COS 330: Great Ideas in Theoretical Computer Science

Fall 2025

Precept 1

My Introduction to 330

Learn

Welcome to COS 330!

In this precept you will learn the format of 330 precepts, and at the same time practice solving some problems and writing proofs.

Precepts will typically have three sections: *Learn*, *Practice* and *Challenge*. The *Learn* section will be an extension of lecture and will typically teach you something new. The *Practice* section will have at least one problem that is similar to problems you will see on exams for you to solve either on your own or in groups. The *Challenge* section will have an extra problem that is more difficult than what you will see on exams, which is intended to entertain you if you finish the *Practice* section early.

Since this is the first precept, the *Learn* section doesn't actually contain any new material. Instead, you'll get to know your TA.

Practice

In today's *Practice* section, you'll work through a guided problem-solving session.

Problem 1

Recall that a bipartite graph has two sets of nodes, L and R, with all edges having one endpoint in L and the other in R. Recall also that a perfect matching is a set of edges such that every node is in exactly one edge.

Let G be a bipartite graph with n nodes on each side. Prove that if every node has degree $\geq n/2$, then G has a perfect matching.

Hint: You may use Hall's Marriage Theorem. Recall that Hall's Marriage Theorem asserts that a bipartite graph has a perfect matching if and only if for every set $S \subseteq L$ of nodes on the left, we have $|N(S)| \geq |S|$, where $N(S)$ denotes the set of nodes with an edge to some node in S.

Step One: Understand what the question is asking. This might seem simple, but is well worth spending time on; it's all too easy to try to prove something hard (or even false) because we misinterpreted a definition. You're not required to know the answer to all of these to answer the question, but reading up on them can help you understand it in more depth:

- What is a graph?
- What makes a graph *bipartite*?[1]
- What is a matching?
- What makes a matching *perfect*?
- What is the degree of a node?

It's absolutely fine not to know these by heart; you can look up the definitions online. Notice, also, that the question itself is only the second paragraph: the first reminds you of definitions, and the third is a hint (which you may or may not use).

Step Two: Try Stuff! Understanding exactly what you're supposed to prove is a (significant) part of a solution, but from then on there's no sure-fire strategy guaranteed to work. General tips are: try to construct a counterexample and see where you get stuck, or working through small examples (or both).

Let's consider the smallest non-trivial example of the problem we're trying to solve: with $n = 2$, call $L = \{u, v\}$ and $R = \{x, w\}$. The hypothesis of the statement says every node has $\geq n/2 = 1$ neighbor; that is, u is connected to x or w (or both), as is v (and x, w are each connected to one or both of u, v).

What would a "bad" G look like? Well, maybe both u and v are connected to w but not to z; this satisfies the degree condition for u, v and w but violates it for z. Of course, since we're trying to prove a true statement,[2] there's no such thing as a "bad" G; but it's a useful exercise to try to come up with graphs that "almost break" the statement.

OK, since we're given Hall's Marriage Theorem as a hint, let's use it. There are three sets $S \subseteq L$ to consider:

1. $S = \{u\}$;
2. $S = \{v\}$; and
3. $S = \{u, v\}$.[3]

Since u and v both have at least one neighbor, that covers the cases $S = \{u\}$ and $S = \{v\}$. But how about $S = \{u, v\}$? Well, the only way for $N(S)$ to be smaller than S is if u and v are both (only) connected to one node on R, say, x. But as we just saw (while trying to come up with a "bad" G), that would violate the degree condition on w, so $N(S) = R$ and the hypothesis for Hall's Marriage Theorem is satisfied; so G does have a perfect matching, and we solved the case $n = 2$.

Step Three: Generalize. Note that **this is not a full solution yet, but it makes a lot of progress**! And the intuition we gathered makes coming up with a full proof much more manageable. Indeed, we can follow the same blueprint of a *proof by contradiction*: show that for small S the degree condition immediately implies $|N(S)| \geq |S|$, then show that $|N(S)| < |S|$ for large S would violate the hypothesis. That is:

[1] Some things can be defined in multiple equivalent ways: for example, a bipartite graph and a 2-colorable graph (where we can color each node, say, red and blue so no edge connects two red or two blue nodes) are the same thing.

