Intro

Today:

- ☐ Course basics
- ☐ Example problems

1 Administrative Stuff

- Go to the course homepage and read the syllabus.
- Get setup on Gradescope and Piazza. Use your campus ID for this.
- Recommended: Get setup on Overleaf. Alternatively or in addition, you can install LaTeX on your own machine.
- Recommended: Get setup on zyBooks, following the instructions in the syllabus.
- The first problem set is due next Wednesday.
- Introduce TAs: **John Chan** and **Zane Rubaii**.
- Introduce Graders: **Teo Anderson** and **Kev Rockwell**.
- Rule A: When online: turn on your camera, turn off your microphone (unless you want to ask or say something), and be engaged!
- Rule 1: When in person: phones and laptops must be off and in your bag. Starts 10 mins before class.
- Rule 2: When in person: proper mask, properly worn. Strongly requested: a well-fit N95 mask.

2 Other Course Basics

- This is Section A. Raissa D'Souza is teaching Section B. The two sections will not be much coordinated.
- Discrete (not discreet!) math. Deals with finite and countably infinite sets. But it's not a term that's routinely used by mathematicians, who say what they do more specifically. Like: ▷ **Set theory** (but no crazy-big infinite sets in our class). ▷ **Logic** (we do quite basic logic, with only a light touch on foundations). ▷ **Combinatorics** (how to count or enumerate things, but we focus on counting, not enumerating). ▷ **Probability** (on finite or countably infinite probability spaces). ▷ **Number theory** (just some basics, and maybe a touch of cryptography). ▷ **Graph theory** (meaning points and two-element subsets of them).

- This is a math class, but very different from calculus, differential equations, and the like. The contents of this course is more elementary in the sense that little is assumed and the material could be taught in high school or middle school. In fact, if you had a really good math teacher in middle school or high school, or if you competed in AMC math competitions, then you might have been exposed to a lot of this stuff. Yet most students will find most of what we do unfamiliar.
- I tend to ask more of my students than other CS professors. If you're wanting an easy class, I'm not your guy.
- On the other hand, I never talk down to my students. I feel like many discrete math books and classes do. I will assume a reasonable level of interest in our subject. I will try to improve what people call your "mathematical maturity." On each topic, I will try to find something non-trivial to say. I try to ask some challenging homework questions. You won't get them all right away. You might not get some at all. That's fine.
- We won't follow some fixed numerical grading scale that you might have become used to. You might be seeing lower percentages than you're used to. It doesn't mean you'll get a worse grade.
- I am not a strong believer in partial credit. Try to get things fully right. If you've made it through classes because of generous partial credit, and if you can't overcome this, then you're not going to do well.
- This term I am recommending a zyBook I set up by Sandy Irani. I wouldn't say that I am going to "follow it," but the topics we talk about are in the book, and I've pruned from the book much that we won't cover. It will be possible to earn credit by spending time with the book. But if you don't want to do that, you can do well in the class without using the book. I've constructed a grading scheme that lets you go either way.
- While I *am* interested in you as a person, it doesn't mean that I can recognize you, particularly if I see you outside this classroom, or you change seat, hair, or clothes. I am totally face blind. Please be kind and tell me who you are if the context doesn't make that clear.

3 Course Goals

1. Learn some standard (and rather useful) material.
2. Gain some mathematical maturity (including your ability to read and write proofs).
3. Improve your ability to think creatively in a rigorous domain.
4. Improve your technical writing.
5. Engage in some introspection.

4 Problem-Solving Techniques

Much of what one learns in this class from struggling working out problems. You are going to get stuck. You might feel stupid or frustrated. Here are some hints that I would like to suggest to help you solve hard problems. I am assuming here that the problem has already been abstracted out for you, so that the question itself is quite concrete and specific.

1. Reformulate to something equivalent
2. Generalize
3. Work out special cases. Small cases. Look for patterns.
4. Name things (e.g., introduce variables)
5. Create tailor-made definitions
6. Draw pictures
7. Think recursively
8. Adopt a playful attitude
9. Forget pattern-matching
10. But look for echoes
11. Know what you know (don't fool yourself, don't try to fool others)
12. Give serious attention to exposition. Never turn in a first draft. Critically read what you write.

For further advice on problem solving I like George Pólya's book *How to Solve It*. He describes a 4-step strategy: (1) Understand the problem; (2) Make a plan; (3) Carry out the plan; and (4) Look back. *Extra credit suggestion:* Read this book and write a 2-page summary and review of it. Post it Piazza.

