**Boy-optimal stable matchings.** Fix an instance of stable marriage. Say that a girl $g$ is a *possible partner* for boy $b$ if there exists any stable matching that matches the pair $(b, g)$. Among the possible partners for boy $b$, call the one that $b$ prefers most the *best possible partner* for $b$. Define a stable matching to be *boy-optimal* if, for each boy $b$, the matching matches $b$ to his best possible partner.

A-priori, it is not clear that there is always a boy-optimal matching — it could be that some girl is not the best possible partner for any boy. But if there is a boy-optimal matching, it has to be unique, because each boy has just one best possible partner.

**Theorem 1.** *Any execution of the stable-marriage algorithm produces a boy-optimal matching.*

*Proof (long form).*

1. Consider an arbitrary execution of the algorithm.
2. By a lemma in a previous note, the algorithm returns a stable matching $M$.
3. Suppose for contradiction that $M$ matches some boy to a girl other than his best-possible partner.
4. That boy proposed to girls in order of preference, so he was rejected previously by his best-possible partner.
5. Consider the *first* iteration in which *any* boy is rejected by his best-possible partner.
6. Let $b_1$ be the rejected boy. Let $g_1$ be his best-possible partner.
7. In this iteration, girl $g_1$ rejects $b_1$, so $g_1$ must be matched to a boy $b_2$ that she prefers to $b_1$.
8. Let $M'$ be a stable matching with couple $(b_1, g_1)$. (It exists by definition of "possible partner".)
9. By the choice of $b_2$ in line 7, girl $g_1$ prefers $b_2$ to $b_1$.
10. Since $M'$ is stable, the pair $(b_2, g_1)$ is not unstable in $M'$.
11. By the previous three lines, $b_2$ prefers his partner in $M'$ to $g_1$.
12. So $b_2$ prefers his best-possible partner to $g_1$.
13. Recall that in this iteration of the algorithm, $b_2$ is matched to $g_1$.
14. But by line 12, $b_2$ likes $g_1$ less than his best-possible partner.
15. Since $b_2$ proposed in order of preference, $b_2$ was previously rejected by his best-possible partner.
16. Hence, this is not the first iteration in which some boy has been rejected by his best-possible partner.
17. This is a contradiction (with line 5). □

Here's an equivalent proof, using induction instead of contradiction.

*Proof (long form).*

1. Consider any execution of the algorithm. *setting up "to show ∀"*

---

**Lemma 1.** *Throughout the execution, the following invariant holds:*

*For every boy $b$ and girl $g$, if $g$ has rejected $b$, then $b$ does not have $g$ as a possible partner.*

*Proof (long form).*

1.1. The invariant is true before the first iteration, because there have been no rejections.
1.2.1. Consider an arbitrary iteration where the invariant holds before the iteration.
1.2.2. Let $g$ be the girl proposed to during the iteration.
1.2.3.1. First consider the case that $g$ doesn't reject anyone during the iteration.
1.2.3.2. In this case, there are no new rejections, so the invariant continues to hold.
1.2.4.1. Next consider the remaining case: $g$ rejects one boy, $b$ (who proposed to $g$, or who $g$ was matched to).
1.2.4.2. Let $b'$ be the boy that $g$ is matched to just after she rejects $b$.
1.2.4.3. Then $g$ prefers $b'$ to $b$ (because she rejected $b$ for $b'$).
1.2.4.4.1. Suppose for contradiction that $b$ has $g$ as a possible partner.
1.2.4.4.2. Let $M$ be a stable matching that matches $b$ to $g$. ($M$ exists by def'n of "possible partner".)
1.2.4.4.3. Let $g'$ be the girl matched to $b'$ in $M$. So, $M$ matches $(b, g)$ and $(b', g')$.
1.2.4.4.4. Since $M$ is stable and $g$ prefers $b'$ to $b$ (line 1.2.4.3), it must be that $b'$ prefers $g'$ to $g$.

1.2.4.4.5. So, in the algorithm, $b'$ was rejected by $g'$, some time before proposing to $g$ in this iteration.
1.2.4.4.6. By the invariant, this means that $b'$ does not have $g'$ as a possible partner.
1.2.4.4.7. This is a contradiction (as stable matching $M$ matches $b'$ to $g'$).
1.2.4.5. By Block 1.2.4.4, $b$ does not have $g$ as a possible partner.
1.2.4.6. So the invariant holds at the end of this iteration.
1.2.5. One of the two cases (Block 1.2.3 or 1.2.4) above must happen, and in either, the invariant continues to hold, so the invariant continues to hold after the iteration.
1.3. By Block 1.2, for each iteration where the invariant holds before the iteration, it holds after. □
1.4. It is true initially (line 1.1), so is maintained throughout.

2. By a lemma in a previous note, the algorithm returns a matching $M$.
3.1. Consider an arbitrary boy $b$. Let $g$ be the girl $M$ matches to $b$.
3.2. Boy $b$ (who proposed in order of preference) was rejected by every girl that he prefers to $g$.
3.3. By this and Lemma 1 above, no girl that he prefers to $g$ is a possible partner for $b$.
3.4. That is, $g$ is $b$'s best possible partner.
4. By block 3, $M$ matches every boy $b$ to his most preferred possible partner — $M$ is boy-optimal. □

Here's the short form of the previous proof.

*Proof (short form).* We argue that the algorithm maintains the following invariant:

*For every boy $b$ and girl $g$, if $g$ has rejected $b$, then $b$ does not have $g$ as a possible partner.*

The invariant holds initially because no rejections have occurred. Suppose the invariant holds before a given iteration. Consider the iteration. Let $g$ be the girl proposed to. If $g$ rejects no-one, by inspection of the invariant, it is maintained. So assume $g$ rejects one boy, say $b$. Let $b'$ be the boy that $g$ rejected $b$ for — $g$'s partner at the end of the iteration, who $g$ prefers to $b$.

To show the invariant is maintained, assume for contradiction that $g$ is a possible partner of $b$ — that is, some stable matching $M$ matches $b$ to $g$. Let $g'$ be the girl that $M$ matches to $b'$. Since pairs $(b, g)$ and $(b', g')$ are each matched in $M$, and $M$ is stable, and $g$ prefers $b'$ to $b$ (per the previous paragraph), it must be that $b'$ prefers $g'$ to $g$ (otherwise $(b', g)$ would be an unstable pair for $M$).

Since $b'$ prefers $g'$ to $g$, but, in this iteration of the algorithm, $b'$ is partnered to $g$, it must be that, in a previous iteration of the algorithm, $b'$ was rejected by $g'$. By this and the invariant (which held at the end of that iteration), girl $g'$ is not a possible partner of $b'$. This contradicts that $M$ matches $b'$ and $g'$. Hence, our assumption that $g$ is a possible partner of $b$ is false. This proves that the invariant is maintained.

Suppose the algorithm finally matches some boy $b$ to some girl $g$. Boy $b$ was rejected by every girl that he prefers to $g$. So, by the invariant, every girl that he prefers to $g$ is not a possible partner. That is, $g$ is $b$'s best possible partner. □

**Exercise.** Define a stable matching $M$ to be *girl-pessimal* if $M$ matches each girl to her least-preferred possible partner. Is the stable-marriage algorithm guaranteed to produce a girl-pessimal matching? Prove it or give a counter-example.