In-Class Problems Week 1, Wed.

Problem 1.

The Pythagorean Theorem says that if a and b are the lengths of the sides of a right triangle, and c is the length of its hypotenuse, then

$$a^2 + b^2 = c^2.$$

This theorem is so fundamental and familiar that we generally take it for granted. But just being familiar doesn't justify calling it "obvious"—witness the fact that people have felt the need to devise different proofs of it for milllenia.[1] In this problem we'll examine a particularly simple "proof without words" of the theorem.

Here's the strategy. Suppose you are given four different colored copies of a right triangle with sides of lengths a, b, and c, along with a suitably sized square, as shown in Figure 1.

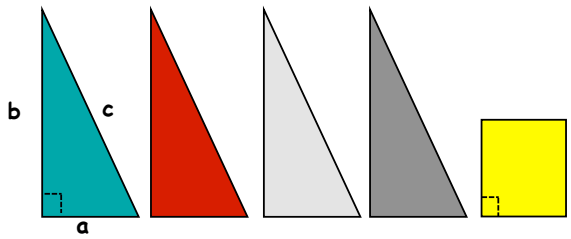


Figure 1 Right triangles and square.

(a) You will first arrange the square and four triangles so they form a $c \times c$ square. From this arrangement you will see that the square is $(b-a) \times (b-a)$.

(b) You will then arrange the same shapes so they form two squares, one $a \times a$ and the other $b \times b$.

You know that the area of an $s \times s$ square is s^2. So appealing to the principle that

Area is Preserved by Rearranging,

you can now conclude that $a^2 + b^2 = c^2$, as claimed.

This really is an elegant and convincing proof of the Pythagorean Theorem, but it has some worrisome features. One concern is that there might be something special about the shape of these particular triangles and square that makes the rearranging possible—for example, suppose $a = b$?

(c) How would you respond to this concern?

[1] Over a hundred different proofs are listed on the mathematics website http://www.cut-the-knot.org/pythagoras/.

(d) Another concern is that a number of facts about right triangles, squares and lines are being *implicitly* assumed in justifying the rearrangements into squares. Enumerate some of these assumed facts.

Problem 2.
What's going on here?!

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = \left(\sqrt{-1}\right)^2 = -1.$$

(a) Precisely identify and explain the mistake(s) in this *bogus* proof.

(b) Prove (correctly) that if $1 = -1$, then $2 = 1$.

(c) Every *positive* real number, r, has two square roots, one positive and the other negative. The standard convention is that the expression $\sqrt{r}$ refers to the *positive* square root of r. Assuming familiar properties of multiplication of real numbers, prove that for positive real numbers r and s,

$$\sqrt{rs} = \sqrt{r}\sqrt{s}.$$

Problem 3.
Identify exactly where the bugs are in each of the following bogus proofs.[2]

(a) Bogus Claim: $1/8 > 1/4$.

Bogus proof.

$$\begin{aligned} 3 &> 2 \\ 3\log_{10}(1/2) &> 2\log_{10}(1/2) \\ \log_{10}(1/2)^3 &> \log_{10}(1/2)^2 \\ (1/2)^3 &> (1/2)^2, \end{aligned}$$

and the claim now follows by the rules for multiplying fractions. ■

(b) *Bogus proof*: $1¢ = \$0.01 = (\$0.1)^2 = (10¢)^2 = 100¢ = \$1.$ ■

(c) Bogus Claim: If a and b are two equal real numbers, then $a = 0$.

Bogus proof.

$$\begin{aligned} a &= b \\ a^2 &= ab \\ a^2 - b^2 &= ab - b^2 \\ (a-b)(a+b) &= (a-b)b \\ a+b &= b \\ a &= 0. \end{aligned}$$

■

[2] From [42], *Twenty Years Before the Blackboard* by Michael Stueben and Diane Sandford in the course textbook

Problem 4.
It's a fact that the Arithmetic Mean is at least as large as the Geometric Mean, namely,

$$\frac{a+b}{2} \geq \sqrt{ab}$$

for all nonnegative real numbers a and b. But there's something objectionable about the following proof of this fact. What's the objection, and how would you fix it?

Bogus proof.

$$\begin{aligned}
\frac{a+b}{2} &\stackrel{?}{\geq} \sqrt{ab}, && \text{so}\\
a+b &\stackrel{?}{\geq} 2\sqrt{ab}, && \text{so}\\
a^2+2ab+b^2 &\stackrel{?}{\geq} 4ab, && \text{so}\\
a^2-2ab+b^2 &\stackrel{?}{\geq} 0, && \text{so}\\
(a-b)^2 &\geq 0 && \text{which we know is true.}
\end{aligned}$$

The last statement is true because $a-b$ is a real number, and the square of a real number is never negative. This proves the claim. ■

Optional (and controversial)

Problem 5.
Albert announces to his class that he plans to surprise them with a quiz sometime next week.

His students first wonder if the quiz could be on Friday of next week. They reason that it can't: if Albert didn't give the quiz *before* Friday, then by midnight Thursday, they would know the quiz had to be on Friday, and so the quiz wouldn't be a surprise any more.

Next the students wonder whether Albert could give the surprise quiz Thursday. They observe that if the quiz wasn't given *before* Thursday, it would have to be given *on* the Thursday, since they already know it can't be given on Friday. But having figured that out, it wouldn't be a surprise if the quiz was on Thursday either. Similarly, the students reason that the quiz can't be on Wednesday, Tuesday, or Monday. Namely, it's impossible for Albert to give a surprise quiz next week. All the students now relax, having concluded that Albert must have been bluffing. And since no one expects the quiz, that's why, when Albert gives it on Tuesday next week, it really is a surprise!

What, if anything, do you think is wrong with the students' reasoning?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 1, Fri.

Problem 1.
Prove that if $a \cdot b = n$, then either a or b must be $\leq \sqrt{n}$, where a, b, and n are nonnegative real numbers.
Hint: by contradiction, Section 1.8 in the course textbook.

Problem 2.
Generalize the proof of Theorem 1.8.1 repeated below that $\sqrt{2}$ is irrational[1] in the course textbook.
For example, how about $\sqrt{3}$?

> **Theorem.** *$\sqrt{2}$ is an irrational number.*
>
> *Proof.* The proof is by contradiction: assume that $\sqrt{2}$ is rational, that is,
>
> $$\sqrt{2} = \frac{n}{d}, \tag{1}$$
>
> where n and d are integers. Now consider the smallest such positive integer denominator, d. We will prove in a moment that the numerator, n, and the denominator, d, are both even. This implies that
>
> $$\frac{n/2}{d/2}$$
>
> is a fraction equal to $\sqrt{2}$ with a smaller positive integer denominator, a contradiction.
>
> > *Since the assumption that $\sqrt{2}$ is rational leads to this contradiction, the assumption must be false. That is, $\sqrt{2}$ is indeed irrational.* This italicized comment on the implication of the contradiction normally goes without saying, but since this is an early example of proof by contradiction, we've said it.
>
> To prove that n and d have 2 as a common factor, we start by squaring both sides of (1) and get $2 = n^2/d^2$, so
>
> $$2d^2 = n^2. \tag{2}$$
>
> So 2 is a factor of n^2, which is only possible if 2 is in fact a factor of n.
>
> This means that $n = 2k$ for some integer, k, so
>
> $$n^2 = (2k)^2 = 4k^2. \tag{3}$$
>
> Combining (2) and (3) gives $2d^2 = 4k^2$, so
>
> $$d^2 = 2k^2. \tag{4}$$
>
> So 2 is a factor of d^2, which again is only possible if 2 is in fact also a factor of d, as claimed. ■

[1] Remember that an irrational number is a number that cannot be expressed as a ratio of two integers.

Problem 3.
If we raise an irrational number to an irrational power, can the result be rational? Show that it can by considering $\sqrt{2}^{\sqrt{2}}$ and arguing by cases.

Problem 4.
The fact that that there are irrational numbers a, b such that a^b is rational was proved earlier by cases. Unfortunately, that proof was *nonconstructive*: it didn't reveal a specific pair, a, b, with this property. But in fact, it's easy to do this: let $a ::= \sqrt{2}$ and $b ::= 2\log_2 3$.

We know $a = \sqrt{2}$ is irrational, and $a^b = 3$ by definition. Finish the proof that these values for a, b work, by showing that $2\log_2 3$ is irrational.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 2, Wed.

Problem 1.
The proof below uses the Well Ordering Principle to prove that every amount of postage that can be assembled using only 6 cent and 15 cent stamps, is divisible by 3. Let the notation "$j \mid k$" indicate that integer j is a divisor of integer k, and let $S(n)$ mean that exactly n cents postage can be assembled using only 6 and 15 cent stamps. Then the proof shows that

$$S(n) \text{ IMPLIES } 3 \mid n, \quad \text{for all nonnegative integers } n. \tag{1}$$

Fill in the missing portions (indicated by "...") of the following proof of (1).

Let C be the set of *counterexamples* to (1), namely[1]

$$C ::= \{n \mid \ldots\}$$

Assume for the purpose of obtaining a contradiction that C is nonempty. Then by the WOP, there is a smallest number, $m \in C$. This m must be positive because....

But if $S(m)$ holds and m is positive, then $S(m-6)$ or $S(m-15)$ must hold, because.....

So suppose $S(m-6)$ holds. Then $3 \mid (m-6)$, because....

But if $3 \mid (m-6)$, then $3 \mid m$, because...,

contradicting the fact that m is a counterexample.

Next, if $S(m-15)$ holds, we arrive at a contradiction in the same way. Since we get a contradiction in both cases, we conclude that...

which proves that (1) holds.

[1] The notation "$\{n \mid \ldots\}$" means "the set of elements, n, such that ...."

Problem 2.
Use the *Well Ordering Principle* [2] to prove that

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}. \tag{2}$$

for all nonnegative integers, n.

Problem 3.
Euler's Conjecture in 1769 was that there are no positive integer solutions to the equation

$$a^4 + b^4 + c^4 = d^4.$$

Integer values for a, b, c, d that do satisfy this equation were first discovered in 1986. So Euler guessed wrong, but it took more than two centuries to demonstrate his mistake.

Now let's consider Lehman's equation, similar to Euler's but with some coefficients:

$$8a^4 + 4b^4 + 2c^4 = d^4 \tag{3}$$

Prove that Lehman's equation (3) really does not have any positive integer solutions.

Hint: Consider the minimum value of a among all possible solutions to (3).

Problem 4.
You are given a series of envelopes, respectively containing $1, 2, 4, \ldots, 2^m$ dollars. Define

> **Property** m: For any nonnegative integer less than 2^{m+1}, there is a selection of envelopes whose contents add up to *exactly* that number of dollars.

Use the Well Ordering Principle (WOP) to prove that Property m holds for all nonnegative integers m.

Hint: Consider two cases: first, when the target number of dollars is less than 2^m and second, when the target is at least 2^m.

Problem 5.
Use the Well Ordering Principle to prove that any integer greater than or equal to 30 can be represented as the sum of nonnegative integer multiples of 6, 10, and 15.

Hint: Use the template for WOP proofs to ensure partial credit. Verify that integers in the interval [30..35] are sums of nonnegative integer multiples of 6, 10, and 15.

[2] Proofs by other methods such as induction or by appeal to known formulas for similar sums will not receive credit.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 2, Fri.

Problem 1.
Prove by truth table that OR distributes over AND, namely,

$$P \text{ OR } (Q \text{ AND } R) \quad \text{is equivalent to} \quad (P \text{ OR } Q) \text{ AND } (P \text{ OR } R) \tag{1}$$

Problem 2.
This problem[1] examines whether the following specifications are *satisfiable*:

1. If the file system is not locked, then
   (a) new messages will be queued.
   (b) new messages will be sent to the messages buffer.
   (c) the system is functioning normally, and conversely, if the system is functioning normally, then the file system is not locked.
2. If new messages are not queued, then they will be sent to the messages buffer.
3. New messages will not be sent to the message buffer.

(a) Begin by translating the five specifications into propositional formulas using four propositional variables:

$$\begin{aligned} L &::= \text{file system locked,} \\ Q &::= \text{new messages are queued,} \\ B &::= \text{new messages are sent to the message buffer,} \\ N &::= \text{system functioning normally.} \end{aligned}$$

(b) Demonstrate that this set of specifications is satisfiable by describing a single truth assignment for the variables L, Q, B, N and verifying that under this assignment, all the specifications are true.

(c) Argue that the assignment determined in part (b) is the only one that does the job.

Problem 3.
Propositional logic comes up in digital circuit design using the convention that **T** corresponds to 1 and **F** to 0. A simple example is a 2-bit *half-adder* circuit. This circuit has 3 binary inputs, a_1, a_0 and b, and 3 binary outputs, c, s_1, s_0. The 2-bit word a_1a_0 gives the binary representation of an integer, k, between 0 and 3. The

[1] Revised from Rosen, 5th edition, Exercise 1.1.36

3-bit word cs_1s_0 gives the binary representation of $k + b$. The third output bit, c, is called the final *carry bit*.

So if k and b were both 1, then the value of a_1a_0 would be `01` and the value of the output cs_1s_0 would `010`, namely, the 3-bit binary representation of $1 + 1$.

In fact, the final carry bit equals 1 only when all three binary inputs are 1, that is, when $k = 3$ and $b = 1$. In that case, the value of cs_1s_0 is `100`, namely, the binary representation of $3 + 1$.

This 2-bit half-adder could be described by the following formulas:

$$\begin{aligned} c_0 &= b \\ s_0 &= a_0 \text{ XOR } c_0 \\ c_1 &= a_0 \text{ AND } c_0 && \text{the carry into column 1} \\ s_1 &= a_1 \text{ XOR } c_1 \\ c_2 &= a_1 \text{ AND } c_1 && \text{the carry into column 2} \\ c &= c_2. \end{aligned}$$

(a) Generalize the above construction of a 2-bit half-adder to an $n+1$ bit half-adder with inputs $a_n, \ldots, a_1, a_0$ and b and outputs $c, s_n, \ldots, s_1, s_0$. That is, give simple formulas for s_i and c_i for $0 \leq i \leq n + 1$, where c_i is the carry into column $i + 1$, and $c = c_{n+1}$.

(b) Write similar definitions for the digits and carries in the sum of two $n+1$-bit binary numbers $a_n \ldots a_1a_0$ and $b_n \ldots b_1b_0$.

Visualized as digital circuits, the above adders consist of a sequence of single-digit half-adders or adders strung together in series. These circuits mimic ordinary pencil-and-paper addition, where a carry into a column is calculated directly from the carry into the previous column, and the carries have to ripple across all the columns before the carry into the final column is determined. Circuits with this design are called *ripple-carry* adders. Ripple-carry adders are easy to understand and remember and require a nearly minimal number of operations. But the higher-order output bits and the final carry take time proportional to n to reach their final values.

(c) How many of each of the propositional operations does your adder from part (b) use to calculate the sum?

Problem 4.

When the mathematician says to his student, "If a function is not continuous, then it is not differentiable," then letting D stand for "differentiable" and C for continuous, the only proper translation of the mathematician's statement would be

$$\text{NOT}(C) \text{ IMPLIES } \text{NOT}(D),$$

or equivalently,

$$D \text{ IMPLIES } C.$$

But when a mother says to her son, "If you don't do your homework, then you can't watch TV," then letting T stand for "can watch TV" and H for "do your homework," a reasonable translation of the mother's statement would be

$$\text{NOT}(H) \text{ IFF } \text{NOT}(T),$$

or equivalently,

$$H \text{ IFF } T.$$

Explain why it is reasonable to translate these two IF-THEN statements in different ways into propositional formulas.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 3, Tue.

Problem 1.
For each of the logical formulas, indicate whether or not it is true when the domain of discourse is $\mathbb{N}$, (the nonnegative integers 0, 1, 2, ...), $\mathbb{Z}$ (the integers), $\mathbb{Q}$ (the rationals), $\mathbb{R}$ (the real numbers), and $\mathbb{C}$ (the complex numbers). Add a brief explanation to the few cases that merit one.

$$\begin{aligned}
\exists x.\, x^2 &= 2 \\
\forall x . \exists y.\, x^2 &= y \\
\forall y . \exists x.\, x^2 &= y \\
\forall x \neq 0 . \exists y.\, xy &= 1 \\
\exists x . \exists y.\, x + 2y = 2 \ &\text{AND}\ 2x + 4y = 5
\end{aligned}$$

Problem 2.
The goal of this problem is to translate some assertions about binary strings into logic notation. The domain of discourse is the set of all finite-length binary strings: λ, 0, 1, 00, 01, 10, 11, 000, 001, ... (Here λ denotes the empty string.) In your translations, you may use all the ordinary logic symbols (including =), variables, and the binary symbols $\mathtt{0}$, $\mathtt{1}$ denoting 0, 1.

A string like $\mathtt{01}x\mathtt{0}y$ of binary symbols and variables denotes the *concatenation* of the symbols and the binary strings represented by the variables. For example, if the value of x is $\mathtt{011}$ and the value of y is $\mathtt{1111}$, then the value of $\mathtt{01}x\mathtt{0}y$ is the binary string $\mathtt{0101101111}$.

Here are some examples of formulas and their English translations. Names for these predicates are listed in the third column so that you can reuse them in your solutions (as we do in the definition of the predicate NO-1S below).

Meaning	Formula	Name
x is a prefix of y	$\exists z\ (xz = y)$	PREFIX(x, y)
x is a substring of y	$\exists u \exists v\ (uxv = y)$	SUBSTRING(x, y)
x is empty or a string of 0's	NOT(SUBSTRING$(\mathtt{1}, x)$)	NO-1S(x)

(a) x consists of three copies of some string.

(b) x is an even-length string of 0's.

(c) x does not contain both a 0 and a 1.

(d) x is the binary representation of $2^k + 1$ for some integer $k \geq 0$.

(e) An elegant, slightly trickier way to define NO-1S(x) is:

$$\text{PREFIX}(x, \mathtt{0}x). \qquad (*)$$

Explain why (*) is true only when x is a string of 0's.

