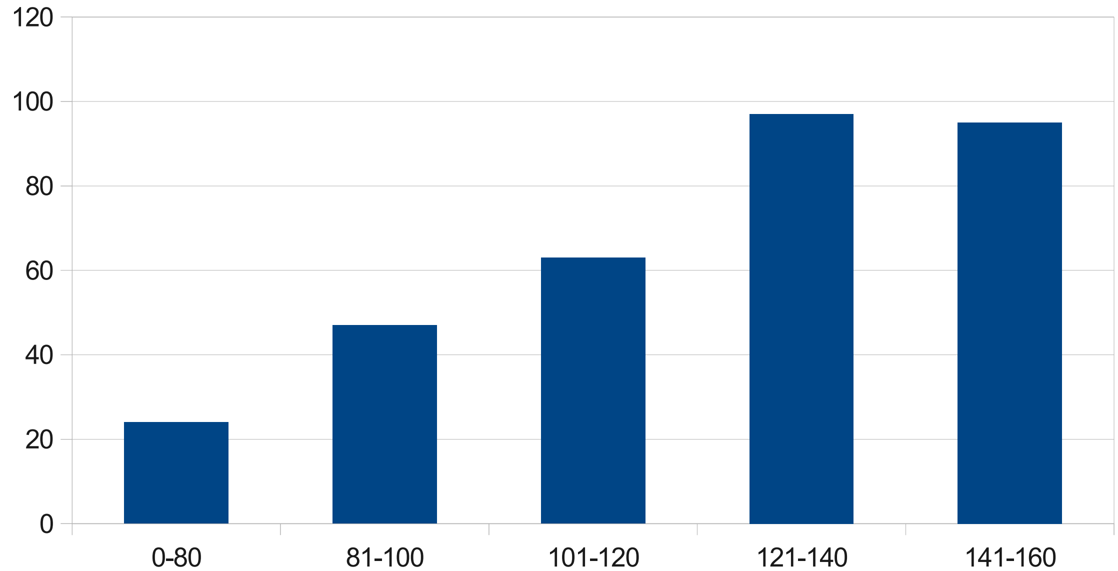
CS103 Midterm Exam Solutions

The distribution of the exam grades was as follows:



Overall, the final statistics were as follows:

75th Percentile: 144 / 160 (90%)

50th Percentile: 127 / 160 (79%)

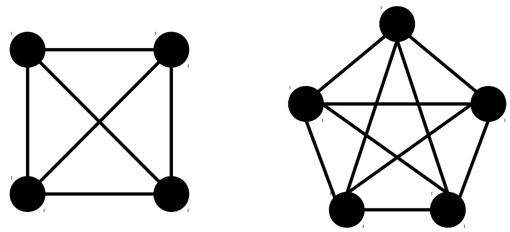
25th Percentile: 104 / 160 (65%)

We are **not** grading this course using raw point totals and will instead be grading on a (fairly generous) curve. Roughly speaking, the median score is the cutoff between a B and a B+, and anything at 75th percentile or higher should translate to a solid A. As always, if you have any comments or questions about the midterm or your grade on the exam, please don't hesitate to drop by office hours! You can also email the staff list with questions.

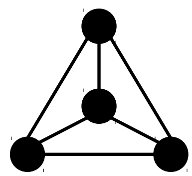
If you think that we made any mistakes in our grading, please feel free to submit a regrade request to us according to the procedure outlined in the Problem Set Policies handout. **All regrade requests must be received by Monday, November 11th at 3:30PM.**

Problem One: Graphs and Relations (40 Points Total)

For any $k \geq 1$, a *k-clique* is an undirected graph $G = (V, E)$ where there are k nodes ($|V| = k$) and where there are edges between each distinct pair of nodes ($E = \{\{u, v\} \mid u, v \in V \text{ and } u \neq v\}$). For example, here is a 4-clique and a 5-clique:



Although in the above drawing of the 4-clique there are two crossing edges, it's possible to draw the 4-clique without any edges crossing, as shown below.