[2] We'll never ask you to prove something false – at least not on purpose!

[3] Can you see why there's no need to consider $S = \varnothing$?

1. if $|S| \le n/2$, any $v \in S$ satisfies $|N(\{v\})| \ge n/2$. Then $|N(S)| \ge |N(\{v\})| \ge n/2 \ge |S|$;
2. if $|S| > n/2$ and $|N(S)| < |S|$, any $w \in L$ that is not connected to S has at most $|R| - |S| < n/2$ neighbors, which contradicts the hypothesis.

Comparing the general case with the $n = 2$ one, you may see they're very similar: we took inspiration from a small example to prove the general case, and it worked!

Now, let's look at another approach to the problem (that will fail – your task is to figure out why). There are many proof strategies you can try, contradiction being only one of them. You can try to use induction and/or prove the contrapositive, for example. Here's a seemingly promising "proof" by induction on the number n of vertices, which doesn't use the hint:

> In the base case, $n = 1$, the single node in L is connected to $\ge 1/2$ nodes in R. Since the number of nodes is an integer, there's one edge from L to R, which is itself a perfect matching.
>
> For the inductive step, take any node $u \in L$ and any edge $\{u, w\}$ with $w \in R$. (There is at least one, since the degree of u is $\ge n/2 \ge 1$.) Consider the graph G' with $L' = L \setminus \{u\}$, $R' = R \setminus \{w\}$ and all the edges from L' to R' that exist in G.
>
> We only removed one node from R, so the degree of all nodes in L' is $\ge \frac{n}{2} - 1 \ge \frac{n-1}{2}$. The same argument shows the degree of nodes in R' is also $\ge \frac{n-1}{2}$. By the inductive hypothesis, G' has a perfect matching; adding the edge $\{u, w\}$ to this matching yields a perfect matching for G.

What is wrong with the "proof" above? Are there special cases when it is correct?

Problem 2

Let I be a set of n intervals described by $\{[s_i, e_i]\}_{i=1}^{n}$, where s_i and e_i are integers representing the start and end times of the i-th interval. Assume that $s_i < e_i$ for all i. We say that two intervals overlap if they share any point. The maximum cardinality of non-overlapping intervals is the largest number of intervals we can select from I so that no two of them overlap.

(a) Consider the following greedy algorithm to find a set of non-overlapping intervals, which we'll call the `start-time` algorithm:

- Sort the intervals in increasing order of their start times.
- Initialize an empty set S to store the selected intervals.

- Iterate through the sorted intervals, and for each interval $[s_i, e_i]$, if it does not overlap with any interval already in S, add it to S.

Show that this algorithm does not always find the maximum cardinality of non-overlapping intervals.

(b) Consider the following greedy algorithm to find a set of non-overlapping intervals, which we'll call the `end-time` algorithm:

- Sort the intervals in increasing order of their *end* times.
- Initialize an empty set S to store the selected intervals.
- Iterate through the sorted intervals, and for each interval $[s_i, e_i]$, if it does not overlap with any interval already in S, add it to S.

Prove the following three things:

1. The `end-time` algorithm always returns a set of non-overlapping intervals.
2. There is no set of non-overlapping intervals larger than the set returned by the `end-time` algorithm.
3. The `end-time` algorithm can be implemented in time $O(n \log n)$.

Challenge

Problem 1

You are boarding a flight at Newark airport that has n seats (assume $n \geq 2$, you can't afford a private jet yet), and it's fully booked (as usual). Every passenger has a unique assigned seat. The first passenger to board is flying for the first time, and sits in a uniformly random seat. The rest of the passengers do the following: when they board, if their assigned seat is free they sit in it, otherwise they choose a uniformly random empty seat and sit in it.

(a) Show that the last person to board sits in their assigned seat with probability $1/2$.

(b) Show that the expected number of people who board to find their assigned seat occupied is $1/2 + 1/3 + \cdots + 1/n$.