Problem 3.
Translate the following sentence into a predicate formula:

> There is a student who has e-mailed at most two other people in the class, besides possibly himself.

The domain of discourse should be the set of students in the class; in addition, the only predicates that you may use are

- equality, and
- $E(x, y)$, meaning that "x has sent e-mail to y."

Problem 4.
Provide a counter model for the implication that is not valid. Informally explain why the other one is valid.

1. $\forall x.\, \exists y.\, P(x, y)$ IMPLIES $\exists y.\, \forall x.\, P(x, y)$
2. $\exists y.\, \forall x.\, P(x, y)$ IMPLIES $\forall x.\, \exists y.\, P(x, y)$

Supplemental Problem[1]

Problem 5.
A certain cabal within the Math for Computer Science course staff is plotting to make the final exam *ridiculously hard*. ("Problem 1. Prove the Poincare Conjecture starting from the axioms of ZFC. Express your answer in khipu—the knot language of the Incas.") The only way to stop their evil plan is to determine exactly who is in the cabal. The course staff consists of seven people:

$$\{\text{Adam, Tom, Albert, Annie, Ben, Elizabeth, Siggi}\}$$

The cabal is a subset of these seven. A membership roster has been found and appears below, but it is deviously encrypted in logic notation. The predicate cabal indicates who is in the cabal; that is, cabal(x) is true if and only if x is a member. Translate each statement below into English and deduce who is in the cabal.

(a) $\exists x, y, z.\, (x \neq y$ AND $x \neq z$ AND $y \neq z$ AND cabal(x) AND cabal(y) AND cabal(z))

(b) NOT(cabal(Siggi) AND cabal(Annie))

(c) cabal(Elizabeth) IMPLIES $\forall x.$ cabal(x)

(d) cabal(Annie) IMPLIES cabal(Siggi)

(e) (cabal(Ben) OR cabal(Albert)) IMPLIES NOT(cabal(Tom))

(f) (cabal(Ben) OR cabal(Siggi)) IMPLIES NOT(cabal(Adam))

(g) Now use these facts to explain exactly who is on the cabal and why.

[1] There is no need to study supplemental problems when preparing for quizzes or exams.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 3, Wed.

Problem 1.

Set Formulas and Propositional Formulas.

(a) Verify that the propositional formula $(P \text{ AND } \overline{Q}) \text{ OR } (P \text{ AND } Q)$ is equivalent to P.

(b) Prove that[1]

$$A = (A - B) \cup (A \cap B)$$

for all sets, A, B, by showing

$$x \in A \text{ IFF } x \in (A - B) \cup (A \cap B)$$

for all elements, x, using the equivalence of part (a) in a chain of IFF's.

Problem 2.

A *formula of set theory* is a predicate formula that only uses the predicate "$x \in y$." The domain of discourse is the collection of sets, and "$x \in y$" is interpreted to mean the set x is one of the elements in the set y.

For example, since x and y are the same set iff they have the same members, here's how we can express equality of x and y with a formula of set theory:

$$(x = y) ::= \forall z.\, (z \in x \text{ IFF } z \in y). \tag{1}$$

Express each of the following assertions about sets by a formula of set theory.

(a) $x = \emptyset$.

(b) $x = \{y, z\}$.

(c) $x \subseteq y$. (x is a subset of y that might equal y.)

Now we can explain how to express "x is a proper subset of y" as a set theory formula using things we already know how to express. Namely, letting "$x \neq y$" abbreviate $\text{NOT}(x = y)$, the expression

$$(x \subseteq y \text{ AND } x \neq y),$$

describes a formula of set theory that means $x \subset y$.

From here on, feel free to use any previously expressed property in describing formulas for the following:

(d) $x = y \cup z$.

(e) $x = y - z$.

[1] The *set difference*, $A - B$, of sets A and B is

$$A - B ::= \{a \in A \mid a \notin B\}.$$

(f) $x = \text{pow}(y)$.

(g) $x = \bigcup_{z \in y} z$.

This means that y is supposed to be a collection of sets, and x is the union of all them. A more concise notation for "$\bigcup_{z \in y} z$" is simply "$\bigcup y$."

Problem 3.

Forming a pair (a, b) of items a and b is a mathematical operation that we can safely take for granted. But when we're trying to show how all of mathematics can be reduced to set theory, we need a way to represent the pair (a, b) as a set.

(a) Explain why representing (a, b) by $\{a, b\}$ won't work.

(b) Explain why representing (a, b) by $\{a, \{b\}\}$ won't work either. *Hint:* What pair does $\{\{1\}, \{2\}\}$ represent?

(c) Define

$$\text{pair}(a, b) ::= \{a, \{a, b\}\}.$$

Explain why representing (a, b) as $\text{pair}(a, b)$ uniquely determines a and b. *Hint:* Sets can't be indirect members of themselves: $a \in a$ never holds for any set a, and neither can $a \in b \in a$ hold for any b.

Problem 4.

Subset take-away[2] is a two player game played with a finite set, A, of numbers. Players alternately choose nonempty subsets of A with the conditions that a player may not choose

- the whole set A, or
- any set containing a set that was named earlier.

The first player who is unable to move loses the game.

For example, if the size of A is one, then there are no legal moves and the second player wins. If A has exactly two elements, then the only legal moves are the two one-element subsets of A. Each is a good reply to the other, and so once again the second player wins.

The first interesting case is when A has three elements. This time, if the first player picks a subset with one element, the second player picks the subset with the other two elements. If the first player picks a subset with two elements, the second player picks the subset whose sole member is the third element. In both cases, these moves lead to a situation that is the same as the start of a game on a set with two elements, and thus leads to a win for the second player.

Verify that when A has four elements, the second player still has a winning strategy.[3]

[2] From Christenson & Tilford, *David Gale's Subset Takeaway Game, American Mathematical Monthly, Oct. 1997*

[3] David Gale worked out some of the properties of this game and conjectured that the second player wins the game for any set A. This remains an open problem.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 3, Fri.

Problem 1.
The *inverse*, R^{-1}, of a binary relation, R, from A to B, is the relation from B to A defined by:

$$b \; R^{-1} \; a \quad \text{iff} \quad a \; R \; b.$$

In other words, you get the diagram for R^{-1} from R by "reversing the arrows" in the diagram describing R. Now many of the relational properties of R correspond to different properties of R^{-1}. For example, R is *total* iff R^{-1} is a *surjection*.

Fill in the remaining entries is this table:

R is	iff	R^{-1} is
total		a surjection
a function		
a surjection		
an injection		
a bijection		

Hint: Explain what's going on in terms of "arrows" from A to B in the diagram for R.

Arrow Properties

Definition. A binary relation, R is

- is a *function* when it has the [≤ 1 arrow **out**] property.
- is *surjective* when it has the [≥ 1 arrows **in**] property. That is, every point in the righthand, codomain column has at least one arrow pointing to it.
- is *total* when it has the [≥ 1 arrows **out**] property.
- is *injective* when it has the [≤ 1 arrow **in**] property.
- is *bijective* when it has both the [$= 1$ arrow **out**] and the [$= 1$ arrow **in**] property.

Problem 2.
Let $A = \{a_0, a_1, \ldots, a_{n-1}\}$ be a set of size n, and $B = \{b_0, b_1, \ldots, b_{m-1}\}$ a set of size m. Prove that $|A \times B| = mn$ by defining a simple bijection from $A \times B$ to the nonnegative integers from 0 to $mn - 1$.

Problem 3.
Assume $f : A \to B$ is total function, and A is finite. Replace the $\star$ with one of $\leq, =, \geq$ to produce the *strongest* correct version of the following statements:

(a) $|f(A)| \star |B|$.

(b) If f is a surjection, then $|A| \star |B|$.

(c) If f is a surjection, then $|f(A)| \star |B|$.

(d) If f is an injection, then $|f(A)| \star |A|$.

(e) If f is a bijection, then $|A| \star |B|$.

Problem 4.
Let $R : A \to B$ be a binary relation. Use an arrow counting argument to prove the following generalization of the Mapping Rule 1 in the course textbook.

Lemma. *If R is a function, and $X \subseteq A$, then*

$$|X| \geq |R(X)|.$$

Problem 5. (a) Prove that if A surj B and B surj C, then A surj C.

(b) Explain why A surj B iff B inj A.

(c) Conclude from (a) and (b) that if A inj B and B inj C, then A inj C.

(d) Explain why A inj B iff there is a total injective *function* ([$=$ 1 out, $\leq$ 1 in]) from A to B. [1]

[1] The official definition of inj is with a total injective *relation* ([$\geq$ 1 out, $\leq$ 1 in])

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 4, Mon.

Problem 1.
Prove by induction:

$$1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} < 2 - \frac{1}{n}, \tag{1}$$

for all $n > 1$.

Problem 2. **(a)** Prove by induction that a $2^n \times 2^n$ courtyard with a 1×1 statue of Bill in *a corner* can be covered with L-shaped tiles. (Do not assume or reprove the (stronger) result of Theorem 5.1.2 in the course textbook that Bill can be placed anywhere. The point of this problem is to show a different induction hypothesis that works.)
(b) Use the result of part (a) to prove the original claim that there is a tiling with Bill in the middle.

Problem 3.
Any amount of 12 or more cents postage can be made using only 3¢ and 7¢ stamps. Prove this *by induction* using the induction hypothesis

$$S(n) ::= n + 12 \text{ cents postage can be made using only } 3¢ \text{ and } 7¢ \text{ stamps.}$$

Problem 4.
The following Lemma is true, but the *proof* given for it below is defective. Pinpoint *exactly* where the proof first makes an unjustified step and explain why it is unjustified.

Lemma. *For any prime p and positive integers $n, x_1, x_2, \ldots, x_n$, if $p \mid x_1 x_2 \ldots x_n$, then $p \mid x_i$ for some $1 \leq i \leq n$.*

Bogus proof. Proof by strong induction on n. The induction hypothesis, $P(n)$, is that Lemma holds for n.

Base case $n = 1$: When $n = 1$, we have $p \mid x_1$, therefore we can let $i = 1$ and conclude $p \mid x_i$.

Induction step: Now assuming the claim holds for all $k \leq n$, we must prove it for $n + 1$.

So suppose $p \mid x_1 x_2 \cdots x_{n+1}$. Let $y_n = x_n x_{n+1}$, so $x_1 x_2 \cdots x_{n+1} = x_1 x_2 \cdots x_{n-1} y_n$. Since the righthand side of this equality is a product of n terms, we have by induction that p divides one of them. If $p \mid x_i$ for some $i < n$, then we have the desired i. Otherwise $p \mid y_n$. But since y_n is a product of the two terms x_n, x_{n+1}, we have by strong induction that p divides one of them. So in this case $p \mid x_i$ for $i = n$ or $i = n + 1$. ■

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 4, Fri.

Problem 1.
Multiplying and dividing an integer n by 2 only requires a one digit left or right shift of the binary representation of n, which are hardware-supported fast operations on most computers. Here is a state machine, R, that computes the product of two nonnegative integers x and y using just these shift operations, along with integer addition:

$$\begin{aligned}
\text{states} &::= \mathbb{N}^3 \qquad \text{(triples of nonnegative integers)} \\
\text{start state} &::= (x, y, 0) \\
\text{transitions} &::= (r, s, a) \longrightarrow \begin{cases} (2r, s/2, a) & \text{for even } s > 0, \\ (2r, (s-1)/2, a + r) & \text{for odd } s > 0. \end{cases}
\end{aligned}$$

(a) Verify that

$$P((r, s, a)) ::= \quad [rs + a = xy] \tag{1}$$

is a preserved invariant of R.

(b) Prove that R is partially correct: if R reachs a final state—a state from which no transition is possible—then $a = xy$.

(c) Briefly explain why this state machine will terminate after a number of transitions bounded by a small constant times the *length* of the binary representation of y.

Problem 2.
In this problem you will establish a basic property of a puzzle toy called the *Fifteen Puzzle* using the method of invariants. The Fifteen Puzzle consists of sliding square tiles numbered $1, \ldots, 15$ held in a 4×4 frame with one empty square. Any tile adjacent to the empty square can slide into it.

The standard initial position is

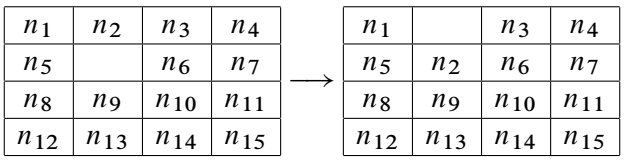
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

We would like to reach the target position (known in the oldest author's youth as "the impossible"):

15	14	13	12
11	10	9	8
7	6	5	4
3	2	1	

A state machine model of the puzzle has states consisting of a 4×4 matrix with 16 entries consisting of the integers $1, \ldots, 15$ as well as one "empty" entry—like each of the two arrays above.

The state transitions correspond to exchanging the empty square and an adjacent numbered tile. For example, an empty at position $(2, 2)$ can exchange position with tile above it, namely, at position $(1, 2)$:



We will use the invariant method to prove that there is no way to reach the target state starting from the initial state.

We begin by noting that a state can also be represented as a pair consisting of two things:

1. a list of the numbers $1, \ldots, 15$ in the order in which they appear—reading rows left-to-right from the top row down, ignoring the empty square, and

2. the coordinates of the empty square—where the upper left square has coordinates $(1, 1)$, the lower right $(4, 4)$.

(a) Write out the "list" representation of the start state and the "impossible" state.

Let L be a list of the numbers $1, \ldots, 15$ in some order. A pair of integers is an *out-of-order pair* in L when the first element of the pair both comes *earlier* in the list and *is larger*, than the second element of the pair. For example, the list $1, 2, 4, 5, 3$ has two out-of-order pairs: (4,3) and (5,3). The increasing list $1, 2 \ldots n$ has no out-of-order pairs.

Let a state, S, be a pair $(L, (i, j))$ described above. We define the *parity* of S to be 0 or 1 depending on whether the sum of the number of out-of-order pairs in L and the row-number of the empty square is even or odd. that is

$$\text{parity}(S) ::= \begin{cases} 0 & \text{if } p(L) + i \text{ is even,} \\ 1 & \text{otherwise.} \end{cases}$$

(b) Verify that the parity of the start state and the target state are different.

(c) Show that the parity of a state is preserved under transitions. Conclude that "the impossible" is impossible to reach.

By the way, if two states have the same parity, then in fact there *is* a way to get from one to the other. If you like puzzles, you'll enjoy working this out on your own.

Problem 3.

The Massachusetts Turnpike Authority is concerned about the integrity of the new Zakim bridge. Their consulting architect has warned that the bridge may collapse if more than 1000 cars are on it at the same time. The Authority has also been warned by their traffic consultants that the rate of accidents from cars speeding across bridges has been increasing.

Both to lighten traffic and to discourage speeding, the Authority has decided to make the bridge *one-way* and to put tolls at *both* ends of the bridge (don't laugh, this is Massachusetts). So cars will pay tolls both on entering and exiting the bridge, but the tolls will be different. In particular, a car will pay \$3 to enter onto the bridge and will pay \$2 to exit. To be sure that there are never too many cars on the bridge, the Authority will let a car onto the bridge only if the difference between the amount of money currently at the entry toll booth and the amount at the exit toll booth is strictly less than a certain threshold amount of $\$T_0$.

The consultants have decided to model this scenario with a state machine whose states are triples of nonnegative integers, (A, B, C), where

- A is an amount of money at the entry booth,
- B is an amount of money at the exit booth, and
- C is a number of cars on the bridge.

Any state with $C > 1000$ is called a *collapsed* state, which the Authority dearly hopes to avoid. There will be no transition out of a collapsed state.

Since the toll booth collectors may need to start off with some amount of money in order to make change, and there may also be some number of "official" cars already on the bridge when it is opened to the public, the consultants must be ready to analyze the system started at *any* uncollapsed state. So let A_0 be the initial number of dollars at the entrance toll booth, B_0 the initial number of dollars at the exit toll booth, and $C_0 \leq 1000$ the number of official cars on the bridge when it is opened. You should assume that even official cars pay tolls on exiting or entering the bridge after the bridge is opened.

(a) Give a mathematical model of the Authority's system for letting cars on and off the bridge by specifying a transition relation between states of the form (A, B, C) above.

(b) Characterize each of the following derived variables

$$A, B, A + B, A - B, 3C - A, 2A - 3B, B + 3C, 2A - 3B - 6C, 2A - 2B - 3C$$

as one of the following

constant	C
strictly increasing	SI
strictly decreasing	SD
weakly increasing but not constant	WI
weakly decreasing but not constant	WD
none of the above	N

and briefly explain your reasoning.

The Authority has asked their engineering consultants to determine T and to verify that this policy will keep the number of cars from exceeding 1000.

The consultants reason that if C_0 is the number of official cars on the bridge when it is opened, then an additional $1000 - C_0$ cars can be allowed on the bridge. So as long as $A - B$ has not increased by $3(1000 - C_0)$, there shouldn't more than 1000 cars on the bridge. So they recommend defining

$$T_0 ::= 3(1000 - C_0) + (A_0 - B_0), \tag{2}$$

where A_0 is the initial number of dollars at the entrance toll booth, B_0 is the initial number of dollars at the exit toll booth.

(c) Use the results of part (b) to define a simple predicate, P, on states of the transition system which is satisfied by the start state —that is $P(A_0, B_0, C_0)$ holds —is not satisfied by any collapsed state, and is a preserved invariant of the system. Explain why your P has these properties. Conclude that the traffic won't cause the bridge to collapse.

(d) A clever MIT intern working for the Turnpike Authority agrees that the Turnpike's bridge management policy will be *safe*: the bridge will not collapse. But she warns her boss that the policy will lead to *deadlock*—a situation where traffic can't move on the bridge even though the bridge has not collapsed.

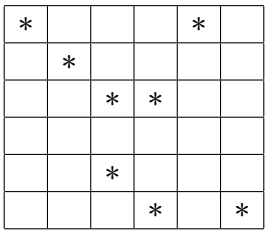
Explain more precisely in terms of system transitions what the intern means, and briefly, but clearly, justify her claim.