This means that the 4-clique is a planar graph. However, there is no corresponding way to draw the 5-clique without at least two edges crossing.

(i) Clique Drawings (15 Points)

Using the Four-Color Theorem, prove that the 5-clique is not a planar graph.

Proof: By contradiction; assume the 5-clique is a planar graph. By the Four-Color Theorem, we know the 5-clique is 4-colorable. Consider any four-coloring of the 5-clique. Since there are five nodes and only four colors, by the pigeonhole principle we know that two of the nodes (call them u and v) must be assigned the same color. But $\{u, v\}$ is an edge in the 5-clique, so there are two nodes connected by an edge that are the same color as one another, which is impossible because we began with a 4-coloring of the 5-clique. We have reached a contradiction, so our assumption must have been wrong. Thus the 5-clique is not a planar graph. ■

Why we asked this question: It seems like cheating, but it's completely legitimate to harness heavyweight theorems in normal mathematics. Proving that the 5-clique isn't planar is messy and difficult, but if you accept that the four-color theorem is true, it's really not so bad!

Common mistakes: Many solutions correctly claimed that for the 5-clique to be planar, every node would have to have a different color, but did so without providing any justification. Similarly, many solutions worked by fixing a node and claiming the remaining 4-clique would have to have each node colored differently, but didn't explain why this would be true.

(ii) Triangular Relations (25 Points)

A binary relation R over a set A is called *triangular* iff for all $x, y, z \in A$, if xRy and xRz, then yRz. For example, the $\equiv_k$ relation is triangular over $\mathbb{Z}$, but the $\leq$ relation over $\mathbb{N}$ is not triangular.

Let R be an arbitrary binary relation over some set A. Prove that R is an equivalence relation if and only if it is reflexive and triangular.

Proof: We prove both directions of implication.

($\Rightarrow$) First, we prove that if R is an equivalence relation, R is reflexive and triangular. Since R is an equivalence relation, we know that it is reflexive, so we just need to prove that it is triangular. To do this, consider any $x, y, z \in A$ such that xRy and xRz. We need to prove yRz. Since R is an equivalence relation, R is symmetric, so from since xRy we get yRx. Since R is an equivalence relation, R is transitive, so from yRx and xRz we get yRz. Thus yRz holds, and R is triangular.

($\Leftarrow$) Next, we prove that if R is reflexive and triangular, then it is an equivalence relation. To do this, we must prove R is reflexive, symmetric, and transitive. Since R is already assumed to be reflexive, this means we just need to show symmetry and transitivity.

To prove that R is symmetric, consider any $x, y \in A$ where xRy. We need to prove yRx. Since R is reflexive, we know that xRx holds. Therefore, since xRy and xRx, by triangularity we know yRx.

To prove that R is transitive, consider any $x, y, z \in A$ where xRy and yRz. We need to prove xRz. Since we've proven R is symmetric and xRy, we know that yRx. Therefore, since yRx and yRz, by triangularity we have that xRz, as required. ■

Why we asked this question: This question is an attractive exam question for several reasons. First, the statement we ask you to prove is a biconditional, which means that you need to prove two separate directions of implication. Second, the proof of transitivity from triangularity and reflexivity builds off of the earlier proof of symmetry, which we thought would be an indication that you were on the right track as you wrote the proof.

Common mistakes: The statement to prove is a biconditional, and unfortunately some solutions only proved one direction of implication. Additionally, many proofs claimed that triangularity and transitivity were identical to one another, which isn't the case (look carefully at the definitions of those properties).