5 First Example Problem

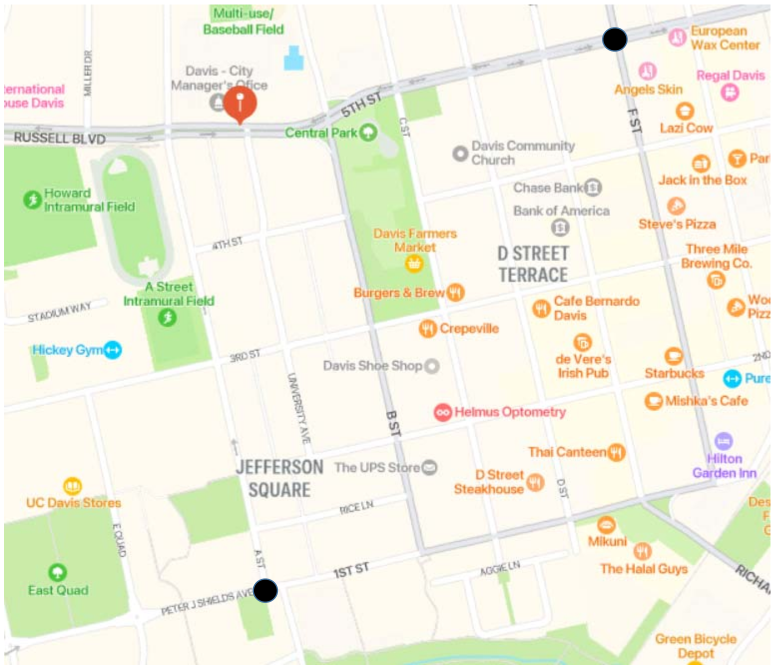
How many paths are there by which you can walk from one corner of downtown Davis, say the NE corner, to the opposite corner?

Step 1: First we need to refine the question. To make it precise. Usually problems in classes come pre-packaged so that, from your vantage, this step doesn't arise. But in the real world, this step is key. It involves making assumptions. Often these assumptions are convenient, but not really true.

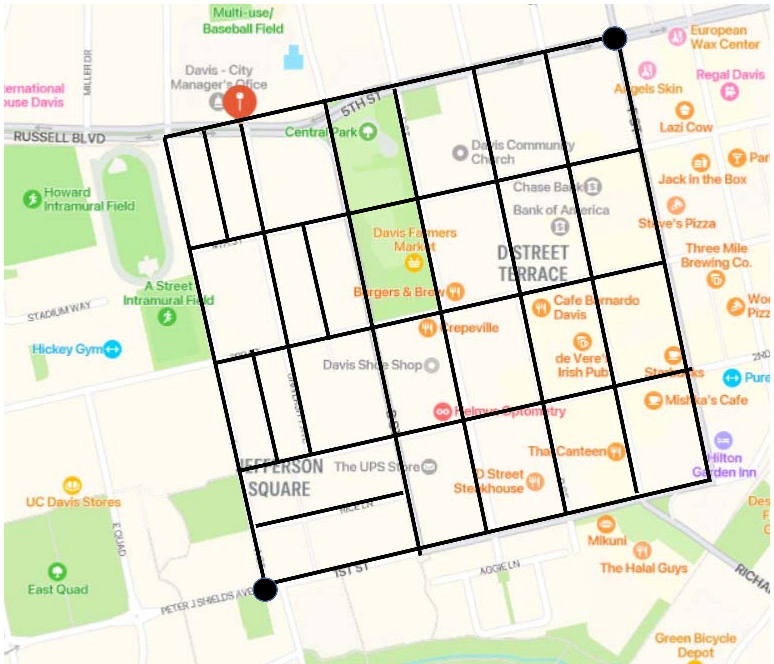
We will say that you are only allowed to walk on streets (including bike or pedestrian walkways) that show up on Google maps. That's our criteria for what a path is. We might call such paths *edges*. You travel from one *vertex* (intersection) to another along these edges.

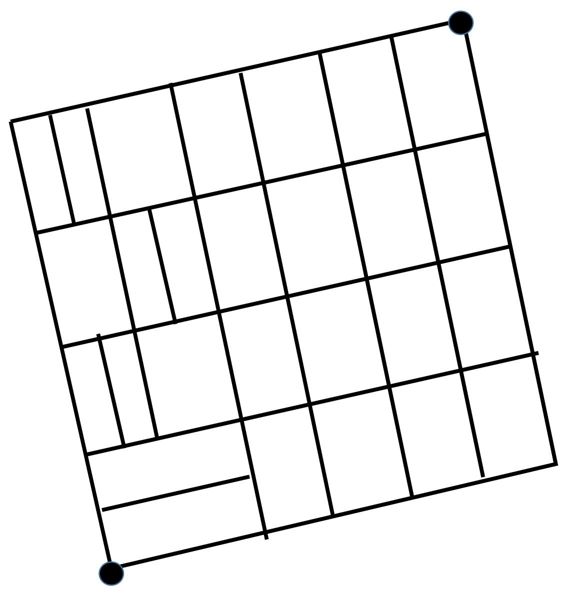
We are only going to attend to paths that are "reasonable": traversing each street should move you in the desired direction, lowering your distance to your destination.

