't be "truthful" – you won't be able to tell people on either side that they won't benefit by lying. Consider the tradeoff between "truthfulness" and "fairness." Of those three options, which do you think is best in your scenario? Explain why in 3-4 sentences.

2.4. Summary

Based on what we've learned from class and the observations you've made so far, write a few sentences on whether you think stable matchings should be used in your scenario. (or if assignments should be made in a decentralized fashion, or some other model should be used to find an algorithm)