CS269I: Incentives in Computer Science
Lecture #1: The Draw and College Admissions*

Tim Roughgarden†

September 26, 2016

1 The Draw

So what do we mean by "incentives," anyway? Let's start with an example that should resonate with many of you — the *Draw*, meaning the process that assigns Stanford students to rooms in dorms and houses. Here's how the Draw works:[1]

The Draw

1. Each student submits a ranked list. Each item in the list is a type of room in a specific dorm or house (like a Toyon two-room double), and they are ordered from most preferred (at the top) to least preferred. (No ties are allowed.) This is how you express your *preferences*—that is, what you want—to the algorithm implementing the Draw.[2]

2. Each student is assigned a number in $\{1, 2, \ldots, 3500\}$.[3]

3. For $i = 1, 2, \ldots, 3500$:

    (a) Student i is assigned to her favorite choice among the options still available. In effect, the algorithm goes down student i's ranked list (from top to bottom), giving her the first option that is available.[4]

†Department of Computer Science, Stanford University, 474 Gates Building, 353 Serra Mall, Stanford, CA 94305. Email: tim@cs.stanford.edu.

[1] We'll ignore some bells and whistles, like pre-assignments.

[2] Importantly, there is no limit on the length of your list. Indeed, students are encouraged to list all of the types of housing accommodations that they'd be willing to accept.

[3] It's not important for us exactly how this is done. But for completeness: 1–1000 are reserved for "Tier 1," 1001–2000 for "Tier 2," 2001–3000 for "Tier 3," and 3001–3500 for "unguaranteed." A student can only use each Tier once (and Tier 1 only as a junior or senior). Each student receives a random number from the appropriate range.

[4] So for example, if your first choice is a Toyon double, you will be assigned to such a room if and only if less than 50 earlier students were already assigned to such rooms.

So that's the Draw. But here's the question.

Is the Draw a good system?

Or, would some alternative system be "better?"

How would you even phrase an argument either for or against the Draw? What kind of language would you use? What are some alternative systems to compare to? One of the primary goals of this course is to provide you with a language and a way of thinking that allows you to reason about these types of questions in a principled way.

1.1 Pareto Optimality

Speaking of language, here's a quick vocabulary lesson.[5]

> **Vocabulary Lesson**
>
> *Pareto optimal* (adj.): the property of an outcome that you can't make anyone better off without making someone else worse off.

It's often reasonable to say that a system that fails to produce a Pareto optimal outcome is not that great. If you can make some people better off without hurting anyone else, why didn't you do it? The converse is not so clear — just because an outcome is Pareto optimal doesn't necessarily mean we should be happy about it. For example, Pareto optimality says nothing about equity or fairness.

We've introduced one system (the Draw) and one property (Pareto optimality). Does the system have the property?

Proposition 1.1 *The outcome of the Draw is Pareto optimal.*[6]

In general, this will not be a proof-heavy class (unlike some of your instructor's other classes). We will still do some proofs, though, when they are short and informative (as is the case here).

Proof: Call the outcome of the Draw the "old" outcome. Consider some alternative "new" outcome. Assume that no student is worse off in the new assignment than in the old (otherwise the new assignment is no threat to the Draw's Pareto optimality). The plan is to show that the two assignments must be identical. Thus every assignment different from the outcome of the Draw makes someone worse off, and this establishes Pareto optimality.

We'll prove the following statement by induction on i (assuming no one is worse off in the new assignment): the first i students are assigned identically in the two assignments. The base case (with $i = 0$) is trivial, since two empty assignments coincide. For the inductive

[5] This is not an economics course, but we'll bring in terminology and tools from economics and game theory on a need-to-know basis. If you've taken at least one economics course, many of these terms will be familiar. In any case, literacy with these terms will serve you well in the future.

[6] No matter how numbers are assigned to the students in step 2.

step, fix a value of $i \geq 1$. By the inductive hypothesis, the first $i-1$ students are assigned identically in the two assignments. Thus, the remaining options for student i in the new assignment are precisely the options remaining when i is considered in the Draw. The Draw gives student i her favorite option among those remaining; if the new assignment does anything different, then student i is worse off. This completes the inductive step and the proof. ■

You should interpret Proposition 1.1 as an initial sanity check that the Draw might be a reasonable system. If the Draw was not Pareto optimal, we would already have some complaints about it.