Supplemental problem:

Problem 4.

A classroom is designed so students sit in a square arrangement. An outbreak of beaver flu sometimes infects students in the class; beaver flu is a rare variant of bird flu that lasts forever, with symptoms including a yearning for more quizzes and the thrill of late night problem set sessions.

Here is an illustration of a 6×6-seat classroom with seats represented by squares. The locations of infected students are marked with an asterisk.

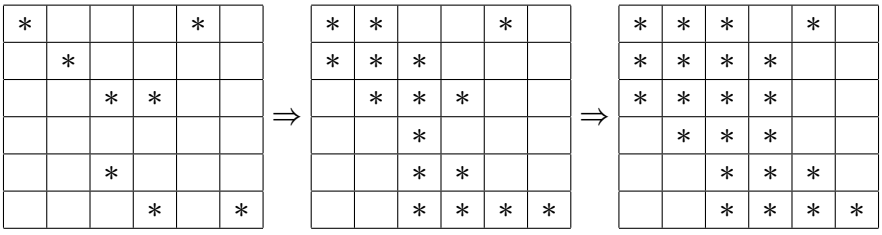


Outbreaks of infection spread rapidly step by step. A student is infected after a step if either

- the student was infected at the previous step (since beaver flu lasts forever), or
- the student was adjacent to *at least two* already-infected students at the previous step.

Here *adjacent* means the students' individual squares share an edge (front, back, left or right); they are not adjacent if they only share a corner point. So each student is adjacent to 2, 3 or 4 others.

In the example, the infection spreads as shown below.



In this example, over the next few time-steps, all the students in class become infected.

Theorem. *If fewer than n students among those in an $n \times n$ arrangment are initially infected in a flu outbreak, then there will be at least one student who never gets infected in this outbreak, even if students attend all the lectures.*

Prove this theorem.

Hint: Think of the state of an outbreak as an $n \times n$ square above, with asterisks indicating infection. The rules for the spread of infection then define the transitions of a state machine. Find a weakly decreasing derived variable that leads to a proof of this theorem. (If you don't see it within 4 minutes, ask your TA for a four word hint.)

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 5, Mon.

Problem 1.
The Elementary 18.01 Functions (F18's) are the set of functions of one real variable defined recursively as follows:

Base cases:

- The identity function, $\mathrm{id}(x) ::= x$ is an F18,
- any constant function is an F18,
- the sine function is an F18,

Constructor cases:
If f, g are F18's, then so are

1. $f + g$, fg, 2^g,
2. the inverse function f^{-1},
3. the composition $f \circ g$.

(a) Prove that the function $1/x$ is an F18.

Warning: Don't confuse $1/x = x^{-1}$ with the inverse id^{-1} of the identity function $\mathrm{id}(x)$. The inverse id^{-1} is equal to id.

(b) Prove by Structural Induction on this definition that the Elementary 18.01 Functions are *closed under taking derivatives*. That is, show that if $f(x)$ is an F18, then so is $f' ::= df/dx$. (Just work out 2 or 3 of the most interesting constructor cases; you may skip the less interesting ones.)

Problem 2.
Let p be the string [ ]. A string of brackets is said to be *erasable* iff it can be reduced to the empty string by repeatedly erasing occurrences of p. For example, here's how to erase the string [[[ ]] [ ]] [ ]:

$$[\,[\,[\,]\,]\,[\,]\,]\,[\,] \to [\,[\,]\,]\,[\,] \to [\,]\,[\,] \to \lambda.$$

On the other hand the string [ ]] [[[[ ]]] is not erasable because when we try to erase, we get stuck:] [[[:

$$[\,]\,]\,[\,[\,[\,[\,]\,]\,]\,] \to]\,[\,[\,[\,[\,]\,] \to]\,[\,[\,[\not\to$$

Let Erasable be the set of erasable strings of brackets. Let RecMatch be the recursive data type of strings of *matched* brackets defined recursively:

- **Base case**: $\lambda \in$ RecMatch.

- **Constructor case**: If $s, t \in \text{RecMatch}$, then $[\, s \,]\, t \in \text{RecMatch}$.

.

(a) Use structural induction to prove that

$$\text{RecMatch} \subseteq \text{Erasable}.$$

(b) Supply the missing parts (labeled by "(*)") of the following proof that

$$\text{Erasable} \subseteq \text{RecMatch}.$$

Proof. We prove by strong induction that every length n string in Erasable is also in RecMatch. The induction hypothesis is

$$P(n) ::= \forall x \in \text{Erasable}.\ |x| = n \text{ IMPLIES } x \in \text{RecMatch}.$$

Base case:

(*) What is the base case? Prove that P is true in this case.

Inductive step: To prove $P(n+1)$, suppose $|x| = n+1$ and $x \in$ Erasable. We need to show that $x \in$ RecMatch.

Let's say that a string y is an *erase* of a string z iff y is the result of erasing a *single* occurrence of p in z.

Since $x \in$ Erasable and has positive length, there must be an erase, $y \in$ Erasable, of x. So $|y| = n - 1 \geq 0$, and since $y \in$ Erasable, we may assume by induction hypothesis that $y \in$ RecMatch.

Now we argue by cases:

Case (y is the empty string):

(*) Prove that $x \in$ RecMatch in this case.

Case ($y = [\, s \,]\, t$ for some strings $s, t \in$ RecMatch): Now we argue by subcases.

- **Subcase**($x = py$):
  (*) Prove that $x \in$ RecMatch in this subcase.
- **Subcase** (x is of the form $[\, s' \,]\, t$ where s is an erase of s'):
  Since $s \in$ RecMatch, it is erasable by part (b), which implies that $s' \in$ Erasable. But $|s'| < |x|$, so by induction hypothesis, we may assume that $s' \in$ RecMatch. This shows that x is the result of the constructor step of RecMatch, and therefore $x \in$ RecMatch.
- **Subcase** (x is of the form $[\, s \,]\, t'$ where t is an erase of t'):
  (*) Prove that $x \in$ RecMatch in this subcase.

(*) Explain why the above cases are sufficient.

This completes the proof by strong induction on n, so we conclude that $P(n)$ holds for all $n \in \mathbb{N}$. Therefore $x \in$ RecMatch for every string $x \in$ Erasable. That is, Erasable $\subseteq$ RecMatch. Combined with part (a), we conclude that

$$\text{Erasable} = \text{RecMatch}.$$

■

Problem 3.
Here is a simple recursive definition of the set, E, of even integers:

Definition. Base case: $0 \in E$.

Constructor cases: If $n \in E$, then so are $n + 2$ and $-n$.

Provide similar simple recursive definitions of the following sets:

(a) The set $S ::= \{2^k 3^m 5^n \in \mathbb{N} \mid k, m, n \in \mathbb{N}\}$.

(b) The set $T ::= \{2^k 3^{2k+m} 5^{m+n} \in \mathbb{N} \mid k, m, n \in \mathbb{N}\}$.

(c) The set $L ::= \{(a, b) \in \mathbb{Z}^2 \mid (a - b) \text{ is a multiple of } 3\}$.

Let L' be the set defined by the recursive definition you gave for L in the previous part. Now if you did it right, then $L' = L$, but maybe you made a mistake. So let's check that you got the definition right.

(d) Prove by structural induction on your definition of L' that

$$L' \subseteq L.$$

(e) Confirm that you got the definition right by proving that

$$L \subseteq L'.$$

(f) See if you can give an *unambiguous* recursive definition of L.

Supplemental problem:

Problem 4.

Definition. The recursive data type, binary-2PTG, of *binary trees* with leaf labels, L, is defined recursively as follows:

- **Base case:** $\langle \texttt{leaf}, l \rangle \in$ binary-2PTG, for all labels $l \in L$.
- **Constructor case:** If $G_1, G_2 \in$ binary-2PTG, then

$$\langle \texttt{bintree}, G_1, G_2 \rangle \in \text{binary-2PTG}.$$

The *size*, $|G|$, of $G \in$ binary-2PTG is defined recursively on this definition by:

- **Base case:**

$$|\langle \texttt{leaf}, l \rangle| ::= 1, \quad \text{for all } l \in L.$$

- **Constructor case:**

$$|\langle \texttt{bintree}, G_1, G_2 \rangle| ::= |G_1| + |G_2| + 1.$$

For example, the size of the binary-2PTG, G, pictured in Figure 1, is 7.

(a) Write out (using angle brackets and labels `bintree`, `leaf`, etc.) the binary-2PTG, G, pictured in Figure 1.

The value of flatten(G) for $G \in$ binary-2PTG is the sequence of labels in L of the leaves of G. For example, for the binary-2PTG, G, pictured in Figure 1,

$$\text{flatten}(G) = (\texttt{win}, \texttt{lose}, \texttt{win}, \texttt{win}).$$

(b) Give a recursive definition of flatten. (You may use the operation of *concatenation* (append) of two sequences.)

(c) Prove by structural induction on the definitions of flatten and size that

$$2 \cdot \text{length}(\text{flatten}(G)) = |G| + 1. \tag{1}$$

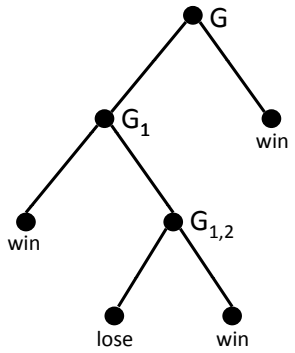


Figure 1 A picture of a binary tree G.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 5, Wed.

Problem 1. (a) Several students felt the proof of Lemma 7.1.7 was worrisome, if not circular in the course textbook. What do you think?

Lemma 7.1.7. Let A be a set and $b \notin A$. If A is infinite, then there is a bijection from $A \cup \{b\}$ to A.

Proof. Here's how to define the bijection: since A is infinite, it certainly has at least one element; call it a_0. But since A is infinite, it has at least two elements, and one of them must not be equal to a_0; call this new element a_1. But since A is infinite, it has at least three elements, one of which must not equal a_0 or a_1; call this new element a_2. Continuing in the way, we conclude that there is an infinite sequence $a_0, a_1, a_2, \ldots, a_n, \ldots$ of different elements of A. Now we can define a bijection $f : A \cup \{b\} \to A$:

$$\begin{aligned} f(b) &::= a_0, \\ f(a_n) &::= a_{n+1} && \text{for } n \in \mathbb{N}, \\ f(a) &::= a && \text{for } a \in A - \{a_0, a_1, \ldots\}. \end{aligned}$$

■

(b) Use the proof of Lemma 7.1.7 to show that if A is an infinite set, then A surj $\mathbb{N}$, that is, every infinite set is "as big as" the set of nonnegative integers in the course textbook.

Problem 2.
Prove that if there is a surjective function ([≤ 1 out, ≥ 1 in] mapping) $f : \mathbb{N} \to S$, then S is countable.

Hint: A Computer Science proof involves filtering for duplicates.

Problem 3.
The rational numbers fill the space between integers, so a first thought is that there must be more of them than the integers, but it's not true. In this problem you'll show that there are the same number of positive rationals as positive integers. That is, the positive rationals are countable.

(a) Define a bijection between the set, $\mathbb{Z}^+$, of positive integers, and the set, $(\mathbb{Z}^+ \times \mathbb{Z}^+)$, of all pairs of positive integers:

$$\begin{array}{c} (1,1), (1,2), (1,3), (1,4), (1,5), \ldots \\ (2,1), (2,2), (2,3), (2,4), (2,5), \ldots \\ (3,1), (3,2), (3,3), (3,4), (3,5), \ldots \\ (4,1), (4,2), (4,3), (4,4), (4,5), \ldots \\ (5,1), (5,2), (5,3), (5,4), (5,5), \ldots \\ \vdots \end{array}$$

(b) Conclude that the set, $\mathbb{Q}^+$, of all positive rational numbers is countable.

Hint: Use Problem 2.

Problem 4.

Let's refer to a programming procedure (written in your favorite programming language —C++, or Java, or Python, ...) as a *string procedure* when it is applicable to data of type **string** and only returns values of type **boolean**. When a string procedure, P, applied to a **string**, s, returns **True**, we'll say that P *recognizes* s. If $\mathcal{R}$ is the set of strings that P recognizes, we'll call P a *recognizer* for $\mathcal{R}$.

(a) Describe how a recognizer would work for the set of strings containing only lowercase Roman letters —`a,b,...,z` —such that each letter occurs twice in a row. For example, `aaccaabbzz`, is such a string, but `abb`, `00bb`, `AAbb`, and `a` are not. (Even better, actually write a recognizer procedure in your favorite programming language).

A set of strings is called *recognizable* if there is a recognizer procedure for it. So the program you described above proves that the set of strings with doubled letters from part (a) is recognizable.

When you actually program a procedure, you have to type the program text into a computer system. This means that every procedure is described by some string of typed characters. If a string, s, is actually the typed description of some string procedure, let's refer to that procedure as P_s. You can think of P_s as the result of compiling s.[1]

In fact, it will be helpful to associate every string, s, with a procedure, P_s. So if string s is not the typed description of a string procedure, we will define P_s to be some fixed string procedure —say one that always returns **False**; so if s is an ill-formed string, P_s will be a recognizer for the empty set of strings.

The result of this is that we have now defined a total function, f, mapping every string, s, to the set, $f(s)$, of strings recognized by P_s. That is we have a total function,

$$f : \textbf{string} \to \text{pow}(\textbf{string}). \tag{1}$$

(b) Explain why range(f) is the set of all recognizable sets of strings.

This is exactly the set up we need to apply the reasoning behind Russell's Paradox to define a set that is not in the range of f, that is, a set of strings, $\mathcal{N}$, that is *not* recognizable.

(c) Let

$$\mathcal{N} ::= \{s \in \textbf{string} \mid s \notin f(s)\}.$$

Prove that $\mathcal{N}$ is not recognizable.

Hint: Similar to Russell's paradox or the proof of Theorem 7.1.11 in the course textbook.

(d) Discuss what the conclusion of part (c) implies about the possibility of writing "program analyzers" that take programs as inputs and analyze their behavior.

[1] The string, s, and the procedure, P_s, have to be distinguished to avoid a type error: you can't apply a string to string. For example, let s be the string that you wrote as your program to answer part (a). Applying s to a string argument, say `aabbccdd`, should throw a type exception; what you need to do is apply the procedure P_s to `aabbccdd`. This should result in a returned value **True**, since `aabbccdd` consists of consecutive pairs of lowercase roman letters.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 5, Fri.

Problem 1.

(a) Use the Pulverizer to find integers x, y such that

$$x30 + y22 = \gcd(30, 22).$$

(b) Now find integers x', y' with $0 \leq y' < 30$ such that

$$x'30 + y'22 = \gcd(30, 22)$$

Problem 2. **(a)** Let $m = 2^9 5^{24} 11^7 17^{12}$ and $n = 2^3 7^{22} 11^{211} 13^1 17^9 19^2$. What is the $\gcd(m, n)$? What is the *least common multiple*, $\text{lcm}(m, n)$, of m and n? Verify that

$$\gcd(m, n) \cdot \text{lcm}(m, n) = mn. \tag{1}$$

(b) Describe in general how to find the $\gcd(m, n)$ and $\text{lcm}(m, n)$ from the prime factorizations of m and n. Conclude that equation (1) holds for all positive integers m, n.

Problem 3.

The *Binary GCD* state machine computes the GCD of integers $a, b > 0$ using only division by 2 and subtraction, which makes it run very efficiently on hardware that uses binary representation of numbers. In practice, it runs more quickly than the more famous Euclidean algorithm described in Section 8.2.1 in the course textbook.

$$\begin{aligned}
\text{states} &::= \mathbb{N}^3 \\
\text{start state} &::= (a, b, 1) \\
\text{transitions} &::= \text{if } \min(x, y) > 0, \text{ then } (x, y, e) \longrightarrow \\
&\quad (x/2, y/2, 2e) && (\text{if } 2 \mid x \text{ and } 2 \mid y) && (2) \\
&\quad (x/2, y, e) && (\text{else if } 2 \mid x) && (3) \\
&\quad (x, y/2, e) && (\text{else if } 2 \mid y) && (4) \\
&\quad (x - y, y, e) && (\text{else if } x > y) && (5) \\
&\quad (y - x, x, e) && (\text{else if } y > x) && (6) \\
&\quad (1, 0, ex) && (\text{otherwise } (x = y)). && (7)
\end{aligned}$$

(a) Use the Invariant Principle to prove that if this machine stops, that is, reaches a state (x, y, e) in which no transition is possible, then $e = \gcd(a, b)$.

(b) Prove that rule (2)

$$(x, y, e) \to (x/2, y/2, 2e)$$

is never executed after any of the other rules is executed.

(c) Prove that the machine reaches a final state in at most $1 + 3(\log a + \log b)$ transitions. (This is a coarse bound; you may be able to get a better one.)

Problem 4.

For nonzero integers, a, b, prove the following properties of divisibility and GCD'S. (You may use the fact that $\gcd(a, b)$ is an integer linear combination of a and b. You may *not* appeal to uniqueness of prime factorization because the properties below are needed to *prove* unique factorization.)

(a) Every common divisor of a and b divides $\gcd(a, b)$.

(b) If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.

(c) If $p \mid bc$ for some prime, p, then $p \mid b$ or $p \mid c$.

(d) Let m be the smallest integer linear combination of a and b that is positive. Show that $m = \gcd(a, b)$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 6, Mon.

Problem 1.
Find

$$\text{remainder}\left(9876^{3456789}\left(9^{99}\right)^{5555} - 6789^{3414259},\ 14\right). \tag{1}$$

Problem 2.
Suppose a, b are relatively prime and greater than 1. In this problem you will prove the *Chinese Remainder Theorem*, which says that for all m, n, there is an x such that

$$x \equiv m \mod a, \tag{2}$$
$$x \equiv n \mod b. \tag{3}$$

Moreover, x is unique up to congruence modulo ab, namely, if x' also satisfies (2) and (3), then

$$x' \equiv x \mod ab.$$

(a) Prove that for any m, n, there is some x satisfying (2) and (3).