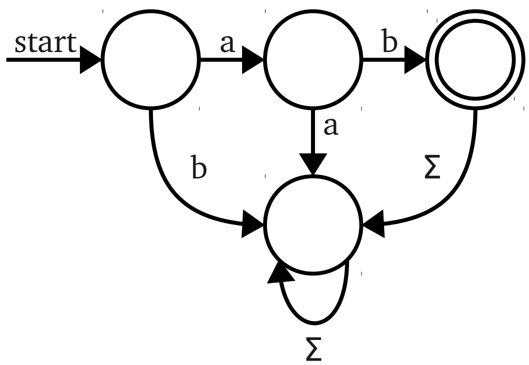
A particularly common mistake was to assume the wrong relations held when trying to prove transitivity from triangularity or vice-versa. For example, to show triangularity, a proof should assume that xRy and xRz and try to prove yRz. Similarly, to show transitivity, a proof should assume that xRy and yRz and try to prove xRz.

One logical error we saw arose when trying to prove symmetry from triangularity. Although triangularity means that $xRy \land xRz \rightarrow yRz$ and that $xRz \land xRy \rightarrow zRy$, these statements alone do not prove that yRz iff zRy, since both statements are predicated on the assumption xRy and xRz, which aren't necessarily true for arbitrary x, y, and z. Finally, many approaches noted that since all equivalence relations are reflexive, it would suffice to prove that any relation that is symmetric and transitive must be triangular and vice-versa. This unfortunately isn't true. Consider the set $A = \{0, 1\}$ where $0R1$ but no other values are related by R. This relation is triangular (do you see why?), but it's not symmetric. It's necessary to assume reflexivity to show that R must be an equivalence relation.

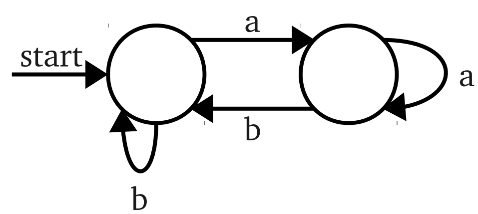
Problem Two: Finite Automata (20 Points)

Below are four DFAs over the alphabet Σ = {**a**, **b**} and seven descriptions of languages over Σ. For each automaton, decide whether the language of that automaton is one of the numbered languages. If so, write that number on the line beneath that automaton. If not, write "none."

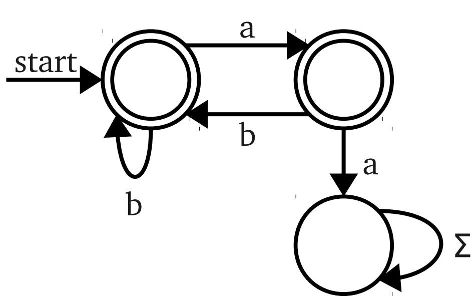
Note that there may be multiple automata with the same language, and some languages might not be used at all.



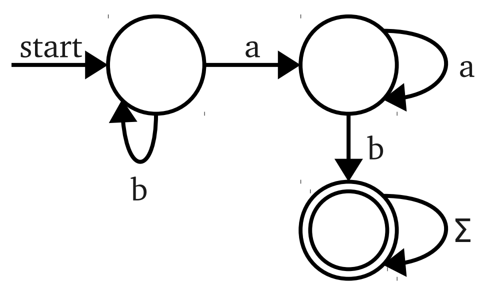
Language: 6



Language: None



Language: 2



Language: 5

Why we asked this question: This question was designed to make sure that you could look at an automaton and come up with a simple explanation for its language. We also wanted to make sure you were comfortable reading set-based definitions of languages.

Common mistakes: Many answers listed multiple languages for the same automaton. Every automaton only has one language, which corresponds to the set of all strings it accepts. If an automaton accepts all strings in a set *S*, but also accepts some strings not in *S*, then *S* is not the language of that automaton. Therefore, it was automatically incorrect to list multiple different languages for the same automaton.

Problem Three: Functions and Cardinality (55 points)

(i) Properties of Finite Sets (20 Points)

Let A and B be finite sets where $|A| = |B|$. This means that there is some natural number n such that $|A| = |B| = n$.

Let $f : A \to B$ be an arbitrary function from A to B. Prove, by contradiction, that if f is an injection, then f is a bijection.