1.2 Strategyproof Mechanisms

Proposition 1.1 and its proof implicitly assume that the ranked list submitted by a student really does reflect her true preferences. The Draw is not telepathic and can never really know anyone's true preferences; all it can do is use the reported preferences as a proxy for the true ones. But do we expect students to submit their ranked lists honestly? Or is there an incentive to "game the system" — is it possible to obtain a better result by lying about your preferences?

Vocabulary Lesson

strategyproof (synonym: *truthful*) (adj.): the property of a mechanism that honesty is always the best policy, meaning that lying about your preferences cannot make you better off.

So is the Draw strategyproof?

Proposition 1.2 *The Draw is strategyproof.*

Proof: Draw numbers are generated independently of the ranked lists that students submit. So you can't affect the draw number i that you get. You also can't affect the choices made by the $i-1$ students before you — their choices are independent of the ranked list that you submit. Thus, the available options at the time you are considered by the Draw does not depend on your submitted list. Since the Draw awards you your highest-ranked available option, it pays to be honest — any lie in your ranked list could only cause you to instead receive a remaining option that is not your favorite. ■

A quick question: does anything change if we swap the order of steps 1 and 2? That is, if you knew your draw number in advance, would it affect which ranked list you would submit? The answer is no, because the ranked list you want to submit (the truthful one) is independent of your draw number.[7]

The Draw is an example of a *mechanism*.

[7] We're assuming that the cost to you of formulating the list is independent of its length. If you were feeling lazy, you could imagine omitting the most popular houses (if you know you have a bad draw number) or the least popular houses (if you know you have a really good draw number).

Vocabulary Lesson

mechanism (n.): a procedure for making a decision or taking an action, as a function of what people want (i.e., of participants' preferences).

For example, the Draw produces an assignment of students to rooms, as a function of students' ranked lists. The specific type of mechanism used in the Draw also has its own name.[8]

Vocabulary Lesson

serial dictatorship (n.): a mechanism that orders the participants, and in this order allows each player to dictate their favorite feasible option (given the choices made by previous players).

2 The Draw in the 1990s

Maybe you're not so impressed with the discussion so far. After all, how else would anybody implement the Draw? (Other than maybe tweaking the process used to assign draw numbers.) Probably the Draw has been implemented this way forever, right?

Actually not. The rules of the Draw have changed many times over the years, most recently in 2009.[9] Here's how the Draw worked when your instructor was a Stanford undergrad (in the mid-90s).

The Draw (Mid-90s Version)

1. Each student submits a ranked list *of at most 8 options*, out of the roughly 60 possibilities.[10]

2. Same as before (modulo unimportant details): each student is assigned a number from some range, and these numbers are chosen independently of the submitted lists.

3. Same as before. But if none of student i's 8 options are still available, then the student is unassigned, or assigned by default to the least popular dorm.[11]

Is this slightly different mechanism strategyproof? Meaning is it always in your best interest to submit your favorite 8 dorms/houses, in order? The answer is *no*. After all, going

[8]Some prefer the less hostile term "priority mechanism."

[9]See the Web site for a 1970 Stanford Daily article about changes to the Draw; the other articles in the issue make for an even more amusing time capsule experience.

[10]We didn't specify a room type, just the house/dorm.

[11]This happened to your instructor's draw group his freshman year. Back then, the least popular dorm was Stern, which hadn't been renovated in a long time.

unassigned is a pretty bad outcome, worth safeguarding against. This motivates putting a "safety" in your eighth slot — a location unpopular enough that you're very likely to get it, but still at least a little better than the worst dorms. Thus in the old Draw, honesty was not always the best policy.