Hint: Let b^{-1} be an inverse of b modulo a and define $e_a ::= b^{-1}b$. Define e_b similarly. Let $x = me_a + ne_b$.

(b) Prove that

$$[x \equiv 0 \mod a \ \text{ AND } \ x \equiv 0 \mod b] \quad \text{implies} \quad x \equiv 0 \mod ab.$$

(c) Conclude that

$$\left[x \equiv x' \mod a \ \text{ AND } \ x \equiv x' \mod b\right] \quad \text{implies} \quad x \equiv x' \mod ab.$$

(d) Conclude that the Chinese Remainder Theorem is true.

(e) What about the converse of the implication in part (c)?

Problem 3.

Definition. The set, P, of integer polynomials can be defined recursively:

Base cases:

- the identity function, $\text{Id}_{\mathbb{Z}}(x) ::= x$ is in P.
- for any integer, m, the constant function, $c_m(x) ::= m$ is in P.

Constructor cases. If $r, s \in P$, then $r + s$ and $r \cdot s \in P$.

(a) Using the recursive definition of integer polynomials given above, prove by structural induction that for all $q \in P$,

$$j \equiv k \pmod{n} \quad \text{IMPLIES} \quad q(j) \equiv q(k) \pmod{n},$$

for all integers j, k, n where $n > 1$.

Be sure to clearly state and label your Induction Hypothesis, Base case(s), and Constructor step.

(b) We'll say that q *produces multiples* if, for every integer greater than one in the range of q, there are infinitely many different multiples of that integer in the range. For example, if $q(4) = 7$ and q produces multiples, then there are infinitely many different multiples of 7 in the range of q.

Prove that if q has positive degree and positive leading coefficient, then q produces multiples. You may assume that every such polynomial is strictly increasing for large arguments.

Hint: Observe that all the elements in the sequence

$$q(k), q(k+v), q(k+2v), q(k+3v), \ldots,$$

are congruent modulo v. Let $v = q(k)$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 6, Wed.

Problem 1.
Find the remainder of $26^{1818181}$ divided by 297.

Hint: $1818181 = (180 \cdot 10101) + 1$; use Euler's theorem.

Problem 2. **(a)** Prove that 2012^{1200} has a multiplicative inverse modulo 77.

(b) What is the value of $\phi(77)$, where ϕ is Euler's function?

(c) What is the remainder of 2012^{1200} divided by 77?

Problem 3.
Prove that for any prime, p, and integer, $k \geq 1$,

$$\phi(p^k) = p^k - p^{k-1},$$

where ϕ is Euler's function. *Hint:* Which numbers between 0 and $p^k - 1$ *are* divisible by p? How many are there?

Note: This is proved in the text. Don't look up that proof.

Problem 4.
At one time, the Guinness Book of World Records reported that the "greatest human calculator" was a guy who could compute 13th roots of 100-digit numbers that were 13th powers. What a curious choice of tasks....

In this problem, we prove

$$n^{13} \equiv n \pmod{10} \tag{1}$$

for all n.

(a) Explain why (1) does not follow immediately from Euler's Theorem.

(b) Prove that

$$d^{13} \equiv d \pmod{10} \tag{2}$$

for $0 \leq d < 10$.

(c) Now prove the congruence (1).

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 6, Fri.

Problem 1.
Let's try out RSA! There is a complete description of the algorithm in the text box. You'll probably need extra paper. **Check your work carefully!**

(a) Go through the **beforehand** steps.

- Choose primes p and q to be relatively small, say in the range 10-40. In practice, p and q might contain hundreds of digits, but small numbers are easier to handle with pencil and paper.
- Try $e = 3, 5, 7, \ldots$ until you find something that works. Use Euclid's algorithm to compute the gcd.
- Find d (using the Pulverizer or Euler's Theorem).

When you're done, put your public key on the board prominentally labelled "Public Key." This lets another team send you a message.

(b) Now send an encrypted message to another team using their public key. Select your message m from the codebook below:

- 2 = Greetings and salutations!
- 3 = Yo, wassup?
- 4 = You guys are slow!
- 5 = All your base are belong to us.
- 6 = Someone on *our* team thinks someone on *your* team is kinda cute.
- 7 = You *are* the weakest link. Goodbye.

(c) Decrypt the message sent to you and verify that you received what the other team sent!

Problem 2. **(a)** Just as RSA would be trivial to crack knowing the factorization into two primes of n in the public key, explain why RSA would also be trivial to crack knowing $\phi(n)$.

(b) Show that if you knew n, $\phi(n)$, and that n was the product of two primes, then you could easily factor n.

Problem 3.
A critical fact about RSA is, of course, that decrypting an encrypted message always gives back the original message, m. Namely, if $n = pq$ where p and q are distinct primes, $m \in [0..pq)$, and

$$d \cdot e \equiv 1 \pmod{(p-1)(q-1)},$$

then

$$\widehat{m}^d ::= \left(m^e\right)^d = m \ (\mathbb{Z}_n). \tag{1}$$

We'll now prove this.

(a) Explain why (1) follows very simply from Euler's theorem when m is *relatively prime to* n.

All the rest of this problem is about removing the restriction that m be relatively prime to n. That is, we aim to prove that equation (1) holds for *all* $m \in [0..n)$.

It is important to realize that, even if it was theoretically necessary, there would be no practical reason to worry about—or to bother to check for—this relative primality condition before sending a message m using RSA. That's because the whole RSA enterprise is predicated on the difficulty of factoring. If an m ever came up that wasn't relatively prime to n, then we could factor n by computing $\gcd(m, n)$. So believing in the security of RSA implies believing that the probability of a message m turning up that was not relatively prime to n is negligible.

But let's be pure, impractical mathematicians and rid of this technically unnecessary relative primality side condition, even if it is harmless. One gain for doing this is that statements about RSA will be simpler without the side condition. More important, the proof below illustrates a useful general method of proving things about a number n by proving them separately for the prime factors of n.

(b) Prove that if p is prime and $a \equiv 1 \pmod{p-1}$, then

$$m^a = m \ (\mathbb{Z}_p). \tag{2}$$

(c) Give an elementary proof[1] that if $a \equiv b \pmod{p_i}$ for distinct primes p_i, then $a \equiv b$ modulo the product of these primes.

(d) Note that (1) is a special case of

Claim. *If n is a product of distinct primes and $a \equiv 1 \pmod{\phi(n)}$, then*

$$m^a = m \ (\mathbb{Z}_n).$$

Use the previous parts to prove the Claim.

[1] There is no need to appeal to the Chinese Remainder Theorem.

The RSA Cryptosystem

A **Receiver** who wants to be able to receive secret numerical messages creates a *private key*, which they keep secret, and a *public key*, which they make publicly available. Anyone with the public key can then be a **Sender** who can publicly send secret messages to the **Receiver**—even if they have never communicated or shared any information besides the public key.

Here is how they do it:

Beforehand The **Receiver** creates a public key and a private key as follows.

1. Generate two distinct primes, p and q. These are used to generate the private key, and they must be kept hidden. (In current practice, p and q are chosen to be hundreds of digits long.)
2. Let $n ::= pq$.
3. Select an integer $e \in [1, n)$ such that $\gcd(e, (p-1)(q-1)) = 1$.
   The *public key* is the pair (e, n). This should be distributed widely.
4. Compute $d \in [1, n)$ such that $de \equiv 1 \pmod{(p-1)(q-1)}$. This can be done using the Pulverizer.
   The *private key* is the pair (d, n). This should be kept hidden!

Encoding To transmit a message $m \in [0, n)$ to **Receiver**, a **Sender** uses the public key to encrypt m into a numerical message

$$\widehat{m} ::= \text{rem}(m^e,\ n).$$

The **Sender** can then publicly transmit $\widehat{m}$ to the **Receiver**.

Decoding The **Receiver** decrypts message $\widehat{m}$ back to message m using the private key:

$$m = \text{rem}(\widehat{m}^d,\ n).$$

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 7, Mon.

Problem 1. **(a)** Give an example of a digraph in which a vertex v is on a positive even-length closed walk, but *no* vertex is on an even-length cycle.

(b) Give an example of a digraph in which a vertex v is on an odd-length closed walk but not on an odd-length cycle.

(c) Prove that every odd-length closed walk contains a vertex that is on an odd-length cycle.

Problem 2.
In the course textbook Lemma 9.2.5 states that dist $(u, v) \leq$ dist (u, x) + dist (x, v). It also states that equality holds iff x is on a shortest path from u to v.

(a) Prove the "iff" statement from left to right.

(b) Prove the "iff" from right to left.

Problem 3.
A 3-bit string is a string made up of 3 characters, each a 0 or a 1. Suppose you'd like to write out, in one string, all eight of the 3-bit strings in any convenient order. For example, if you wrote out the 3-bit strings in the usual order starting with 000 001 010..., you could concatenate them together to get a length $3 \cdot 8 = 24$ string that started 000001010....

But you can get a shorter string containing all eight 3-bit strings by starting with 00010.... Now 000 is present as bits 1 through 3, and 001 is present as bits 2 through 4, and 010 is present as bits 3 through 5, ...

(a) Say a string is *3-good* if it contains every 3-bit string as 3 consecutive bits somewhere in it. Find a 3-good string of length 10, and explain why this is the minimum length for any string that is 3-good.

(b) Explain how any walk that includes every edge in the graph shown in Figure 1 determines a string that is 3-good. Find the walk in this graph that determines your 3-good string from part (a).

(c) Explain why a walk in the graph of Figure 1 that includes every every edge *exactly once* provides a minimum-length 3-good string.[1]

(d) Generalize the 2-bit graph to a k-bit digraph, B_k, for $k \geq 2$, where $V(B_k) ::= \{0,1\}^k$, and any walk through B_k that contains every edge exactly once determines a minimum length $(k+1)$-good bit-string.[2]

What is this minimum length?

Define the transitions of B_k. Verify that the in-degree and out-degree of every vertex is even, and that there is a positive path from any vertex to any other vertex (including itself) of length at most k.

[1] The 3-good strings explained here generalize to n-good strings for $n \geq 3$. They were studied by the great Dutch mathematician/logician Nicolaas de Bruijn, and are known as *de Bruijn sequences*. de Bruijn died in February, 2012 at the age of 94.

[2] Problem 9.23 explains why such "Eulerian" paths exist.

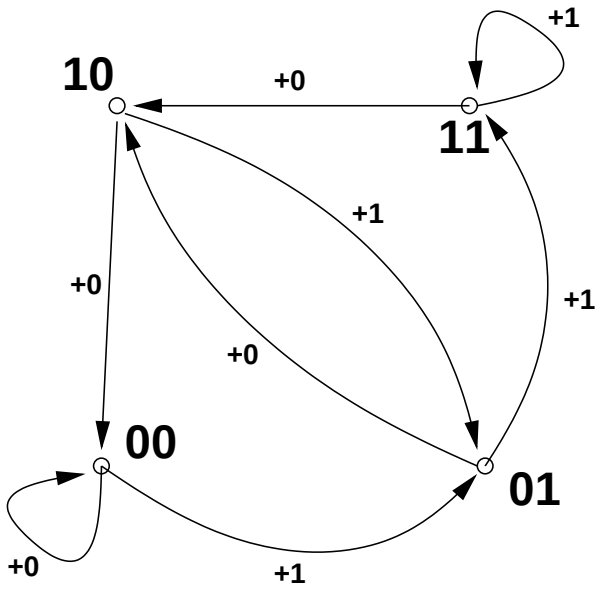


Figure 1 The 2-bit graph.

Suppemental Problem:

Problem 4.
In a round-robin tournament, every two distinct players play against each other just once. For a round-robin tournament with no tied games, a record of who beat whom can be described with a *tournament digraph*, where the vertices correspond to players and there is an edge $\langle x \to y \rangle$ iff x beat y in their game.

A *ranking* is a path that includes all the players. So in a ranking, each player won the game against the next ranked player, but may very well have lost their games against players ranked later—whoever does the ranking may have a lot of room to play favorites.

(a) Give an example of a tournament digraph with more than one ranking.

(b) Prove that every finite tournament digraph has a ranking.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 7, Fri.

Problem 1.
The table below lists some prerequisite information for some subjects in the MIT Computer Science program (in 2006). This defines an indirect prerequisite relation that is a DAG with these subjects as vertices.

$18.01 \to 6.042$	$18.01 \to 18.02$
$18.01 \to 18.03$	$6.046 \to 6.840$
$8.01 \to 8.02$	$6.001 \to 6.034$
$6.042 \to 6.046$	$18.03, 8.02 \to 6.002$
$6.001, 6.002 \to 6.003$	$6.001, 6.002 \to 6.004$
$6.004 \to 6.033$	$6.033 \to 6.857$

(a) Explain why exactly six terms are required to finish all these subjects, if you can take as many subjects as you want per term. Using a *greedy* subject selection strategy, you should take as many subjects as possible each term. Exhibit your complete class schedule each term using a greedy strategy.

(b) In the second term of the greedy schedule, you took five subjects including 18.03. Identify a set of five subjects not including 18.03 such that it would be possible to take them in any one term (using some nongreedy schedule). Can you figure out how many such sets there are?

(c) Exhibit a schedule for taking all the courses—but only one per term.

(d) Suppose that you want to take all of the subjects, but can handle only two per term. Exactly how many terms are required to graduate? Explain why.

(e) What if you could take three subjects per term?

Problem 2.
A pair of Math for Computer Science Teaching Assistants, Lisa and Annie, have decided to devote some of their spare time this term to establishing dominion over the entire galaxy. Recognizing this as an ambitious project, they worked out the following table of tasks on the back of Annie's copy of the lecture notes.

1. **Devise a logo** and cool imperial theme music - 8 days.
2. **Build a fleet** of Hyperwarp Stardestroyers out of eating paraphernalia swiped from Lobdell - 18 days.
3. **Seize control** of the United Nations - 9 days, after task #1.
4. **Get shots** for Lisa's cat, Tailspin - 11 days, after task #1.
5. **Open a Starbucks chain** for the army to get their caffeine - 10 days, after task #3.

6. **Train an army** of elite interstellar warriors by dragging people to see *The Phantom Menace* dozens of times - 4 days, after tasks #3, #4, and #5.

7. **Launch the fleet** of Stardestroyers, crush all sentient alien species, and establish a Galactic Empire - 6 days, after tasks #2 and #6.

8. **Defeat Microsoft** - 8 days, after tasks #2 and #6.

We picture this information in Figure 1 below by drawing a point for each task, and labelling it with the name and weight of the task. An edge between two points indicates that the task for the higher point must be completed before beginning the task for the lower one.

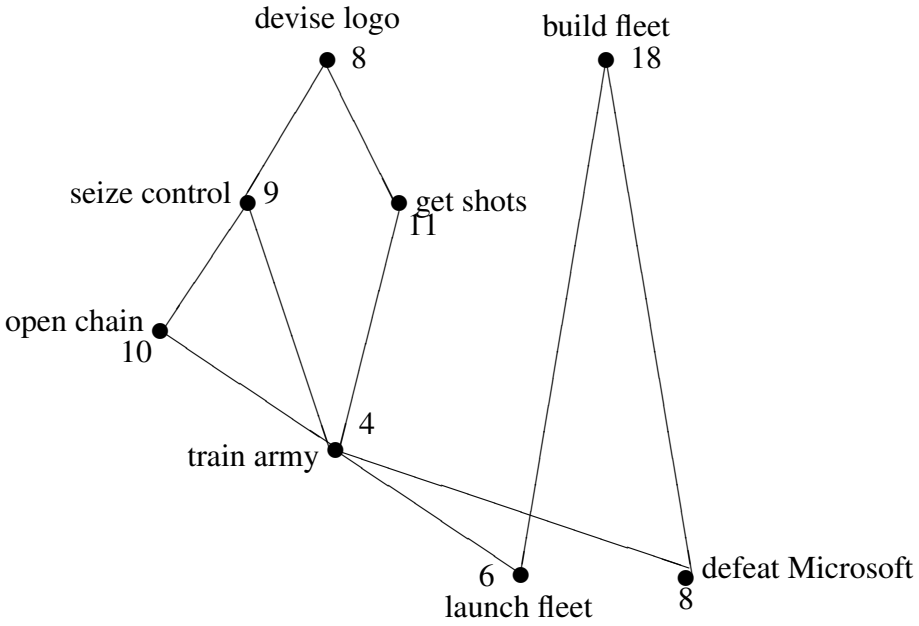


Figure 1 Graph representing the task precedence constraints.

(a) Give some valid order in which the tasks might be completed.

Lisa and Annie want to complete all these tasks in the shortest possible time. However, they have agreed on some constraining work rules.

- Only one person can be assigned to a particular task; they cannot work together on a single task.
- Once a person is assigned to a task, that person must work exclusively on the assignment until it is completed. So, for example, Lisa cannot work on building a fleet for a few days, run to get shots for Tailspin, and then return to building the fleet.

(b) Lisa and Annie want to know how long conquering the galaxy will take. Annie suggests dividing the total number of days of work by the number of workers, which is two. What lower bound on the time to conquer the galaxy does this give, and why might the actual time required be greater?

(c) Lisa proposes a different method for determining the duration of their project. She suggests looking at the duration of the *critical path*, the most time-consuming sequence of tasks such that each depends on the one before. What lower bound does this give, and why might it also be too low?

(d) What is the minimum number of days that Lisa and Annie need to conquer the galaxy? No proof is required.

Problem 3.
Sauron finds that conquering Middle Earth breaks down into a bunch of tasks. Each task can be completed by a horrible creature called a *Ringwraith* in exactly one week. Sauron realizes the prerequisite structure among the tasks defines a DAG. He has n tasks in his DAG, with a maximum length chain of t tasks.

(a) Sauron is trying to describe various features of his scheduling problem using standard terminology. For each feature below, indicate the number of the corresponding term.