We will say that the NE corner of downtown Davis is 5th and F; while the SW corner is 1st and A.



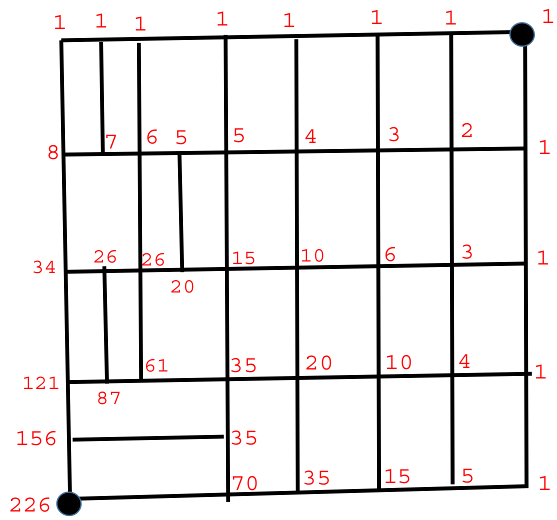
As part of this abstracting process, we can now throw away the information that no longer concerns us. But, at least in the back of your head, retain sight of the inaccuracies we incur when we make our abstractions. All too often people abstract problems and then treat the abstraction as though it were the real thing. It is not just students who do this, but professors and professionals. Indeed there is good reason to expect that increased training makes one *less* aware, and less skeptical, about mismatches between real-world problems and their abstractions.





Step 2: With an abstracted problem in hand, we must find an approach to its solution. Here, good techniques from our list include generalizing and thinking recursively. We can try to find the number of valid paths from the start vertex to each and every other vertex. The key insight is that the number of paths to a vertex is going to be the sum of the number of paths to its

nearer neighbors. The number of paths from the starting point to itself is 1. From this simple idea, we can iteratively but quickly label every vertex with the number of paths to it, as marked below. We get a concrete answer: that there are 226 paths from the start to the finish.



Step 3. Look back and think on what you have done. Ask related questions. Would one get the same number of paths going from the Finish to the Start? From the NW corner to the SE corner? What happens if you do this on an $n \times n$ regular grid? Or on all pairs of integers, moving away from the origin? Will all the Start-Finish paths have the same travel distance? In what ways was our initial model inaccurate model, and how might it be tweaked to improve its accuracy?

It is possible to be good at solving abstracted problems and lousy at generating them. Or lousy at contextualizing what a solution means. Computer Science departments are replete with such individual. Don't be too easily impressed by technical skill alone.

6 Second Example Problem

I claim that at least six riffle shuffles are needed to well mix a deck of 52 cards. Five is inadequate.

Here's what I mean by a *riffle shuffle*: You cut the deck into two approximately equal halves. You decide which hand the cards will first fall from. Then you let fall some number of cards from that hand. Then you switch and do the same. Keep going back and forth until all the cards have have fallen. Collect them neatly up and smile.

Think of doing this single riffle shuffle as a probabilistic way RIFFLE to mix up the deck. Think of doing it multiple times as a different mixing. For example, we could let MIX-5 denote doing RIFFLE five times. With independent probabilistic choices for each RIFFLE.

If you want to investigate in any depth how well mixed the cards are after n riffle shuffles then you will need a better model of RIFFLE: you would need to know the probability of different ways to cut the deck, and then the probabilities for how cards subsequently fall from your hands. But we're not going to do that. I prefer to be vague about the distribution on cards that RIFFLE induces, and I prefer to be vague on what it might mean to *well mix* our deck of cards.

Still, I can tell you something that would imply that we have *not* mixed the cards well: if some ordering of the cards is just *impossible* to reach.

That is, suppose you start with the cards in a fixed order, call it

$$\sigma = (1,\ 2,\ \ldots,\ 52).$$

If there's some ordering τ that is *impossible* to get to after some mixing, then we might say that this mixing is *deficient*. Cards aren't well-mixed after a deficient mixing.

> As an aside, let's count how many ways are there to order our 52 cards. I claim it is $52! = 52 \cdot 51 \cdot 50 \cdot \cdots \cdot 3 \cdot 2 \cdot 1$. (This is a big number—nearly 10^{68}.) This because there are 52 cards that we can put first; and then, having made some choice, there are 51 remaining cards that we can place second; and so on.
>
> In an ideal shuffle, if you start with the cards in the fixed order σ, then, at the end of the shuffle, all 52! ordering should be equally likely. That is, each should occur with probability 1/52!. If a shuffle is deficient, there is some sequence τ that is not arising with probability 1/52!, but with probability 0. That's a pretty bad defect.