Proof: By contradiction; suppose that A and B are finite sets where $|A| = |B|$ and that $f : A \to B$ is an injection but not a bijection. Since f is an injection but not a bijection, we know that f must not be surjective. Therefore, there must be some $b \in B$ such that $f(a) \neq b$ for any $a \in A$. This means that there are at most $|B|$ - 1 possible values for $f(a)$, since b is never mapped to. However, there are $|A| = |B|$ possible values of a, so by the pigeonhole principle there must be two values a_1, a_2 such that $f(a_1) = f(a_2)$ but $a_1 \neq a_2$. But this is impossible, since f is injective. We have reached a contradiction, so our assumption must have been wrong. Thus f is a bijection. ■

Why we asked this question: This question was designed to make sure you were comfortable manipulating injections, surjections, and bijections. You need to call down to the definitions of these types of functions in order to make the proof work. Additionally, the pigeonhole principle ends up making an appearance, even though it's not immediately obvious that this is the case.

Common mistakes: One of the more common mistakes we saw on this problem was negating the statement "if f is an injection, then f is a bijection" and getting "if f is an injection, then f is not a bijection." Remember: the negation of the statement $P \to Q$ is not the implication $P \to \neg Q$; it's the statement $P \land \neg Q$.

One of the trickier parts of this problem was using the fact that A and B are specifically *finite* sets to show why f must be a bijection. To be correct, a proof needs to rely on the fact that A and B don't contain infinitely many elements, since in that case it's possible that f is an injection but not a bijection. As an example, the proof approach we outlined above uses the fact that A and B are finite by noting that there are $n - 1$ covered elements of B, which forces some element to be covered twice.

The most common mistakes we saw were failing to leverage finiteness in the proof to explain why f necessarily must be a bijection. For example, the fact that if f is injective does indeed mean that $|A|$ elements of B must be covered by f, but that in of itself doesn't imply that all elements of B are covered unless B is finite (for example, let $f : \mathbb{N} \to \mathbb{Z}$ be the function $f(n) = n$).

When we proved that $|\mathbb{N}| \neq |\mathbb{R}|$, we did so by assuming there was some bijection $f : \mathbb{N} \to \mathbb{R}$ and then constructing a diagonal number d that was different from $f(n)$ for any $n \in \mathbb{N}$. We defined this diagonal number d by giving its decimal representation as follows:

$$d[n] = \begin{cases} 1 & \text{if } f(n)[n] = 0 \\ 0 & \text{otherwise} \end{cases}$$

As a reminder, the notation $r[0]$ refers to the integer part of r and $r[n]$ refers to the nth decimal digit of r for any $n > 0$.

Let $\mathbb{Q}$ be the set of all rational numbers. Interestingly, $|\mathbb{N}| = |\mathbb{Q}|$, meaning that there is at least one bijection from $\mathbb{N}$ to $\mathbb{Q}$.

(ii) Constructing Irrational Numbers (15 Points)

Prove that for any bijection $f : \mathbb{N} \to \mathbb{Q}$, the real number d defined as

$$d[n] = \begin{cases} 1 & \text{if } f(n)[n] = 0 \\ 0 & \text{otherwise} \end{cases}$$

is irrational. *(Hint: Although the proofs of irrationality we've done so far have worked by writing $d = p / q$ for some integers p and q and deriving a contradiction, for this problem we recommend that you use a different proof strategy.)*

Proof: Let d be defined as above and assume for the sake of contradiction that d is rational. Since $f : \mathbb{N} \to \mathbb{Q}$ is a bijection, it is surjective, so there must be some $n \in \mathbb{N}$ such that $f(n) = d$. We consider two cases:

Case 1: $f(n)[n] = 0$. Then $d[n] = 1$, so $f(n) \neq d$.

Case 2: $f(n)[n] \neq 0$. Then $d[n] = 0$, so $f(n) \neq d$.

In both cases we find $f(n) \neq d$, contradicting the fact that $f(n) = d$. We have reached a contradiction, so our assumption must have been wrong. Thus d, as defined above, must be irrational. ■

Why we asked this question: We've seen diagonalization used to prove that two sets don't have the same cardinality, but it can be used for multiple other purposes as well. In fact, this shows that if you can find any bijection between $\mathbb{N}$ and $\mathbb{Q}$, you can use it to discover an irrational number.