Standard Terminology

1.	Indirect prerequisite	2.	Topological sort
3.	Chain	4.	Antichain
5.	Size of the largest antichain	6.	Size of the smallest antichain
7.	Length of the longest chain	8.	Length of the shortest chain

1. A set of tasks that can be worked on simultaneously.
2. A possible order in which all the tasks could be completed, if only one Ringwraith were available.
3. The minimum number of weeks required to complete all tasks, if an unlimited number of Ringwraiths were available.

(b) If Sauron is lucky, he will be able to get away with a small crew of Ringwraiths. Write a simple formula involving n and t for the smallest number of Ringwraiths that could possibly be able to complete all n tasks in t weeks. (Do not make any additional assumptions about the relative sizes of n and t besides $t \leq n$.) Given any n and t, describe a DAG that can be completed in t weeks using this number of Ringwraiths.

(c) On the other hand, if Sauron is unlucky, he may need a large crew of Ringwraiths in order to conquer Middle Earth in t weeks. Write a simple formula involving n and t for the largest number of Ringwraiths that Sauron would ever need in order to be sure of completing all n tasks in t weeks—no matter how unlucky he was. Given any n and t, describe a DAG that can be completed in t weeks and requires this number of Ringwraiths.

Problem 4.
If a and b are distinct nodes of a digraph, then a is said to *cover* b if there is an edge from a to b and every path from a to b includes this edge. If a covers b, the edge from a to b is called a *covering edge*.

(a) What are the covering edges in the DAG in Figure 2?

(b) Let covering (D) be the subgraph of D consisting of only the covering edges. Suppose D is a finite DAG. Explain why covering (D) has the same positive walk relation as D.

Hint: Consider *longest* paths between a pair of vertices.

(c) Show that if two *DAG*'s have the same positive walk relation, then they have the same set of covering edges.

(d) Conclude that covering (D) is the *unique* DAG with the smallest number of edges among all digraphs with the same positive walk relation as D.

The following examples show that the above results don't work in general for digraphs with cycles.

(e) Describe two graphs with vertices $\{1, 2\}$ which have the same set of covering edges, but not the same positive walk relation (*Hint:* Self-loops.)

(f) (i) The *complete digraph* without self-loops on vertices $1, 2, 3$ has edges between every two distinct vertices. What are its covering edges?

(ii) What are the covering edges of the graph with vertices $1, 2, 3$ and edges $\langle 1 \to 2 \rangle$, $\langle 2 \to 3 \rangle$, $\langle 3 \to 1 \rangle$?

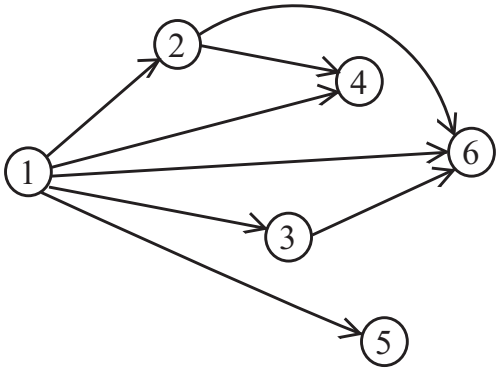


Figure 2 DAG with edges not needed in paths

(iii) What about their positive walk relations?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 8, Mon.

Problem 1.
For each of the binary relations below, state whether it is a strict partial order, a weak partial order, an equivalence relation, or none of these. If it is a partial order, state whether it is a linear order. If it is none, indicate which of the axioms for partial-order and equivalence relations it violates.

(a) The superset relation $\supseteq$ on the power set pow $\{1, 2, 3, 4, 5\}$.

(b) The relation between any two nonnegative integers a and b such that $a \equiv b \pmod 8$.

(c) The relation between propositional formulas G and H such that [G IMPLIES H] is valid.

(d) The relation between propositional formulas G and H such that [G IFF H] is valid.

(e) The relation 'beats' on Rock, Paper, and Scissors (for those who don't know the game Rock, Paper, Scissors, Rock beats Scissors, Scissors beats Paper, and Paper beats Rock).

(f) The empty relation on the set of real numbers.

(g) The identity relation on the set of integers.

(h) The divisibility relation on the integers, $\mathbb{Z}$.

Problem 2.
The proper subset relation, $\subset$, defines a strict partial order on the subsets of [1..6], that is, on pow([1..6]).

(a) What is the size of a maximal chain in this partial order? Describe one.

(b) Describe the largest antichain you can find in this partial order.

(c) What are the maximal and minimal elements? Are they maximum and minimum?

(d) Answer the previous part for the $\subset$ partial order on the set pow $[1..6] - \emptyset$.

Problem 3.
Let S be a sequence of n different numbers. A *subsequence* of S is a sequence that can be obtained by deleting elements of S.

For example, if

$$S = (6, 4, 7, 9, 1, 2, 5, 3, 8)$$

Then 647 and 7253 are both subsequences of S (for readability, we have dropped the parentheses and commas in sequences, so 647 abbreviates $(6, 4, 7)$, for example).

An *increasing subsequence* of S is a subsequence of whose successive elements get larger. For example, 1238 is an increasing subsequence of S. Decreasing subsequences are defined similarly; 641 is a decreasing subsequence of S.

(a) List all the maximum-length increasing subsequences of S, and all the maximum-length decreasing subsequences.

Now let A be the *set* of numbers in S. (So A is the integers [1..9] for the example above.) There are two straightforward linear orders for A. The first is numerical order where A is ordered by the $<$ relation. The second is to order the elements by which comes first in S; call this order $<_S$. So for the example above, we would have

$$6 <_S 4 <_S 7 <_S 9 <_S 1 <_S 2 <_S 5 <_S 3 <_S 8$$

Let $\prec$ be the product relation of the linear orders $<_s$ and $<$. That is, $\prec$ is defined by the rule

$$a \prec a' \quad ::= \quad a < a' \text{ AND } a <_S a'.$$

So $\prec$ is a partial order on A (Section 9.9 in the course textbook).

(b) Draw a diagram of the partial order, $\prec$, on A. What are the maximal and minimal elements?

(c) Explain the connection between increasing and decreasing subsequences of S, and chains and anti-chains under $\prec$.

(d) Prove that every sequence, S, of length n has an increasing subsequence of length greater than $\sqrt{n}$ or a decreasing subsequence of length at least $\sqrt{n}$.

Problem 4.

For any total function $f : A \to B$ define a relation $\equiv_f$ by the rule:

$$a \equiv_f a' \quad \text{iff} \quad f(a) = f(a'). \tag{1}$$

(a) Observe (and sketch a proof) that $\equiv_f$ is an equivalence relation on A.

(b) Prove that every equivalence relation, R, on a set, A, is equal to $\equiv_f$ for the function $f : A \to \text{pow}(A)$ defined as

$$f(a) ::= \{a' \in A \mid a \; R \; a'\}.$$

That is, $f(a) = R(a)$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 8, Wed.

Problem 1.
A researcher analyzing data on heterosexual sexual behavior in a group of m males and f females found that within the group, the male average number of female partners was 10% larger that the female average number of male partners.

(a) Comment on the following claim. "Since we're assuming that each encounter involves one man and one woman, the average numbers should be the same, so the males must be exaggerating."

(b) For what constant c is $m = c \cdot f$?

(c) The data shows that approximately 20% of the females were virgins, while only 5% of the males were. The researcher wonders how excluding virgins from the population would change the averages. If he knew graph theory, the researcher would realize that the nonvirgin male average number of partners will be $x(f/m)$ times the nonvirgin female average number of partners. What is x?

(d) For purposes of further research, it would be helpful to pair each female in the group with a unique male in the group. Explain why this is not possible.

Problem 2. **(a)** Prove that in every simple graph, there are an even number of vertices of odd degree.

(b) Conclude that at a party where some people shake hands, the number of people who shake hands an odd number of times is an even number.

(c) Call a sequence of people at the party a *handshake sequence* if each person in the sequence has shaken hands with the next person, if any, in the sequence.

Suppose George was at the party and has shaken hands with an odd number of people. Explain why, starting with George, there must be a handshake sequence ending with a different person who has shaken an odd number of hands.

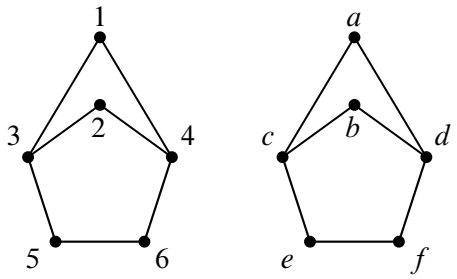
Problem 3.
List all the isomorphisms between the two graphs given in Figure 1. Explain why there are no others.

Problem 4.
Which of the items below are simple-graph properties preserved under isomorphism?

(a) The vertices can be numbered 1 through 7.

(b) There is a cycle that includes all the vertices.

(c) There are two degree 8 vertices.



(d) Two edges are of equal length.

(e) No matter which edge is removed, there is a path between any two vertices.

(f) There are two cycles that do not share any vertices.

(g) One vertex is a subset of another one.

(h) The graph can be pictured in a way that all the edges have the same length.

(i) The OR of two properties that are preserved under isomorphism.

(j) The negation of a property that is preserved under isomorphism.

Supplemental problem

Problem 5.
Let G be a digraph. The neighbors of a vertex v are the endpoints of the edges out of v. Since a digraph is formally the same as a binary relation on $V(G)$, the set of neighbors of v is simply the image, $G(v)$, of v under the relation G.

(a) Suppose f is an isomorphism from G to another digraph H. Prove that

$$f(G(v)) = H(f(v)).$$

Your proof should follow by simple reasoning using the definitions of isomorphism and image of a vertex under the edge relation—no pictures or handwaving.

Hint: Prove by a chain of iff's that

$$h \in H(f(v)) \text{ IFF } h \in f(G(v))$$

for every $h \in V(H)$.

(b) Conclude that if G and H are isomorphic graphs, then they have the same number of vertices of out-degree k, for all $k \in \mathbb{N}$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 8, Fri.

Problem 1.

A portion of a computer program consists of a sequence of calculations where the results are stored in variables, like this:

	Inputs:		a, b
Step 1.	c	$=$	$a + b$
2.	d	$=$	$a * c$
3.	e	$=$	$c + 3$
4.	f	$=$	$c - e$
5.	g	$=$	$a + f$
6.	h	$=$	$f + 1$
	Outputs:		d, g, h

A computer can perform such calculations most quickly if the value of each variable is stored in a *register*, a chunk of very fast memory inside the microprocessor. Programming language compilers face the problem of assigning each variable in a program to a register. Computers usually have few registers, however, so they must be used wisely and reused often. This is called the *register allocation* problem.

In the example above, variables a and b must be assigned different registers, because they hold distinct input values. Furthermore, c and d must be assigned different registers; if they used the same one, then the value of c would be overwritten in the second step and we d get the wrong answer in the third step. On the other hand, variables b and d may use the same register; after the first step, we no longer need b and can overwrite the register that holds its value. Also, f and h may use the same register; once $f + 1$ is evaluated in the last step, the register holding the value of f can be overwritten.

(a) Recast the register allocation problem as a question about graph coloring. What do the vertices correspond to? Under what conditions should there be an edge between two vertices? Construct the graph corresponding to the example above.

(b) Color your graph using as few colors as you can. Call the computer s registers $R1$, $R2$, etc. Describe the assignment of variables to registers implied by your coloring. How many registers do you need?

(c) Suppose that a variable is assigned a value more than once, as in the code snippet below:

$$\begin{aligned} &\ldots \\ t &= r + s \\ u &= t * 3 \\ t &= m - k \\ v &= t + u \\ &\ldots \end{aligned}$$

How might you cope with this complication?

Problem 2.

False Claim. *If every vertex in a graph has positive degree, then the graph is connected.*

(a) Prove that this Claim is indeed false by providing a counterexample.

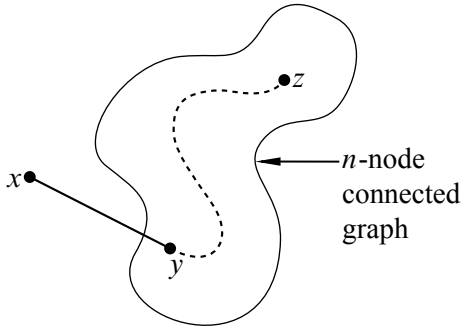
(b) Since the Claim is false, there must be a logical mistake in the following bogus proof. Pinpoint the *first* logical mistake (unjustified step) in the proof.

Bogus proof. We prove the Claim above by induction. Let $P(n)$ be the proposition that if every vertex in an n-vertex graph has positive degree, then the graph is connected.

Base cases: ($n \leq 2$). In a graph with 1 vertex, that vertex cannot have positive degree, so $P(1)$ holds vacuously.

$P(2)$ holds because there is only one graph with two vertices of positive degree, namely, the graph with an edge between the vertices, and this graph is connected.

Inductive step: We must show that $P(n)$ implies $P(n+1)$ for all $n \geq 2$. Consider an n-vertex graph in which every vertex has positive degree. By the assumption $P(n)$, this graph is connected; that is, there is a path between every pair of vertices. Now we add one more vertex x to obtain an $(n+1)$-vertex graph:



All that remains is to check that there is a path from x to every other vertex z. Since x has positive degree, there is an edge from x to some other vertex, y. Thus, we can obtain a path from x to z by going from x to y and then following the path from y to z. This proves $P(n+1)$.

By the principle of induction, $P(n)$ is true for all $n \geq 0$, which proves the Claim.

■

Problem 3.

In this problem, we examine an interesting connection between propositional logic and 3-colorings of certain special graphs. Consider the graph in Figure 1. We designate the vertices connected in the triangle on the left as *color-vertices*; since they form a triangle, they are forced to have different colors in any coloring of the graph. The colors assigned to the color-vertices will be called **T**, **F** and **N**. The dotted lines indicate edges to the color-vertex **N**.

(a) Prove that there exists a 3-coloring of the graph iff neither P nor Q are colored N.

(b) Argue that the graph in Figure 1 acts like a two-input OR-gate: a valid 3-coloring of the graph has the vertex labelled (P OR Q) colored **T** iff at least one of the vertices labelled P and Q are colored **T**.

(c) Changing the endpoint of one edge in Figure 1 will turn it into a two-input AND simulator. Explain.

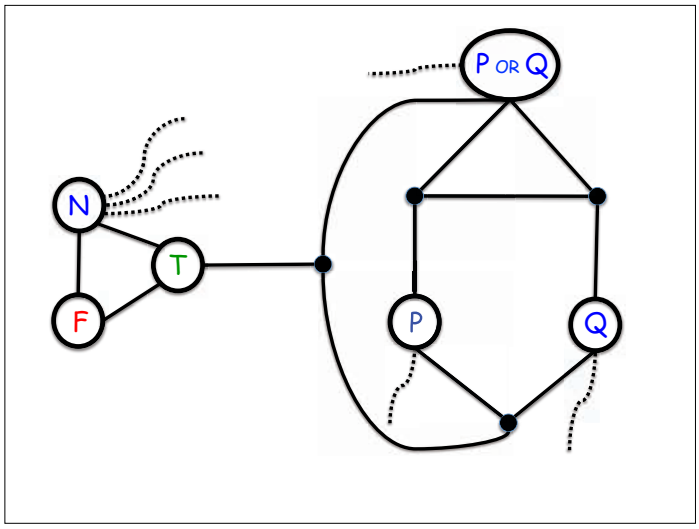


Figure 1 A 3-color OR-gate

Problem 4.

The n-dimensional *hypercube*, H_n, is a graph whose vertices are the binary strings of length n. Two vertices are adjacent if and only if they differ in exactly 1 bit. For example, in H_3, vertices `111` and `011` are adjacent because they differ only in the first bit, while vertices `101` and `011` are not adjacent because they differ at both the first and second bits.

(a) Verify that for any two vertices $x \neq y$ of H_3, there are 3 paths from x to y in H_3, such that, besides x and y, no two of those paths have a vertex in common.

(b) Conclude that the connectivity of H_3 is 3.

(c) Try extending your reasoning to H_4. (In fact, the connectivity of H_n is n for all $n \geq 1$. A proof appears in the problem solution.)

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 9, Mon.

Problem 1.
Let G be a 4×4 grid with vertical and horizontal edges between neighboring vertices. Formally,

$$V(G) = [0,3]^2 ::= \{(k,j) \mid 0 \le k, j \le 3\}.$$

Letting $h_{i,j}$ be the horizontal edge $\langle (i,j)\text{—}(i+1,j)\rangle$ and $v_{j,i}$ be the vertical edge $\langle (j,i)\text{—}(j,i+1)\rangle$ for $i \in [0,2]$, $j \in [0,3]$, the weights of these edges are

$$w(h_{i,j}) ::= \frac{4i+j}{100},$$
$$w(v_{j,i}) ::= 1 + \frac{i+4j}{100}.$$

(A picture of G would help; you might like to draw one.)

(a) Construct a minimum weight spanning tree (MST) for G by initially selecting the minimum weight edge, and then successively selecting the minimum weight edge that does not create a cycle with the previously selected edges. Stop when the selected edges form a spanning tree of G. (This is Kruskal's MST algorithm.) Explain how the "gray edge" Lemma 11.10.11 justifies this algorithm in the course textbook.

(b) Grow an MST for G starting with the tree consisting of the single vertex $(1,2)$ and successively adding the minimum weight edge with exactly one endpoint in the tree. Stop when the tree spans G. (This is Prim's MST algorithm.) Explain how the "gray edge" Lemma 11.10.11 justifies this algorithm in the course textbook.

(c) Grow an MST for G by treating the vertices $(0,0)$, $(0,3)$, $(2,3)$ as 1-vertex trees and then successively adding, for each tree in parallel, the minimum weight edge among the edges with one endpoint in the tree. Continue as long as there is no edge between two trees, then go back to applying the general gray edge method until the parallel trees merge to form a spanning tree of G. (This is 6.042's parallel MST algorithm.)

(d) Verify that you got the same MST each time.