I claim that MIX-5 is a deficient shuffle. Proving this requires inventing a clever definition. Inventing a good definition was one of the techniques I enumerated for solving hard problems.

Let $s = (a_1, a_2, \ldots, a_n)$ be list of values. Then a **subsequence** t from s is a sequence obtained from s by excising (removing) any of its entries. The rest stay put. Formally, it's a sequence $t = (a_{i_1}, a_{i_2}, \ldots, a_{i_k})$ where $1 \le i_1 < i_2 < \cdots < i_k \le n$. So, for example, $(3, 5, 7)$ is a subsequence of $(10, 3, 4, 5, 10, 7)$, while $(3, 5, 4)$ is not.

Definition 1 *A* ***rising sequence*** r *from a sequence of integers* s *is a maximal subsequence from* s *of consecutive integers.*

Saying that $r = (i, i+1, \ldots, i+k)$ is *maximal* means that s does not have a subsequence $(i-1, i, i+1, \ldots, i+k)$, nor a subsequence $(i, i+1, \ldots, i+k, i+k+1)$: the consecutive integers of r can't be *extended* on the left or on the right.

Here is an example: the sequence $(1, 6, 2, 3, 7, 8, 4, 5)$ has two rising sequences: $(1, 2, 3, 4, 5)$ and $(6, 7, 8)$. In fact, you should be able to convince yourself that every finite sequence s of integers has some well-defined number $R(s)$ of rising sequences. Can you figure out a procedure to compute $R(s)$? I am going to let you think on that.

We will using rising sequences to prove that RIFFLE-5 is deficient. First consider the ordering of cards $\sigma = (1, 2, \ldots, 52)$. How many rising sequences does it have? Just 1. What about the sequence $\tau = (52, 51, 50 \ldots, 3, 2, 1)$? It has 52 rising sequences.

I claim that if we riffle shuffle the deck σ then the resulting deck, which previously had 1 rising sequence, will now have at most 2 rising sequences. If we riffle shuffle what results one more time—so now we've done 2 riffle times starting from σ—then the resulting deck will have at most 4 rising sequences. Riffle shuffle 3 times starting from σ and you'll get at most 8 rising sequences. Riffle shuffle 4 times starting from σ and you'll get at most 16 rising sequences. Riffle shuffle 5 times starting from σ and you'll get at most 32 rising sequences. In particular, with 5 riffle shuffles, starting from σ, there's no way you'll get ordering τ. Thus MIX-5 is deficient. It's not a good way to mix a deck of cards.

OK, but why the at-most-doubling behavior? I am claiming that s' is obtained from s by a single riffle shuffle, then the number of rising sequences in s' is at most twice the number of rising sequences in s: in symbols, $R(s') \leq 2R(s)$. For consider some rising sequence $r = (a, a+1, \ldots, a+k)$ in s. If the cut of the deck occurs after card $a+i-1$ but on or before card $a+i$, then, after the riffle shuffle the deck s' will include subsequences $(a, \ldots, a+i-1)$ and $(a+i, \ldots, a+k)$. Now these subsequences might not be maximal—it is possible that we can extend one or both of them. But even if we can't extend *any* of these still we have identified $2R(s)$ subsequences of consecutive values in s' that, together, include all the cards of s. Thus $R(s') \leq 2R(s)$.

> In 1992, Dave Bayer and Persi Diaconis showed that after seven shuffles of a deck of 52 cards, every sequence of cards is nearly equally likely. It's a quantitative results: with seven riffle shuffles you pass a threshold of 0.5 in a standard measure of how far your are from a perfectly randomized deck of cards.
>
> In general, the authors show that you need about $1.5 \lg n$ riffle shuffles to well-shuffle a deck of n cards (in the sense of getting the "total-variation distance to uniform" drop under 1/2 at that point). By "lg" we denote log base-2.
>
> Here's a chart from the Bayer-Diaconis paper for how well 52 cards are mixed after 1 to 10 riffle shuffles. All numbers are rounded to two decimal places. An exact value of 0 would mean perfect mixing (all orderings occur with probability 1/52!); an exact value of 1 would mean that no mixing has taken place—as when the initial distribution s remains the only one reachable after the mixing.

Num of Shuffles	1	2	3	4	5	6	7	8	9	10
Dist to uniform	1.00	1.00	1.00	1.00	0.92	0.61	0.33	0.17	0.08	0.04