Common mistakes: One common mistake was claiming that all rational numbers must have decimal representations that terminate in infinitely many zeros, meaning that d must terminate with infinitely many 1s and thus can't be rational. This isn't true, though; the number $^1/_3 = 0.333...$ terminates with infinitely many 3's. Another common error was correctly stating that d wasn't mapped to by f, but not explaining why this was the case.

Suppose that $f : \mathbb{N} \to \mathbb{R}$ is a function from natural numbers to real numbers. We can use limits to formally define what it means for $f(n)$ to approach 0 as n goes to infinity.

The statement

$$\lim_{n \to \infty} f(n) = 0$$

is, by definition, equivalent to the following statement:

For any positive real number ε, there is a natural number n_0 such that for every natural number $n \geq n_0$, the inequality $|f(n)| \leq \varepsilon$ holds.

(iii) To The Limit **(15 Points)**

Given the predicates

Real(x), which states that x is a real number,

Natural(n), which states that n is a natural number, and

$x \leq y$, which states that x is less than y;

the functions

$f(x)$, which represents the value of the function f given input x, and

$|x|$, which represents the absolute value of x;

and the constant 0, write a statement in first-order logic that says $\lim_{n \to \infty} f(n) = 0$. You should only use the functions, predicates, and constants defined above.

Here is one possible solution:

$$\begin{aligned}
&\forall \varepsilon. \, (Real(\varepsilon) \land \varepsilon \neq 0 \land \varepsilon \geq 0 \to \\
&\qquad \exists n_0. \, (Natural(n) \land \\
&\qquad\qquad \forall n. \, (Natural(n) \land n \geq n_0 \to |f(n)| \leq \varepsilon) \\
&\qquad) \\
&)
\end{aligned}$$

Why we asked this question: We asked this question for a few reasons. First, we wanted to show that there is *yet another* definition of infinity – the infinity that appears in limits – that's totally different from the infinities we've seen elsewhere. Second, this problem is a good exercise in structuring first-order logic statements. It's important to determine where to use ∧ and where to use →, as well as how to express concepts like "ε is positive" in terms of simpler operators. Third, this problem tested whether you remembered that = is a part of first-order logic, which is useful (but not necessary) in translating the statement "ε is positive."

Common mistakes: The most common mistake on this problem was incorrectly translating the statement "ε is positive." The statement $\varepsilon \geq 0$ only shows that ε is nonnegative, which isn't sufficient. Additionally, many answers incorrectly mixed up → and ∧ or used quantifiers incorrectly.

Problem Four: Graphs and Relations (45 Points Total)

Prove, by induction on *n*, that

$$\sum_{i=0}^{n} \frac{i}{2^i} = 2 - \frac{n+2}{2^n}$$

Proof: By induction. Let $P(n)$ be defined as

$$P(n) \equiv \text{“} \sum_{i=0}^{n} \frac{i}{2^i} = 2 - \frac{n+2}{2^n} \text{”}$$

We will prove $P(n)$ holds for all $n \in \mathbb{N}$. As our base case, we will prove $P(0)$, that

$$\sum_{i=0}^{0} \frac{i}{2^i} = 2 - \frac{0+2}{2^0}$$

Note that

$$\sum_{i=0}^{0} \frac{i}{2^i} = \frac{0}{2^0} = 0 = 2 - 2 = 2 - \frac{0+2}{2^0}$$