Problem 2.
Prove that a graph is a tree iff it has a unique path between every two vertices.

Problem 3.
Let G be a weighted graph and suppose there is a unique edge $e \in E(G)$ with smallest weight, that is, $w(e) < w(f)$ for all edges $f \in E(G) - \{e\}$. Prove that any minimum weight spanning tree (MST) of G must include e.

Problem 4.

A simple graph, G, is said to have *width* w iff there is a way to list all its vertices so that each vertex is adjacent to at most w vertices that appear earlier in the list. All the graphs mentioned below are assumed to be finite.

(a) Prove that every graph with width one is a forest.

Hint: By induction, removing the last vertex.

(b) Prove that every finite tree has width one. Conclude that a graph is a forest iff it has width one.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 9, Wed.

Problem 1.
Four Students want separate assignments to four VI-A Companies. Here are their preference rankings:

Student	Companies
Albert:	HP, Bellcore, AT&T, Draper
Sarah:	AT&T, Bellcore, Draper, HP
Tasha:	HP, Draper, AT&T, Bellcore
Elizabeth:	Draper, AT&T, Bellcore, HP

Company	Students
AT&T:	Elizabeth, Albert, Tasha, Sarah
Bellcore:	Tasha, Sarah, Albert, Elizabeth
HP:	Elizabeth, Tasha, Albert, Sarah
Draper:	Sarah, Elizabeth, Tasha, Albert

(a) Use the Mating Ritual to find *two* stable assignments of Students to Companies.

(b) Describe a simple procedure to determine whether any given stable marriage problem has a unique solution, that is, only one possible stable matching.

Problem 2.
Suppose that Harry is one of the boys and Alice is one of the girls in the *Mating Ritual*. Which of the properties below are preserved invariants? Why?

a. Alice is the only girl on Harry's list.

b. There is a girl who does not have any boys serenading her.

c. If Alice is not on Harry's list, then Alice has a suitor that she prefers to Harry.

d. Alice is crossed off Harry's list, and Harry prefers Alice to anyone he is serenading.

e. If Alice is on Harry's list, then she prefers Harry to any suitor she has.

Problem 3.
A preserved invariant of the Mating Ritual is:

> For every girl, G, and every boy, B, if G is crossed off B's list, then G has a favorite suitor, and she prefers him over B.

Use the invariant to prove that the Mating Algorithm produces stable marriages. (Don't look up the proof in the Notes or slides.)

Problem 4.

The most famous application of stable matching was in assigning graduating medical students to hospital residencies. Each hospital has a preference ranking of students, and each student has a preference ranking of hospitals, but unlike finding stable marriages between an equal number of boys and girls, hospitals generally have differing numbers of available residencies, and the total number of residencies may not equal the number of graduating students.

Explain how to adapt the Stable Matching problem with an equal number of boys and girls to this more general situation. In particular, modify the definition of stable matching so it applies in this situation, and explain how to adapt the Mating Ritual to handle it.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 9, Fri.

Problem 1.
We begin with two large glasses. The first glass contains a pint of water, and the second contains a pint of wine. We pour 1/3 of a pint from the first glass into the second, stir up the wine/water mixture in the second glass, and then pour 1/3 of a pint of the mix back into the first glass and repeat this pouring back-and-forth process a total of n times.

(a) Describe a closed-form formula for the amount of wine in the first glass after n back-and-forth pourings.

(b) What is the limit of the amount of wine in each glass as n approaches infinity?

Problem 2.
An explorer is trying to reach the Holy Grail, which she believes is located in a desert shrine d days walk from the nearest oasis. In the desert heat, the explorer must drink continuously. She can carry at most 1 gallon of water, which is enough for 1 day. However, she is free to make multiple trips carrying up to a gallon each time to create water caches out in the desert.

For example, if the shrine were $2/3$ of a day's walk into the desert, then she could recover the Holy Grail after two days using the following strategy. She leaves the oasis with 1 gallon of water, travels $1/3$ day into the desert, caches $1/3$ gallon, and then walks back to the oasis—arriving just as her water supply runs out. Then she picks up another gallon of water at the oasis, walks $1/3$ day into the desert, tops off her water supply by taking the $1/3$ gallon in her cache, walks the remaining $1/3$ day to the shrine, grabs the Holy Grail, and then walks for $2/3$ of a day back to the oasis—again arriving with no water to spare.

But what if the shrine were located farther away?

(a) What is the most distant point that the explorer can reach and then return to the oasis, with no water precached in the desert, if she takes a total of only 1 gallon from the oasis?

(b) What is the most distant point the explorer can reach and still return to the oasis if she takes a total of only 2 gallons from the oasis? No proof is required; just do the best you can.

(c) The explorer will travel using a recursive strategy to go far into the desert and back, drawing a total of n gallons of water from the oasis. Her strategy is to build up a cache of $n-1$ gallons, plus enough to get home, a certain fraction of a day's distance into the desert. On the last delivery to the cache, instead of returning home, she proceeds recursively with her $n-1$ gallon strategy to go farther into the desert and return to the cache. At this point, the cache has just enough water left to get her home.

Prove that with n gallons of water, this strategy will get her $H_n/2$ days into the desert and back, where H_n is the nth Harmonic number:

$$H_n ::= \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Conclude that she can reach the shrine, however far it is from the oasis.

(d) Suppose that the shrine is $d = 10$ days walk into the desert. Use the asymptotic approximation $H_n \sim \ln n$ to show that it will take more than a million years for the explorer to recover the Holy Grail.

Problem 3.
Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a weakly decreasing function. Define

$$S ::= \sum_{i=1}^{n} f(i)$$

and

$$I ::= \int_1^n f(x)\,dx.$$

Prove that

$$I + f(n) \leq S \leq I + f(1).$$

(Proof by very clear picture is OK.)

Problem 4.
Sammy the Shark is a financial service provider who offers loans on the following terms.

- Sammy loans a client m dollars in the morning. This puts the client m dollars in debt to Sammy.
- Each evening, Sammy first charges a service fee which increases the client's debt by f dollars, and then Sammy charges interest, which multiplies the debt by a factor of p. For example, Sammy might charge a "modest" ten cent service fee and 1% interest rate per day, and then f would be 0.1 and p would be 1.01.

(a) What is the client's debt at the end of the first day?

(b) What is the client's debt at the end of the second day?

(c) Write a formula for the client's debt after d days and find an equivalent closed form.

(d) If you borrowed \$10 from Sammy for a year, how much would you owe him?

Supplemental problem

Problem 5.
You've seen this neat trick for evaluating a geometric sum:

$$\begin{aligned} S &= 1 + z + z^2 + \ldots + z^n \\ zS &= z + z^2 + \ldots + z^n + z^{n+1} \\ S - zS &= 1 - z^{n+1} \\ S &= \frac{1 - z^{n+1}}{1 - z} \text{ (where } z \neq 1) \end{aligned}$$

Use the same approach to find a closed-form expression for this sum:

$$T = 1z + 2z^2 + 3z^3 + \ldots + nz^n$$

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 10, Mon.

Problem 1.
Recall that for functions f, g on $\mathbb{N}$, $f = O(g)$ iff

$$\exists c \in \mathbb{N}\, \exists n_0 \in \mathbb{N}\, \forall n \geq n_0 \quad c \cdot g(n) \geq |f(n)|\,. \tag{1}$$

For each pair of functions below, determine whether $f = O(g)$ and whether $g = O(f)$. In cases where one function is O() of the other, indicate the *smallest nonnegative integer*, c, and for that smallest c, the *smallest corresponding nonnegative integer* n_0 ensuring that condition (1) applies.

(a) $f(n) = n^2, g(n) = 3n$.

(b) $f(n) = (3n - 7)/(n + 4), g(n) = 4$

(c) $f(n) = 1 + (n \sin(n\pi/2))^2, g(n) = 3n$

Problem 2.

(a) Indicate which of the following asymptotic relations below on the set of nonnegative real-valued functions are equivalence relations (**E**), strict partial orders (**S**), weak partial orders (**W**), or *none* of the above (**N**).

- $f \sim g$, the "asymptotically equal" relation.
- $f = o(g)$, the "little Oh" relation.
- $f = O(g)$, the "big Oh" relation.
- $f = \Theta(g)$, the "Theta" relation.
- $f = O(g)$ AND NOT$(g = O(f))$.

(b) Indicate the implications among the assertions in part (a). For example,

$$f = o(g) \text{ IMPLIES } f = O(g).$$

Problem 3.

False Claim.

$$2^n = O(1). \tag{2}$$

Explain why the claim is false. Then identify and explain the mistake in the following bogus proof.

Bogus proof. The proof is by induction on n where the induction hypothesis, $P(n)$, is the assertion (2).

base case: $P(0)$ holds trivially.

inductive step: We may assume $P(n)$, so there is a constant $c > 0$ such that $2^n \leq c \cdot 1$. Therefore,

$$2^{n+1} = 2 \cdot 2^n \leq (2c) \cdot 1,$$

which implies that $2^{n+1} = O(1)$. That is, $P(n+1)$ holds, which completes the proof of the inductive step.

We conclude by induction that $2^n = O(1)$ for all n. That is, the exponential function is bounded by a constant.

■

Supplemental problems

Problem 4.
Assign true or false for each statement and prove it.

- $n^2 \sim n^2 + n$
- $3^n = O\left(2^n\right)$
- $n^{\sin(n\pi/2)+1} = o\left(n^2\right)$
- $n = \Theta\left(\dfrac{3n^3}{(n+1)(n-1)}\right)$

Problem 5.
Give an elementary proof (without appealing to Stirling's formula) that $\log(n!) = \Theta(n \log n)$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 10, Fri.

Problem 1. (a) How many of the billion numbers in the range from 1 to 10^9 contain the digit 1? (*Hint:* How many don't?)

(b) There are 20 books arranged in a row on a shelf. Describe a bijection between ways of choosing 6 of these books so that no two adjacent books are selected and 15-bit strings with exactly 6 ones.

Problem 2.
An n-vertex *numbered tree* is a tree whose vertex set is $\{1, 2, \ldots, n\}$ for some $n > 2$. We define the *code* of the numbered tree to be a sequence of $n - 2$ integers from 1 to n obtained by the following recursive process:[1]

> If there are more than two vertices left, write down the *father* of the largest leaf, delete this *leaf*, and continue this process on the resulting smaller tree. If there are only two vertices left, then stop —the code is complete.

For example, the codes of a couple of numbered trees are shown in the Figure 1.

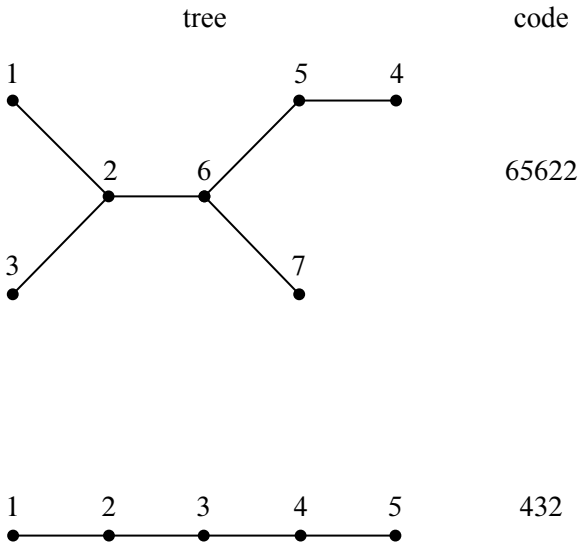


Figure 1

(a) Describe a procedure for reconstructing a numbered tree from its code.

(b) Conclude there is a bijection between the n-vertex numbered trees and $\{1, \ldots, n\}^{n-2}$, and state how many n-vertex numbered trees there are.

Problem 3.

(a) Let $\mathcal{S}_{n,k}$ be the possible nonnegative integer solutions to the inequality

$$x_1 + x_2 + \cdots + x_k \leq n. \tag{1}$$

That is

$$\mathcal{S}_{n,k} ::= \{(x_1, x_2, \ldots, x_k) \in \mathbb{N}^k \mid (1) \text{ is true}\}.$$

Describe a bijection between $\mathcal{S}_{n,k}$ and the set of binary strings with n zeroes and k ones.

(b) Let $\mathcal{L}_{n,k}$ be the length k weakly increasing sequences of nonnegative integers $\leq n$. That is

$$\mathcal{L}_{n,k} ::= \{(y_1, y_2, \ldots, y_k) \in \mathbb{N}^k \mid y_1 \leq y_2 \leq \cdots \leq y_k \leq n\}.$$

Describe a bijection between $\mathcal{L}_{n,k}$ and $\mathcal{S}_{n,k}$.

Supplemental problem

Problem 4.
Let X and Y be finite sets.

(a) How many binary relations from X to Y are there?

(b) Define a bijection between the set $[X \to Y]$ of all total functions from X to Y and the set $Y^{|X|}$. (Recall Y^n is the Cartesian product of Y with itself n times.) Based on that, what is $|[X \to Y]|$?

(c) Using the previous part, how many *functions*, not necessarily total, are there from X to Y? How does the fraction of functions vs. total functions grow as the size of X grows? Is it $O(1)$, $O(|X|)$, $O(2^{|X|})$,...?

(d) Show a bijection between the powerset, $\text{pow}(X)$, and the set $[X \to \{0,1\}]$ of 0-1-valued total functions on X.

(e) Let X be a set of size n and B_X be the set of all bijections from X to X. Describe a bijection from B_X to the set of permutations of X.[2] This implies that there are how may bijections from X to X?

[2] A sequence in which all the elements of a set X appear exactly once is called a *permutation* of X.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 11, Wed.

Problem 1.
Your class tutorial has 12 students, who are supposed to break up into 4 groups of 3 students each. Your Teaching Assistant (TA) has observed that the students waste too much time trying to form balanced groups, so he decided to pre-assign students to groups and email the group assignments to his students.

(a) Your TA has a list of the 12 students in front of him, so he divides the list into consecutive groups of 3. For example, if the list is ABCDEFGHIJKL, the TA would define a sequence of four groups to be $(\{A, B, C\}, \{D, E, F\}, \{G, H, I\}, \{J, K, L\})$. This way of forming groups defines a mapping from a list of twelve students to a sequence of four groups. This is a k-to-1 mapping for what k?

(b) A group assignment specifies which students are in the same group, but not any order in which the groups should be listed. If we map a sequence of 4 groups,

$$(\{A, B, C\}, \{D, E, F\}, \{G, H, I\}, \{J, K, L\}),$$

into a group assignment

$$\{\{A, B, C\}, \{D, E, F\}, \{G, H, I\}, \{J, K, L\}\},$$

this mapping is j-to-1 for what j?

(c) How many group assignments are possible?

(d) In how many ways can $3n$ students be broken up into n groups of 3?

Problem 2. **(a)** There are 30 books arranged in a row on a shelf. In how many ways can eight of these books be selected so that there are at least two unselected books between any two selected books?

(b) How many nonnegative integer solutions are there for the following equality?

$$x_1 + x_2 + \cdots + x_m = k. \tag{1}$$

(c) How many nonnegative integer solutions are there for the following inequality?

$$x_1 + x_2 + \cdots + x_m \leq k. \tag{2}$$

(d) How many length m weakly increasing sequences of nonnegative integers $\leq k$ are there?

Problem 3.
The Tao of BOOKKEEPER: we seek enlightenment through contemplation of the word $BOOKKEEPER$.

(a) In how many ways can you arrange the letters in the word $POKE$?

(b) In how many ways can you arrange the letters in the word BO_1O_2K? Observe that we have subscripted the O's to make them distinct symbols.

(c) Suppose we map arrangements of the letters in BO_1O_2K to arrangements of the letters in $BOOK$ by erasing the subscripts. Indicate with arrows how the arrangements on the left are mapped to the arrangements on the right.

O_2BO_1K
KO_2BO_1
O_1BO_2K
KO_1BO_2
BO_1O_2K
BO_2O_1K
...

$BOOK$
$OBOK$
$KOBO$
...

(d) What kind of mapping is this, young grasshopper?

(e) In light of the Division Rule, how many arrangements are there of $BOOK$?

(f) Very good, young master! How many arrangements are there of the letters in $KE_1E_2PE_3R$?

(g) Suppose we map each arrangement of $KE_1E_2PE_3R$ to an arrangement of $KEEPER$ by erasing subscripts. List all the different arrangements of $KE_1E_2PE_3R$ that are mapped to $REPEEK$ in this way.

(h) What kind of mapping is this?

(i) So how many arrangements are there of the letters in $KEEPER$?

Now you are ready to face the BOOKKEEPER!

(j) How many arrangements of $BO_1O_2K_1K_2E_1E_2PE_3R$ are there?

(k) How many arrangements of $BOOK_1K_2E_1E_2PE_3R$ are there?

(l) How many arrangements of $BOOKKE_1E_2PE_3R$ are there?

(m) How many arrangements of $BOOKKEEPER$ are there?

Remember well what you have learned: subscripts on, subscripts off.
This is the Tao of Bookkeeper.

(n) How many arrangements of $VOODOODOLL$ are there?

(o) How many length 52 sequences of digits contain exactly 17 two's, 23 fives, and 12 nines?

Problem 4.
Find the coefficients of

(a) x^5 in $(1+x)^{11}$

(b) x^8y^9 in $(3x+2y)^{17}$

(c) a^6b^6 in $(a^2+b^3)^5$

Problem 5.
Solve the following counting problems. Define an appropriate mapping (bijective or k-to-1) between a set whose size you know and the set in question.

(a) An independent living group is hosting nine new candidates for membership. Each candidate must be assigned a task: 1 must wash pots, 2 must clean the kitchen, 3 must clean the bathrooms, 1 must clean the common area, and 2 must serve dinner. Write a multinomial coefficient for the number of ways this can be done.

(b) How many nonnegative integers less than 1,000,000 have exactly one digit equal to 9 and have a sum of digits equal to 17?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 11, Fri.