2.1 Why Care About Strategyproofness?

Is it so bad that a mechanism is not strategyproof? It's not necessarily a dealbreaker — after all, the old form of the Draw did an OK job for many years. But the strategyproof property is a nice one when you can get it. Think of it this way: today, when you submit your housing preferences, you have to do a fair amount of soul-searching to figure out what you really want. But this is a noble type of hard thinking — any mechanism worth its salt needs to solicit your preferences. Back in the 1990s, you additionally had to think hard about what options were likely to be available at the time you were considered. This second type of hard thinking arises as a byproduct of the specific mechanism chosen, and it is not fundamental to the underlying problem (as the current Draw shows). And why force participants to solve inessential hard problems?

Another problem with non-strategyproof mechanisms like the mid-90s Draw is *regret*, meaning that in hindsight a participant wishes they could go back and submit different preferences. (In today's Draw, you might wish you could have gotten a better draw number, but you'll never regret submitting your true ranked list.) Certainly your instructor's draw group regretted in hindsight not putting a safer safety in the eighth slot. Relatedly, the mid-90s Draw could easily produce non-Pareto optimal outcomes (can you find an example?).

Finally, let's ask a previous question again: would we expect participants' behavior in the mid-90s version of the Draw to change if we reversed the order of the first two steps?[12] That is, would knowing your draw number in advance change the ranked list that you submit? Sure — if you have a really good draw number you can focus on the most popular locations, while if you have a really bad number, you won't waste any of your slots on popular options.[13]

2.2 The Take-Away

Comparing the two versions of the Draw, we see a vivid illustration that *the rules of the game matter*. Seemingly small changes to a system can make a big difference in the experience of the participants, and in how well the system functions. We'll see this same phenomenon over and over again in this course, as we hop from one application domain to another.

Can you think of other systems that you regularly participate in that are clearly not strategyproof, or that do not produce Pareto optimal outcomes?

[12]This was in fact done at some point between the two versions of the Draw described here.

[13]Thus the problem faced in the mid-90s was particularly challenging: you had to estimate what would still be available when you were considered, "on average" over the draw numbers that you might get.

3 College Admissions

Let's move on to another process that you're intimately familiar with — the process of applying to colleges. Before discussing the system in place, let's be clear on the problem being solved.[14] The problem is to assign students to colleges; ideally, in a way respectful of what everybody wants. Whereas the Draw solves the problem of assigning students to rooms. Sounds like two versions of the same problem, no? So why don't we just use a serial dictatorship to decide who attends which college?

A big difference between room assignment and college admissions is that *colleges have preferences too.* (Presumably, rooms didn't care which students were assigned to them.)

> **Vocabulary Lesson**
>
> *two-sided market* (n.): a market with two distinct groups of participants, each with their own preferences.

College admissions is a two-sided market, with students on one side of the market and colleges on the other. In a two-sided market, it's not enough to choose—you must also be chosen. Room assignment is a one-sided market, since only one side of the market (the students) has preferences.

So, do you have any complaints about the current system for college admissions in the U.S.? For example, do you have to think hard about any issues other than the fundamental one of where you actually want to go?

Unlike the Draw, where there weren't any obvious problems to talk about, there are some justified complaints one can make about our college admissions system. For example, you need to think hard about how many colleges to apply to. (There is no hard limit like in the mid-90s draw, but each application is costly.) Given that you're not applying everywhere, the problems with 8-slot lists in the mid-90s Draw are also relevant here. A still more difficult problem is posed by the option to apply for early admission to a college. For example, imagine that your second-favorite college offers binding early admission (you have to accept or reject before hearing back from other colleges), but your favorite does not. Should you apply early to your second choice or not? The correct answer depends on the unknowable, namely on whether or not you would be accepted by your first choice.

Another distinction between the Draw and U.S. college admissions is that the former is centralized, with all assignments being made at the same time by a single clearinghouse, while the latter is decentralized, with each college making decisions independently. The decentralized solution may be more convenient and flexible from a logistical perspective, but it also suffers from otherwise avoidable inefficiencies. For example, if too many colleges independently go after the same small set of the best students, some colleges will have lower-

[14] Always differentiate between the problem being solved, and the specific solution that is being advocated. Many times people skip straight to the solution, without thinking carefully about what it is they're actually trying to accomplish.

than-expected yields and be unable to fill their class.[15] This phenomenon is quite prevalent in faculty hiring, where often all the top departments go after the same young hotshot, and almost all of them fail to hire anybody.