Thus $P(0)$ holds. For the inductive step, assume that for some $n \in \mathbb{N}$ that $P(n)$ holds and that

$$\sum_{i=0}^{n} \frac{i}{2^i} = 2 - \frac{n+2}{2^n}$$

We will prove $P(n+1)$, that

$$\sum_{i=0}^{n+1} \frac{i}{2^i} = 2 - \frac{n+3}{2^{n+1}}$$

To see this, note that

$$\sum_{i=0}^{n+1} \frac{i}{2^i} = \sum_{i=0}^{n} \frac{i}{2^i} + \frac{n+1}{2^{n+1}} = 2 - \frac{n+2}{2^n} + \frac{n+1}{2^{n+1}} = 2 - \frac{2(n+2)}{2^{n+1}} + \frac{n+1}{2^{n+1}} = 2 - \frac{2n+4-n-1}{2^{n+1}} = 2 - \frac{n+3}{2^{n+1}}$$

Thus $P(n+1)$ holds, completing the induction. ■

Why we asked this question: This question is the only question on the exam that tests equational reasoning (i.e. manipulating equalities to get a desired result). We wanted to see if you were able to manipulate an unusual summation in the context of a proof by induction. Additionally, this particular summation comes up in many algorithmic contexts, and in case you were interested in continuing onward in CS (especially CS161) we wanted to help prepare you!

Common mistakes: Most people got this one right! The most common mistake we saw was justifying the base case incorrectly, either by assuming that the sum was the empty sum (it's not; it's a sum of one element, which is zero) or by assuming the equality held and working backwards. A common error in the inductive step was to assume that the equality suggested by $P(n+1)$ was true and working backwards to arrive at the equality for $P(n)$, which isn't valid reasoning.

Problem Five: The Pigeonhole Principle (30 Points)

In Problem Set Three, you proved that any group of 6 people will have a group of three mutual acquaintances, a group of three mutual strangers, or both.

It's possible to rephrase this in graph theoretic terms. Suppose you take a 6-clique (recall from Problem Four that a k-clique is a graph of k nodes where every pair of nodes has an edge between them) and color each edge red or blue. The result you proved is equivalent to the following:

Theorem: In a 6-clique where each edge is colored either red or blue, there exists a blue triangle or a red triangle.

Here, a "triangle" refers to a collection of three nodes that are all joined together by edges of the same color. To see why this is equivalent to the original problem, take any group of six people and draw a blue edge between them if they're acquaintances and a red edge between them if they're strangers. Now, a group of "three mutual acquaintances" is a blue triangle and a group of "three mutual strangers" is a red triangle.

Suppose you have a 17-clique (that is, an undirected graph with 17 nodes where there's an edge between each pair of nodes) where each edge is colored one of three different colors (say, red, green, and blue). Prove that there must be a blue triangle, a red triangle, or a green triangle.

(Hint: As with the question on Problem Set Three, start off by choosing some node out of the group and looking at the colors of the edges it's connected to. We strongly recommend using the theorem above – which you've already proven in Problem Set Three – in your proof.)

Proof: Consider any 17-clique where the edges are colored red, green, and blue. Choose any node v and look at the 16 edges connected to it. By the pigeonhole principle, since there are three colors and 16 edges, there must be some color for which there are at least $\lceil 16 / 3 \rceil = 6$ edges of that color connected to v. Without loss of generality, let that color be green.

Look at some collection of 6 nodes connected by green edges to v. If any of the edges between those nodes are green, then we have a green triangle formed by taking v and any two nodes in the 6 that are connected by a green edge. Otherwise, all the edges in the group of six are either red or blue, so by the earlier theorem we know there must be a red triangle or a blue triangle in that group. ■

Why we asked this question: Question 5.ii on Problem Set Three was one of the trickier pigeonhole principle proofs we covered. In this problem, we wanted you to see how you could use the same reasoning to solve this more complicated problem.

Common mistakes: Most of the proofs we saw were either completely correct or definitely on the right track. Common mistakes included claiming that there must be *exactly* six edges of some color connected to a node (rather than six or more) or claiming that in any group of six nodes where edges are colored red or blue there must be a blue triangle or red triangle, rather than that in any *6-clique* there must be a red triangle or blue triangle. Another mistake we saw was finding the set of 6 nodes connected to the node singled out in the first step and incorrectly claiming that it was necessary that only two colors of edges would link them together.