Problem 1.
Solve the following problems using the pigeonhole principle. For each problem, try to identify the *pigeons*, the *pigeonholes*, and a *rule* assigning each pigeon to a pigeonhole.

(a) In a certain Institute of Technology, every ID number starts with a 9. Suppose that each of the 75 students in a class sums the nine digits of their ID number. Explain why two people must arrive at the same sum.

(b) In every set of 100 integers, there exist two whose difference is a multiple of 37.

(c) For any five points inside a unit square (not on the boundary), there are two points at distance *less than* $1/\sqrt{2}$.

(d) Show that if $n+1$ numbers are selected from $\{1, 2, 3, \ldots, 2n\}$, two must be consecutive, that is, equal to k and $k+1$ for some k.

Problem 2.
To ensure password security, a company requires their employees to choose a password. A length 10 word containing each of the characters:

a, d, e, f, i, l, o, p, r, s,

is called a *cword*. A password can be a cword which does not contain any of the subwords "fails", "failed", or "drop."

For example, the following two words are passwords: adefiloprs, srpolifeda,
but the following three cwords are not: a**drop**eflis, **failedrop**s, **drop**e**fails**.

(a) How many cwords contain the subword "drop"?

(b) How many cwords contain both "drop" and "fails"?

(c) Use the Inclusion-Exclusion Principle to find a simple arithmetic formula involving factorials for the number of passwords.

Problem 3.
How many paths are there from point $(0, 0)$ to $(50, 50)$ if each step along a path increments one coordinate and leaves the other unchanged? How many are there when there are impassable boulders sitting at points $(10, 11)$ and $(21, 20)$? (You do not have to calculate the number explicitly; your answer may be an expression involving binomial coefficients.)

Hint: Inclusion-Exclusion.

Supplemental problems

Problem 4. (a) Prove that every positive integer divides a number such as 70, 700, 7770, 77000, whose decimal representation consists of one or more 7's followed by one or more 0's.

Hint: $7, 77, 777, 7777, \ldots$

(b) Conclude that if a positive number is not divisible by 2 or 5, then it divides a number whose decimal representation is all 7's.

Problem 5.
Show that for any set of 201 positive integers less than 300, there must be two whose quotient is a power of three (with no remainder).

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 12, Mon.

Problem 1.
[The Four-Door Deal]

Let's see what happens when *Let's Make a Deal* is played with **four** doors. A prize is hidden behind one of the four doors. Then the contestant picks a door. Next, the host opens an unpicked door that has no prize behind it. The contestant is allowed to stick with their original door or to switch to one of the two unopened, unpicked doors. The contestant wins if their final choice is the door hiding the prize.

Let's make the same assumptions as in the original problem:

1. The prize is equally likely to be behind each door.
2. The contestant is equally likely to pick each door initially, regardless of the prize's location.
3. The host is equally likely to reveal each door that does not conceal the prize and was not selected by the player.

Use The Four Step Method to find the following probabilities. The tree diagram may become awkwardly large, in which case just draw enough of it to make its structure clear. Also, indicate the set of outcomes in each of the events below. A numerical probability without a demonstration of the Method is not a satisfactory answer.

(a) Contestant Stu, a sanitation engineer from Trenton, New Jersey, stays with his original door. What is the probability that Stu wins the prize?

(b) Contestant Zelda, an alien abduction researcher from Helena, Montana, switches to one of the remaining two doors with equal probability. What is the probability that Zelda wins the prize?

Problem 2.

Suppose there is a system, built by Caltech graduates, with n components. We know from past experience that any particular component will fail in a given year with probability p. That is, letting F_i be the event that the ith component fails within one year, we have

$$\Pr[F_i] = p$$

for $1 \leq i \leq n$. The *system* will fail if *any one* of its components fails. What can we say about the probability that the system will fail within one year?

Let F be the event that the system fails within one year. Without any additional assumptions, we can't get an exact answer for $\Pr[F]$. However, we can give useful upper and lower bounds, namely,

$$p \leq \Pr[F] \leq np. \tag{1}$$

We may as well assume $p < 1/n$, since the upper bound is trivial otherwise. For example, if $n = 100$ and $p = 10^{-5}$, we conclude that there is at most one chance in 1000 of system failure within a year and at least one chance in 100,000.

Let's model this situation with the sample space $\mathcal{S} ::= \text{pow}([1, n])$ whose outcomes are subsets of positive integers $\leq n$, where $s \in \mathcal{S}$ corresponds to the indices of exactly those components that fail within one year. For example, $\{2, 5\}$ is the outcome that the second and fifth components failed within a year and none of the other components failed. So the outcome that the system did not fail corresponds to the empty set, $\emptyset$.

(a) Show that the probability that the system fails could be as small as p by describing appropriate probabilities for the outcomes. Make sure to verify that the sum of your outcome probabilities is 1.

(b) Show that the probability that the system fails could actually be as large as np by describing appropriate probabilities for the outcomes. Make sure to verify that the sum of your outcome probabilities is 1.

(c) Prove inequality (1).

Problem 3.

To determine which of two people gets a prize, a coin is flipped twice. If the flips are a Head and then a Tail, the first player wins. If the flips are a Tail and then a Head, the second player wins. However, if both coins land the same way, the flips don't count and the whole process starts over.

Assume that on each flip, a Head comes up with probability p, regardless of what happened on other flips. Use the four step method to find a simple formula for the probability that the first player wins. What is the probability that neither player wins?

Hint: The tree diagram and sample space are infinite, so you're not going to finish drawing the tree. Try drawing only enough to see a pattern. Summing all the winning outcome probabilities directly is cumbersome. However, a neat trick solves this problem—and many others. Let s be the sum of all winning outcome probabilities in the whole tree. Notice that *you can write the sum of all the winning probabilities in certain subtrees as a function of s*. Use this observation to write an equation in s and then solve.

Problem 4.

Prove the following probabilistic inequality, referred to as the *Union Bound*.

Let $A_1, A_2, \ldots, A_n, \ldots$ be events. Then

$$\Pr\left[\bigcup_{n\in\mathbb{N}} A_n\right] \leq \sum_{n\in\mathbb{N}} \Pr[A_n].$$

Hint: Replace the A_n's by pairwise disjoint events and use the Sum Rule.

Supplemental problems

Problem 5.

Here are some handy rules for reasoning about probabilities that all follow directly from the Disjoint Sum Rule. Prove them.

$$\Pr[A - B] = \Pr[A] - \Pr[A \cap B] \qquad \text{(Difference Rule)}$$

$$\Pr[\overline{A}] = 1 - \Pr[A] \qquad \text{(Complement Rule)}$$

$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] \qquad \text{(Inclusion-Exclusion)}$$

$$\Pr[A \cup B] \leq \Pr[A] + \Pr[B] \qquad \text{(2-event Union Bound)}$$

$$\text{If } A \subseteq B, \text{ then } \Pr[A] \leq \Pr[B] \qquad \text{(Monotonicity)}$$

Problem 6.
The New York Yankees and the Boston Red Sox are playing a two-out-of-three series. In other words, they play until one team has won two games. Then that team is declared the overall winner and the series ends. Assume that the Red Sox win each game with probability $3/5$, regardless of the outcomes of previous games.

Answer the questions below using the four step method. You can use the same tree diagram for all three problems.

(a) What is the probability that a total of 3 games are played?

(b) What is the probability that the winner of the series loses the first game?

(c) What is the probability that the *correct* team wins the series?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 12, Wed.

Problem 1.
There is an unpleasant degenerative disease called Beaver Fever which causes people to tell unrelentingl math jokes in social settings, believing other people would think they're funny. Fortunately, Beaver Fever is rare, afflicting only about 1 in 1000 people. Doctor Meyer has a fairly reliable diagnostic test to determine who is going to suffer from this disease:

- If a person will suffer from Beaver Fever, the probability that Dr. Meyer diagnoses this is 0.99.
- If a person will not suffer from Beaver Fever, the probability that Dr. Meyer diagnoses this is 0.97.

Let B be the event that a randomly chosen person will suffer Beaver Fever, and Y be the event that Dr. Meyer's diagnosis is "Yes, this person will suffer from Beaver Fever," with $\overline{B}$ and $\overline{Y}$ being the complements of these events.

(a) The description above explicitly gives the values of the following quantities. What are their values?

$$\Pr[B] \qquad \Pr\left[Y \mid B\right] \qquad \Pr\left[\overline{Y} \mid \overline{B}\right]$$

(b) Write formulas for $\Pr[\overline{B}]$ and $\Pr\left[Y \mid \overline{B}\right]$ solely in terms of the explicitly given quantities in part (a)—literally use their expressions, not their numeric values.

(c) Write a formula for the probability that Dr. Meyer says a person will suffer from Beaver Fever solely in terms of $\Pr[B]$, $\Pr[\overline{B}]$, $\Pr\left[Y \mid B\right]$ and $\Pr\left[Y \mid \overline{B}\right]$.

(d) Write a formula solely in terms of the expressions given in part (a) for the probability that a person will suffer Beaver Fever given that Doctor Meyer says they will. Then calculate the numerical value of the formula.

Suppose there was a vaccine to prevent Beaver Fever, but the vaccine was expensive or slightly risky itself. If you were sure you were going to suffer from Beaver Fever, getting vaccinated would be worthwhile, but by part (d), even if Dr. Meyer diagnosed you as a future sufferer of Beaver Fever, the probability you actually will suffer Beaver Fever remains low (less than 1/30).

In this case, you might sensibly decide not to be vaccinated (after all, Beaver Fever is not *that* bad an affliction). So the diagnostic test serves no purpose in your case—you may as well not have bothered to get diagnosed. Even so, the test may be useful:

(e) Suppose Dr. Meyer had enough vaccine to treat 2% of the population. If he randomly chose people to vaccinate, he could expect to vaccinate only 2% of the people who needed it. But by testing everyone and only vaccinating those diagnosed as future sufferers, he can expect to vaccinate a much larger fraction people who were going to suffer from Beaver Fever. Estimate this fraction.

Problem 2.
There are three prisoners in a maximum-security prison for fictional villains: the Evil Wizard Voldemort, the Dark Lord Sauron, and Little Bunny Foo-Foo. The parole board has declared that it will release two of the three, chosen uniformly at random, but has not yet released their names. Naturally, Sauron figures that he will be released to his home in Mordor, where the shadows lie, with probability $2/3$.

A guard offers to tell Sauron the name of one of the other prisoners who will be released (either Voldemort or Foo-Foo). If the guard has a choice of naming either Voldemort or Foo-Foo (because both are to be released), he names one of the two with equal probability.

Sauron knows the guard to be a truthful fellow. However, Sauron declines this offer. He reasons that knowing what the guards says will reduce his chances, so he is better off not knowing. For example, if the guard says, "Little Bunny Foo-Foo will be released", then his own probability of release will drop to $1/2$ because he will then know that either he or Voldemort will also be released, and these two events are equally likely.

Dark Lord Sauron has made a typical mistake when reasoning about conditional probability. Using a tree diagram and the four-step method, **explain his mistake**. What is the probability that Sauron is released given that the guard says Foo-Foo is released?

Hint: Define the events S, F, and "F" as follows:

$$\begin{aligned} \text{“}F\text{”} &= \text{Guard says Foo-Foo is released} \\ F &= \text{Foo-Foo is released} \\ S &= \text{Sauron is released} \end{aligned}$$

Problem 3.
There are two decks of cards. One is complete, but the other is missing the Ace of spades. Suppose you pick one of the two decks with equal probability and then select a card from that deck uniformly at random. What is the probability that you picked the complete deck, given that you selected the eight of hearts? Use the four-step method and a tree diagram.

Supplemental problem

Problem 4.
Suppose you repeatedly flip a fair coin until you see the sequence `HTT` or `HHT`. What is the probability you see the sequence `HTT` first?

Hint: Try to find the probability that `HHT` comes before `HTT` conditioning on whether you first toss an `H` or a `T`. The answer is not $1/2$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 12, Fri.

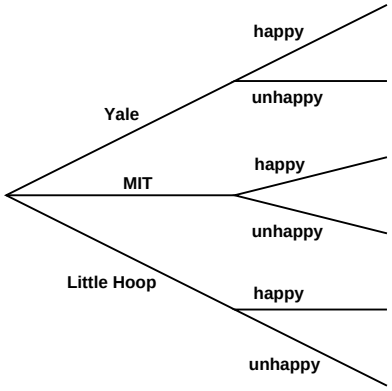
Problem 1.
Sally Smart just graduated from high school. She was accepted to three reputable colleges.

- With probability $4/12$, she attends Yale.
- With probability $5/12$, she attends MIT.
- With probability $3/12$, she attends Little Hoop Community College.

Sally is either happy or unhappy in college.

- If she attends Yale, she is happy with probability $4/12$.
- If she attends MIT, she is happy with probability $7/12$.
- If she attends Little Hoop, she is happy with probability $11/12$.

(a) A tree diagram to help Sally project her chance at happiness is shown below. On the diagram, fill in the edge probabilities, and at each leaf write the probability of the corresponding outcome.



(b) What is the probability that Sally is happy in college?

(c) What is the probability that Sally attends Yale, given that she is happy in college?

(d) Show that the event that Sally attends Yale **is not** independent of the event that she is happy.

(e) Show that the event that Sally attends MIT **is** independent of the event that she is happy.

Problem 2.
Suppose you flip three fair, mutually independent coins. Define the following events:

- Let A be the event that *the first* coin is heads.
- Let B be the event that *the second* coin is heads.
- Let C be the event that *the third* coin is heads.
- Let D be the event that *an even number of* coins are heads.

(a) Use the four step method to determine the probability space for this experiment and the probability of each of A, B, C, D.

(b) Show that these events are not mutually independent.

(c) Show that they are 3-way independent.

Problem 3.
Graphs, Logic & Probability

Let G be an undirected simple graph with $n > 3$ vertices. Let $E(x, y)$ mean that G has an edge between vertices x and y, and let $P(x, y)$ mean that there is a length 2 walk in G between x and y.

(a) Write a predicate-logic formula defining $P(x, y)$ in terms of $E(x, y)$.

For the following parts (b)–(d), let V be a fixed set of $n > 3$ vertices, and let G be a graph with these vertices constructed randomly as follows: for all distinct vertices $x, y \in V$, independently include edge $\langle x—y\rangle$ as an edge of G with probability p. In particular, $\Pr[E(x, y)] = p$ for all $x \neq y$.

(b) For distinct vertices w, x, y and z in V, circle the event pairs that are independent.

1. $E(w, x)$ versus $E(x, y)$
2. $[E(w, x)$ AND $E(w, y)]$ versus $[E(z, x)$ AND $E(z, y)]$
3. $E(x, y)$ versus $P(x, y)$
4. $P(w, x)$ versus $P(x, y)$
5. $P(w, x)$ versus $P(y, z)$

(c) Write a simple formula in terms of n and p for $\Pr[\text{NOT } P(x, y)]$, for distinct vertices x and y in V.

Hint: Use part (b), item 2.

(d) What is the probability that two distinct vertices x and y lie on a three-cycle in G? Answer with a simple expression in terms of p and r, where $r ::= \Pr[\text{NOT}(P(x, y))]$ is the correct answer to part (c).

Hint: Express x and y being on a three-cycle as a simple formula involving $E(x, y)$ and $P(x, y)$.

Supplemental Problem

Problem 4.
Let A, B, C be events. For each of the following statements, prove it or give a counterexample.

(a) If A is independent of B, then A is also independent of $\overline{B}$.

(b) If A is independent of B, and A is independent of C, then A is independent of $B \cap C$.

Hint: Choose A, B, C pairwise but not 3-way independent.

(c) If A is independent of B, and A is independent of C, then A is independent of $B \cup C$.

Hint: Part (b).

(d) If A is independent of B, and A is independent of C, and A is independent of $B \cap C$, then A is independent of $B \cup C$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 13, Mon.

Guess the Bigger Number Game

Team 1:

- Write two different integers between 0 and 7 on separate pieces of paper.
- Put the papers face down on a table.

Team 2:

- Turn over one paper and look at the number on it.
- Either stick with this number or switch to the other (unseen) number.

Team 2 wins if it chooses the larger number; else, Team 1 wins.

Problem 1.
The analysis given before class implies that Team 2 has a strategy that wins 4/7 of the time no matter how Team 1 plays. Can Team 2 do better? The answer is "no," because Team 1 has a strategy that guarantees that it wins at least 3/7 of the time, no matter how Team 2 plays. Describe such a strategy for Team 1 and explain why it works.

Problem 2.
Let I_A and I_B be the indicator variables for events A and B. Prove that I_A and I_B are independent iff A and B are independent.

Hint: Let $A^1 ::= A$ and $A^0 ::= \overline{A}$, so the event $[I_A = c]$ is the same as A^c for $c \in \{0,1\}$; likewise for B^1, B^0.

Problem 3.
Let $R_1, R_2, \ldots, R_m$, be mutually independent random variables with uniform distribution on $[1,n]$. Let $M ::= \max\{R_i \mid i \in [1,m]\}$.

(a) Write a formula for $\text{PDF}_M(1)$.

(b) More generally, write a formula for $\Pr[M \leq k]$.

(c) For $k \in [1,n]$, write a formula for $\text{PDF}_M(k)$ in terms of expressions of the form "$\Pr[M \leq j]$" for $j \in [1,n]$.

Problem 4.
Suppose you have a biased coin that has probability p of flipping heads. Let J be the number of heads in n independent coin flips. So J has the general binomial distribution:

$$\mathrm{PDF}_J(k) = \binom{n}{k} p^k q^{n-k}$$

where $q ::= 1 - p$.

(a) Show that

$$\begin{aligned} \mathrm{PDF}_J(k-1) &< \mathrm{PDF}_J(k) && \text{for } k < np + p, \\ \mathrm{PDF}_J(k-1) &> \mathrm{PDF}_J(k) && \text{for } k > np + p. \end{aligned}$$

(b) Conclude that the maximum value of PDF_J is asymptotically equal to

$$\frac{1}{\sqrt{2\pi npq}}.$$

Hint: For the asymptotic estimate, it's ok to assume that np is an integer, so by part (a), the maximum value is $\mathrm{PDF}_J(np)$. Use Stirling's Formula.

