In lecture, we saw the interval scheduling problem, a greedy algorithm that solves it, and a proof that the algorithm produces an optimal solution. Of course, how we present it in a lecture is very different from how we would present it in writing. What follows is what would be a good write-up for the problem, if it were being written as a homework or exam proof instead.

Here is what the prompt might look like:

We are given a set of n intervals, each of which has a start time s_i and a finish time f_i. Our goal is to select as large of a subset of the intervals such that no two selected intervals overlap. Prove that the following **greedy algorithm** is optimal: Sign up for the class that ends earliest. Remove it and all overlapping classes from the set of available classes. Repeat this process until no classes remain.

For CompSci 161, it is sufficient for such a proof to show that our first selected interval is part of *some* optimal solution. This (obviously) is not a complete proof of the algorithm, but it is something of a meta-proof; it is a proof of the statement "if the professor asked me to write a complete proof of this algorithm's correct, I could do so."

Before reading further, I encourage you to take your lecture notes and convert them into what you believe would be a good homework or exam submission for the same problem. You can think of your lecture notes as your first draft and that as your submission draft.

. . . .

Suppose I have an optimal set of intervals OPT that does not include our first chosen interval. I claim that the set OPT$'$, which I form by removing the first-ending interval from OPT and replacing it with the first-ending interval from the input, is also an optimal set of intervals.

I will refer to the input's first-ending interval as x and the first-ending interval from OPT as y. This means that OPT$'$ is OPT with y removed and x added.

Clearly, OPT and OPT$'$ are both of the same size (number of intervals). What remains to be seen is that OPT$'$ is also a *valid* set of intervals: that is, we need to show that none overlap.

When we remove the first ending interval from OPT, but before we add the overall first-ending interval, there is no overlap among intervals in this set, because OPT was a valid output, so none of its intervals overlap.

In addition, we know that, because no intervals from OPT overlap, we know that all intervals *other than the first-ending interval* begin no earlier the first-ending interval ends. Because x ends before y ends, and all $z \in$ OPT-$\{y\}$ begin no earlier than y ends, this means (by transitivity) that x does not overlap with any elements of OPT-$\{y\}$. So OPT$'$is a valid set.

Because OPT$'$ is both valid (no intervals overlap) and equal in size to any optimal solution, it must be optimal. Therefore, there is an optimal solution that includes the first ending interval.