In light of our complaints, it's natural to ask: could there be a "better" mechanism for college admissions? For example, could there be some analog of the Draw, but for two-sided markets?

4 The National Resident Matching Program (NRMP)

When you're brainstorming for novel solutions to a problem, it's often useful to see how analogous problems have been tackled in other domains. To think about alternative systems for college admissions, let's examine the process by which newly minted doctors are assigned to hospital residencies.[16]

Medical residencies became widespread in the U.S. around 1900. (Before that, medical school graduates immediately started practicing.) From 1900 to 1945, hospitals competed for doctors in an ad hoc and decentralized way. As time went on, hospitals made offers to doctors earlier and earlier during their tenure at medical school. Why would this happen? Well if you're a hospital, a natural strategy is to try to make offers slightly earlier than everyone else to get an edge on your competition. Compounded over all the hospitals and multiple years, this incentive to make offers early produces absurd if predictable results. In 1945, it was standard to extend residency offers to medical students who had just finished their first year (i.e., two years before graduation). This wasn't good for anyone. The hospitals had little information on such young medical students and so were making potentially risky offers. Meanwhile, many of the medical students were committing to a position well before they knew what they actually wanted. When a market reaches this point, it is said to have *unraveled*. Unraveling is pretty common; for example, the market for matching law school graduates to clerkships is plagued by similar problems to this day. It also seems to be happening now with tech companies and computer science graduates.

Returning to the market for medical residents, in 1945 the situation was so bad that change was called for. Medical schools agreed that they wouldn't release any information about their students (enrollment status, transcripts, letters of recommendation, etc.) until an appropriate date, early in the final year of medical school. This stemmed the unraveling, but introduced different tricky incentives for residency programs. As you can imagine, on the magic date when medical student information was released, there was a mad dash by programs to secure the best doctors. Suppose you're running a hospital, and you extend an offer to your first choice. (No email in 1945; if you're lucky you can call them, otherwise we're talking about snail mail here.) Here's the worst-case scenario: after a period of deliberation, your first choice declines your offer in favor of a different program that she prefers; and meanwhile, your second, third, fourth, ... choices all accept offers from other programs.

[15] This is why colleges run waiting lists. Note that the Draw doesn't need waiting lists, except to accommodate students dropping out or entering the system after the assignments have been made.

[16] This discussion is drawn from Roth [1].

Because of timing issues, you might strike out with all the doctors that you want. This dilemma inevitably led residency programs to start making "exploding offers" which expire very soon after being extended. This way, the program still has a good shot at recruiting some good doctors, even if its first choice declines.[17] Thinking back to the college admissions process, one can regard binding early admissions as a form of unraveling and a type of exploding offer. We now see that the situation could conceivably be a lot worse than it is, with colleges admitting high-school students in their junior year, or sophomore year, or. . . (Actually, maybe this has already happened for highly coveted student-athletes.)

Fed up with the exploding-offer rat race, in 1952 hospitals decided to do something radical and move to a centralized clearinghouse—a single procedure to determine all assignments (analogous to the Draw). They proposed an algorithm for computing assignments as a function of the preference lists submitted by both medical school graduates and residency programs, and circulated it for comments. A committee of medical school students took umbrage with the proposed procedure, arguing that it was unfair. (See Exercise Set #1.) This committee went further and proposed an alternative procedure, which they argued was better. The powers-to-be were convinced and implemented the alternative procedure. Much of this remarkable algorithm persists in "the Match" of today, and we'll discuss it and its properties in detail next lecture.

References

[1] A. E. Roth. *Who Gets What And Why.* Mariner, 2015.

[17]The market for law school graduates can be particularly extreme. Roth [1, P.90] quotes a law school student who in 2005, on a flight from her first interview to her second interview, got 3 voicemail messages: the first extending an offer from where she just interviewed; the second to urge her to return the call soon; and the third to rescind the offer. The length of the flight? 35 minutes!