Supplemental problem

Problem 5.
You have just been married and you both want to have children. Of course, any child is a blessing, but your spouse prefers girls, so you decide to keep having children until you have a girl. In other words, if your 1st child is a girl, you'll stop there. If it's a boy, then you'll have a 2nd child, too. If that one is a girl, you'll stop there. Otherwise, you'll have a 3rd child, and so on. Assume that you will never abandon this ingenious plan and that every child is equally likely to be a boy or a girl, independently of the number of its brothers so far. Let B be the *boys* that you will eventually have to put up with to enjoy the company of your beloved daughter.

(a) For $i = 0, 1, 2, \ldots$, what is the value of $\mathrm{PDF}_B(i)$?

(b) For $i = 0, 1, 2, \ldots$, what is the value of $\mathrm{CDF}_B(i)$?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 8, Wed.

Problem 1.
A researcher analyzing data on heterosexual sexual behavior in a group of m males and f females found that within the group, the male average number of female partners was 10% larger that the female average number of male partners.

(a) Comment on the following claim. "Since we're assuming that each encounter involves one man and one woman, the average numbers should be the same, so the males must be exaggerating."

(b) For what constant c is $m = c \cdot f$?

(c) The data shows that approximately 20% of the females were virgins, while only 5% of the males were. The researcher wonders how excluding virgins from the population would change the averages. If he knew graph theory, the researcher would realize that the nonvirgin male average number of partners will be $x(f/m)$ times the nonvirgin female average number of partners. What is x?

(d) For purposes of further research, it would be helpful to pair each female in the group with a unique male in the group. Explain why this is not possible.

Problem 2. **(a)** Prove that in every simple graph, there are an even number of vertices of odd degree.

(b) Conclude that at a party where some people shake hands, the number of people who shake hands an odd number of times is an even number.

(c) Call a sequence of people at the party a *handshake sequence* if each person in the sequence has shaken hands with the next person, if any, in the sequence.

Suppose George was at the party and has shaken hands with an odd number of people. Explain why, starting with George, there must be a handshake sequence ending with a different person who has shaken an odd number of hands.

Problem 3.
List all the isomorphisms between the two graphs given in Figure 1. Explain why there are no others.

Problem 4.
Which of the items below are simple-graph properties preserved under isomorphism?

(a) The vertices can be numbered 1 through 7.

(b) There is a cycle that includes all the vertices.

(c) There are two degree 8 vertices.

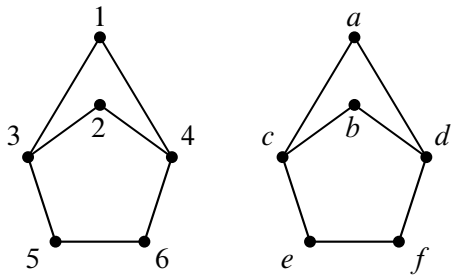


Figure 1 Graphs with several isomorphisms

(d) Two edges are of equal length.

(e) No matter which edge is removed, there is a path between any two vertices.

(f) There are two cycles that do not share any vertices.

(g) One vertex is a subset of another one.

(h) The graph can be pictured in a way that all the edges have the same length.

(i) The OR of two properties that are preserved under isomorphism.

(j) The negation of a property that is preserved under isomorphism.

Supplemental problem

Problem 5.
Let G be a digraph. The neighbors of a vertex v are the endpoints of the edges out of v. Since a digraph is formally the same as a binary relation on $V(G)$, the set of neighbors of v is simply the image, $G(v)$, of v under the relation G.

(a) Suppose f is an isomorphism from G to another digraph H. Prove that

$$f(G(v)) = H(f(v)).$$

Your proof should follow by simple reasoning using the definitions of isomorphism and image of a vertex under the edge relation—no pictures or handwaving.

Hint: Prove by a chain of iff's that

$$h \in H(f(v)) \text{ IFF } h \in f(G(v))$$

for every $h \in V(H)$.

(b) Conclude that if G and H are isomorphic graphs, then they have the same number of vertices of out-degree k, for all $k \in \mathbb{N}$.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 13, Fri.

Problem 1.
A herd of cows is stricken by an outbreak of *cold cow disease*. The disease lowers a cow's body temperature from normal levels, and a cow will die if its temperature goes below 90 degrees F. The disease epidemic is so intense that it lowered the average temperature of the herd to 85 degrees. Body temperatures as low as 70 degrees, **but no lower**, were actually found in the herd.

(a) Use Markov's Bound to prove that at most 3/4 of the cows could survive.

(b) Suppose there are 400 cows in the herd. Show that the bound from part (a) is the best possible by giving an example set of temperatures for the cows so that the average herd temperature is 85 and 3/4 of the cows will have a high enough temperature to survive.

(c) Notice that the results of part (b) are purely arithmetic facts about averages, not about probabilities. But you verified the claim in part (a) by applying Markov's bound on the deviation of a random variable. Justify this approach by regarding the temperature, T, of a cow as a random variable. Carefully specify the probability space on which T is defined: what are the outcomes? what are their probabilities? Explain the precise connection between properties of T and average herd temperature that justifies the application of Markov's Bound.

Problem 2.
A gambler plays 120 hands of draw poker, 60 hands of black jack, and 20 hands of stud poker per day. He wins a hand of draw poker with probability 1/6, a hand of black jack with probability 1/2, and a hand of stud poker with probability 1/5.

(a) What is the expected number of hands the gambler wins in a day?

(b) What would the Markov bound be on the probability that the gambler will win at least 108 hands on a given day?

(c) Assume that the outcomes of the card games are *pairwise*, but possibly *not* mutually, independent. What is the variance of the number of hands won per day? You may answer with a numerical expression that is not completely evaluated.

(d) What would the Chebyshev bound be on the probability that the gambler will win at least 108 hands on a given day? You may answer with a numerical expression that is not completely evaluated.

Problem 3.
The hat-check staff has had a long day serving at a party, and at the end of the party they simply return the n checked hats in a random way such that the probability that any particular person gets their own hat back is $1/n$.

Let X_i be the indicator variable for the ith person getting their own hat back. Let S_n be the total number of people who get their own hat back.

(a) What is the expected number of people who get their own hat back?

(b) Write a simple formula for $\text{Ex}[X_i \cdot X_j]$ for $i \neq j$.

Hint: What is $\Pr[X_j = 1 \mid X_i = 1]$?

(c) Explain why you cannot use the variance of sums formula to calculate $\text{Var}[S_n]$.

(d) Show that $\text{Ex}[(S_n)^2] = 2$. *Hint:* $(X_i)^2 = X_i$.

(e) What is the variance of S_n?

(f) Show that there is at most a 1% chance that more than 10 people get their own hat back.

Supplementary Problems

Problem 4.
Let K_n be the complete graph with n vertices. Each of the edges of the graph will be randomly assigned one of the colors red, green, or blue. The assignments of colors to edges are mutually independent, and the probability of an edge being assigned red is r, blue is b, and green is g (so $r + b + g = 1$).

A set of three vertices in the graph is called a *triangle*. A triangle is *monochromatic* if the three edges connecting the vertices are all the same color.

(a) Let m be the probability that any given triangle, T, is monochromatic. Write a simple formula for m in terms of r, b, and g.

(b) Let I_T be the indicator variable for whether T is monochromatic. Write simple formulas in terms of m, r, b, and g for $\text{Ex}[I_T]$ and $\text{Var}[I_T]$.

Let T and U be distinct triangles.

(c) What is the probability that T and U are both monochromatic if they do not share an edge?... if they do share an edge?

$$\textbf{Now assume } r = b = g = \frac{1}{3}\textbf{.}$$

(d) Show that I_T and I_U are independent random variables.

(e) Let M be the number of monochromatic triangles. Write simple formulas in terms of n and m for $\text{Ex}[M]$ and $\text{Var}[M]$.

(f) Let $\mu ::= \text{Ex}[M]$. Use Chebyshev's Bound to prove that

$$\Pr[|M - \mu| > \sqrt{\mu \log \mu}] \leq \frac{1}{\log \mu}.$$

(g) Conclude that

$$\lim_{n \to \infty} \Pr[|M - \mu| > \sqrt{\mu \log \mu}] = 0$$

Problem 5.
Let R be a positive integer valued random variable.

(a) If $\text{Ex}[R] = 2$, how large can $\text{Var}[R]$ be?

(b) How large can $\text{Ex}[1/R]$ be?

(c) If $R \leq 2$, that is, the only values of R are 1 and 2, how large can $\text{Var}[R]$ be?

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 14, Mon.

Problem 1.
A recent Gallup poll found that 35% of the adult population of the United States believes that the theory of evolution is "well-supported by the evidence." Gallup polled 1928 Americans selected uniformly and independently at random. Of these, 675 asserted belief in evolution, leading to Gallup's estimate that the fraction of Americans who believe in evolution is $675/1928 \approx 0.350$. Gallup claims a margin of error of 3 percentage points, that is, he claims to be confident that his estimate is within 0.03 of the actual percentage.

(a) What is the largest variance an indicator variable can have?

(b) Use the Pairwise Independent Sampling Theorem to determine a confidence level with which Gallup can make his claim.

(c) Gallup actually claims greater than 99% confidence in his estimate. How might he have arrived at this conclusion? (Just explain what quantity he could calculate; you do not need to carry out a calculation.)

(d) Accepting the accuracy of all of Gallup's polling data and calculations, can you conclude that there is a high probability that the percentage of adult Americans who believe in evolution is 35 ± 3 percent?

Problem 2.
Let $B_1, B_2, \ldots, B_n$ be mutually independent random variables with a uniform distribution on the integer interval $[1, d]$. Let D equal to the number of events $[B_i = B_j]$ that happen where $i \neq j$. It was observed in Section 16.4 (and proved in Problem 18.2) that $\Pr[B_i = B_j] = 1/d$ for $i \neq j$ and that the events $[B_i = B_j]$ are pairwise independent in the course textbook.

Let $E_{i,j}$ be the indicator variable for the event $[B_i = B_j]$.

(a) What are $\mathrm{Ex}[E_{i,j}]$ and $\mathrm{Var}[E_{i,j}]$ for $i \neq j$?

(b) What are $\mathrm{Ex}[D]$ and $\mathrm{Var}[D]$?

(c) In a 6.01 class of 500 students, the youngest student was born 15 years ago and the oldest 35 years ago. Show that more than half the time, there will be will be between 12 and 23 pairs of students who have the same birth date. (For simplicity, assume that the distribution of birthdays is uniform over the 7305 days in the two decade interval from 35 years ago to 15 years ago.)

Hint: Let D be the number of pairs of students in the class who have the same birth date. Note that $|D - \mathrm{Ex}[D]| < 6$ IFF $D \in [12, 23]$.

Problem 3.
Let $G_1, G_2, G_3, \ldots$, be an infinite sequence of pairwise independent random variables with the same expectation, μ, and the same finite variance. Let

$$f(n, \epsilon) ::= \Pr\left[\left|\frac{\sum_{i=1}^{n} G_i}{n} - \mu\right| \leq \epsilon\right].$$

The Weak Law of Large Numbers can be expressed as a logical formula of the form:

$$\forall \epsilon > 0\, Q_1\, Q_2 \ldots\, [f(n, \epsilon) \geq 1 - \delta]$$

where $Q_1 Q_2 \ldots$ is a sequence of quantifiers from among:

$$\begin{array}{cccccc} \forall n & \exists n & \forall n_0 & \exists n_0 & \forall n \geq n_0 & \exists n \geq n_0 \\ & \forall \delta > 0 & \exists \delta > 0 & \forall \delta \geq 0 & \exists \delta \geq 0 & \end{array}$$

Here the n and n_0 range over nonnegative integers, and δ and ϵ range over real numbers.

Write out the proper sequence $Q_1 Q_2 \ldots$

Problem 4.

An *International Journal of Epidemiology* has a policy of publishing papers about drug trial results only if the conclusion about the drug's effectiveness (or lack thereof) holds at the 95% confidence level. The editors and reviewers carefully check that any trial whose results they publish was *properly performed and accurately reported*. They are also careful to check that trials whose results they publish have been conducted independently of each other.

The editors of the Journal reason that under this policy, their readership can be confident that at most 5% of the published studies will be mistaken. Later, the editors are embarrassed—and astonished—to learn that *every one* of the 20 drug trial results they published during the year was wrong. The editors thought that because the trials were conducted independently, the probability of publishing 20 wrong results was negligible, namely, $(1/20)^{20} < 10^{-25}$.

Write a brief explanation to these befuddled editors explaining what's wrong with their reasoning and how it could be that all 20 published studies were wrong.

Hint: xkcd comic: "significant" xkcd.com/882/

Supplementary Problems

Problem 5.

A defendent in traffic court is trying to beat a speeding ticket on the grounds that—since virtually everybody speeds on the turnpike—the police have unconstitutional discretion in giving tickets to anyone they choose. (By the way, we don't recommend this defense :-).)

To support his argument, the defendent arranged to get a random sample of trips by 3,125 cars on the turnpike and found that 94% of them broke the speed limit at some point during their trip. He says that as a consequence of sampling theory (in particular, the Pairwise Independent Sampling Theorem), the court can be 95% confident that the actual percentage of all cars that were speeding is $94 \pm 4\%$.

The judge observes that the actual number of car trips on the turnpike was never considered in making this estimate. He is skeptical that, whether there were a thousand, a million, or 100,000,000 car trips on the turnpike, sampling only 3,125 is sufficient to be so confident.

Suppose you were were the defendent. How would you explain to the judge why the number of randomly selected cars that have to be checked for speeding *does not depend on the number of recorded trips*? Remember that judges are not trained to understand formulas, so you have to provide an intuitive, nonquantitative explanation.

Problem 6.

The proof of the Pairwise Independent Sampling Theorem 19.4.1 was given for a sequence $R_1, R_2, \ldots$ of pairwise independent random variables with the same mean and variance in the course textbook.

The theorem generalizes straighforwardly to sequences of pairwise independent random variables, possibly with *different* distributions, as long as all their variances are bounded by some constant.

Theorem (Generalized Pairwise Independent Sampling)**.** *Let $X_1, X_2, \ldots$ be a sequence of pairwise independent random variables such that* $\text{Var}[X_i] \leq b$ *for some* $b \geq 0$ *and all* $i \geq 1$*. Let*

$$A_n ::= \frac{X_1 + X_2 + \cdots + X_n}{n},$$
$$\mu_n ::= \text{Ex}[A_n].$$

Then for every $\epsilon > 0$*,*

$$\Pr[|A_n - \mu_n| \geq \epsilon] \leq \frac{b}{\epsilon^2} \cdot \frac{1}{n}. \tag{1}$$

(a) Prove the Generalized Pairwise Independent Sampling Theorem.

(b) Conclude that the following holds:

Corollary (Generalized Weak Law of Large Numbers)**.** *For every* $\epsilon > 0$*,*

$$\lim_{n \to \infty} \Pr[|A_n - \mu_n| \leq \epsilon] = 1.$$

6.042J / 18.062J Mathematics for Computer Science
Spring 2015

In-Class Problems Week 14, Wed.

Problem 1. **(a)** Find a stationary distribution for the random walk graph in Figure 1.

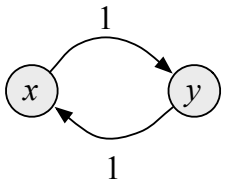


Figure 1

(b) Explain why a long random walk starting at node x in Figure 1 will not converge to a stationary distribution. Characterize which starting distributions will converge to the stationary one.

(c) Find a stationary distribution for the random walk graph in Figure 2.

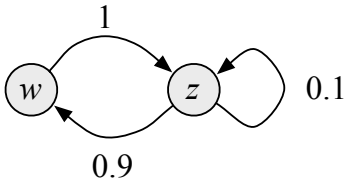


Figure 2

(d) If you start at node w Figure 2 and take a (long) random walk, does the distribution over nodes ever get close to the stationary distribution? You needn't prove anything here, just write out a few steps and see what's happening.

(e) Explain why the random walk graph in Figure 3 has an uncountable number of stationary distributions.

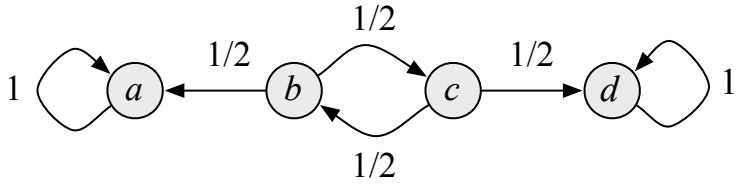


Figure 3

(f) If you start at node b in Figure 3 and take a long random walk, the probability you are at node d will be close to what fraction? Explain.

(g) Give an example of a random walk graph that is not strongly connected but has a unique stationary distribution. *Hint:* There is a trivial example.

Problem 2.
Prove that for finite random walk graphs, the uniform distribution is stationary if and only the probabilities of the edges coming into each vertex always sum to 1, namely

$$\sum_{u \in \text{into}(v)} p(u, v) = 1, \tag{1}$$

where $\text{into}(w) ::= \{v \mid \langle v \to w \rangle \text{ is an edge}\}$.

Problem 3.
A Google-graph is a random-walk graph such that every edge leaving any given vertex has the same probability. That is, the probability of each edge $\langle v \to w \rangle$ is $1/\text{outdeg}(v)$.

A digraph is *symmetric* if, whenever $\langle v \to w \rangle$ is an edge, so is $\langle w \to v \rangle$. Given any finite, symmetric Google-graph, let

$$d(v) ::= \frac{\text{outdeg}(v)}{e},$$

where e is the total number of edges in the graph.

(a) If d was used for webpage ranking, how could you hack this to give your page a high rank? ...and explain informally why this wouldn't work for "real" page rank using digraphs?

(b) Show that d is a stationary distribution.

6.042J / 18.062J Mathematics for Computer Science
Spring 2015