Document Purpose

In this document, we more carefully present the proof of correctness given for the Gale-Shapley algorithm.

Pseudocode

Here is a version of pseudocode for the algorithm.

Algorithm 1 Gale-Shapley (1962)

```
1: procedure StableMatching(Doctor Preferences, Hospital Preferences)
2:     while There is an unmatched hospital do
3:         The unmatched hospital that is first alphabetically makes an offer to their favorite doctor (that
           they haven't already made an offer to)
4:         if The doctor is unmatched then
5:             The doctor accepts the offer.
6:         else
7:             The doctor chooses their preferred offer between their current one and this new one.
8:         end if
9:     end while
10:    return all pairs
11: end procedure
```

Proof Overview

To prove the correctness of this algorithm, we must show that, no matter the input, the algorithm outputs a stable matching.

In order to do so, we should show that the output matching is stable (or has no blocking pairs). However, note that it's not clear that the output will be a full matching: after all, each hospital is only allowed to propose to each doctor once – what if they were rejected by everyone? For that reason, we actually need to show two things:

1. When it terminates, every hospital (and every doctor) has been matched.

2. The matching that was output is stable.

Let's go through them in order. These proofs are good models for the types of proofs we'll be writing in this course.

Proof of Claim 1

I want you to pay special attention to **how** we use the algorithm to write our proof.

Proof. We aim to prove that, at the end of the algorithm, every hospital is matched. We will proceed by **contradiction**.

Assume, for the sake of contradiction, that the algorithm terminates with an unmatched hospital h. Then there must also be an unmatched doctor d. However, according to line 3, h eventually must make an offer to every doctor, which means it must make an offer to doctor d. By lines 4-5, an unmatched doctor must accept an offer, so d would have accepted that offer. There is also no way for a doctor to become unmatched, so d must remain matched at the end. This is a contradiction to the fact that d was unmatched.

Therefore, it is impossible for there to be an unmatched hospital h. □

Notice how we explicitly cite lines of the algorithm to make our claims. The algorithm is used as source material to support our claims when helpful. You should view it as a reference that the reader has also read.

One very common mistake is to write out what the algorithm does. In general, you should assume that the reader understands the algorithm as well, so you don't need to tell them what it does. You only need to show that it is correct.

Proof of Claim 2 (Direct)

First is the direct proof presented in class, with more clarity.

Proof. We aim to prove that the matching output by the algorithm is stable, which means there are no blocking pairs. We will make a direct proof.

Take any pair (d, h) that isn't in the matching. If it's not in the matching, that means either (1) it was never proposed or (2) it was proposed and then rejected. We will show that, in either case, it is not a blocking pair.

For the first case, assume it was never proposed. Note that, by line 3, the hospital always proposes to its favorite remaining doctor. If (d, h) was never proposed, then the hospital has only proposed to doctors it prefers over d, so it is matched with a doctor it prefers over d. Thus, (d, h) can not be blocking.

For the second case, assume that (d, h) was proposed and then rejected. Rejections only occur in line 7, and this only occurs if the doctor has an offer they prefer over h. Thus, the doctor prefers its final offer over h, so (d, h) is not blocking.

We have shown that, for any pair (d, h) that isn't in the matching, no matter whether or not the pair (d, h) was proposed it can not be blocking. Thus, the matching is stable. □

Again, note that we're not rerunning the algorithm in this proof. We're only using the lines that are relevant to the argument, and assuming that the reader understands how the algorithm is run.

Proof of Claim 2 (Contradiction)

This direct proof was not the only way to prove the claim. Here is a proof by contradiction for you to compare.

Proof. We will show that there is no blocking pair. We will proceed by contradiction.

Assume that the pair (d, h) is a blocking pair. Say that h is matched with d' instead. By line 3, we know that h must have made an offer to d before it made the offer to d'. Thus, if (d, h) is not in the final matching, then d must have rejected the offer by h. However, by line 7, d can only reject the offer by h if it receives an offer from a hospital it prefers over h. This means that d is currently matched with a hospital it prefers over h, which contradicts the claim that (d, h) is a blocking pair. □

Here too, we are only using the algorithm where needed. Lines 3 and 7 are enough to prove that the matching is stable, so that's all we need to talk about.