Disclaimer

The TAs do not know what is on the final. The following is our guide for what we believe will be helpful in preparation. We do not have a solution key for this review. We recommend using the practice exam if you would like to self-assess against reference solutions. Good luck!

1 Proof Techniques

1.1 Direct Proof

We directly use our hypotheses to reason that our conclusion is correct.

Practice Problem(s)

1. Prove the claim that the product of two odd numbers is odd.
2. Give a direct proof that if m and n are both perfect squares, then nm is also a perfect square. (An integer a is a perfect square if there is an integer b such that $a = b^2$.)

1.2 Proof by Cases

Proof by exhaustion, also known as proof by cases, is a method of mathematical proof in which the statement to be proved is split into a finite number of cases and each case is solved to show that, for every possible "angle" in the domain of a claim, we can exhaustively show that the claim can be proved.

Practice Problem(s)

1. Prove the claim that there exists irrational $x, y \in R$ such that x^y is rational.
2. Let's agree that given any two people, they have either met or not. If every pair of people in a group has met, we'll call the group a club. If every pair of people in a group has not met, we'll call it a group of strangers.

   Prove that every collection of 6 people includes a club of 3 people or a group of 3 strangers.[a]

[a] *Mathematics for Computer Science* Eric Lehman. 1.7.

1.3 Counterexample

Counterexamples help us prove that a certain claim is not true. A counterexample is a tangible example, that fits appropriately within the domain of a problem, that disproves the claim being made. Note that not every negative statement can be shown by counterexample (e.g., statements of the form “there does not exist. . . ”).

However, you **cannot** prove a claim by showing one example of it. Counterexamples are used to *disprove*. (Alternatively, used to prove an inequality, as of sets.)

For example, the claim “all CS22 students like dinosaurs" can be disproved by finding a student who does not like dinosaurs. Finding this counterexample, however, will not prove that no 22 students like dinosaurs.

Practice Problem(s)

1. Prove or disprove the claim that for all sets A and B, $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$.
2. Prove or disprove via counterexample: $\forall x \in \mathbb{Z}, -1 \leq x \leq 1 \rightarrow x^2 = x$.
3. Consider set A as being the set of positive even integers. $(A_1, A_2) \in R$ if $A_1 = 3 \cdot A_2$. Example, $(18, 6) \in R$. Prove via counterexample that $(A_1, A_2) \in R \wedge (A_2, A_3) \in R \not\rightarrow (A_1, A_3) \in R$.

1.4 Contradiction

To prove the *negation* of a statement $\neg p$, we show that it is impossible for p to hold. This is known as *proof by contradiction*. It proceeds as follows:

1. Assume p is true.
2. Given p is true, use a direct proof to obtain a contradiction.
3. Since p being true leads us to a contradiction, p must be false, i.e., $\neg p$ must be true.

Occasionally, we can also use a similar technique to prove a positive (i.e., not negated) statement. *Before using contradiction, see if a direct approach would suffice.*

Here is how we would prove a (positive) proposition p by contradiction:

1. Assume p is not true.
2. Given p is false, use a direct proof to obtain a contradiction.
3. Since p being false leads us to a contradiction, p must be true.

Practice Problem(s)

1. Prove that there is no least positive real number.
2. Prove “If $3n + 2$ is odd, then n is odd." via contradiction.
3. Show that if n is an integer and $n^3 + 5$ is odd, then n is even using

> a) a proof by contraposition.
>
> b) a proof by contradiction.
>
> 4. Consider a set $A = \{a_1, ..., a_n\}$ with cardinality n. Consider $f : \mathcal{P}(A) \to \{0, 1\}^n$ where $f(X) = s_1 s_2 ... s_n$ and $s_i = 1$ if $a_i \in X$ and $s_i = 0$ if $a_i \notin X$.
>
>   Prove the claim that if $f(X_1) = f(X_2)$ then $X_1 = X_2$.

Samples of different proof types can be found in the resources section of the 22 website.

2 Logic

2.1 Preliminary Definitions

1. A **propositional formula** is a condensed representation of a truth table using logical operators and variables. We call a propositional formula a *proposition* for short.

2. The term **logical expression** is often used synonymously with the word proposition.

3. Two propositions are **logically equivalent** when they represent the same truth table. We can prove propositions are logically equivalent by either comparing their truth tables or using logical rewrite rules. A full list of the rules you can use is on our course website.

4. A **valid proposition** is one that evaluates to true on any choice of inputs; it is true no matter what. It is also sometimes called a tautology. The classic example of a valid proposition is $b \lor \neg b$ (thanks, Shakespeare).

5. A proposition is **satisfiable** if it evaluates to true on *some* choice of inputs; that is, that there is some assignment of the input variables to true and false that makes the proposition true.

6. A proposition is **unsatisfiable** if it is false on any choice of inputs; it is false no matter what. It is also sometimes called a contradiction. The classic example of an unsatisfiable proposition is $p \land \neg p$.

Let's now review the interpretation of each of the following logical operators:

P	Q	P	$P \land Q$	$P \lor Q$	$P \oplus Q$	$P \to Q$	$P \leftrightarrow Q$
T	T	F	T	T	F	T	T
T	F	F	F	T	T	F	F
F	T	T	F	T	T	T	F
F	F	T	F	F	F	T	T

2.2 Implication

In the formula $P \to Q$, we call P the **hypothesis** and Q the **conclusion**. $P \to Q$ is logically equivalent to $\neg P \lor Q$. In words, this means that for $P \to Q$ to be true, Q must be true or P must be false.

This choice can seem a little strange at first. Why is $P \to Q$ true when P is false? Consider the following statement: "If it is raining, I will bring my umbrella." Here are the events that could possibly occur.

- It rains, and I bring my umbrella. That seems fine. The statement is consistent with the situation.
- It rains, and I don't bring my umbrella. The statement does not fit with the situation.
- It doesn't rain, and I bring my umbrella. This situation doesn't seem to directly conflict with the statement. After all, what if I brought my umbrella to block the sun instead? As a result, we say the statement is still consistent with the situation.
- It doesn't rain, and I don't bring my umbrella. The statement seems consistent with this situation, too.

The only scenario where the statement doesn't fit is the second, which is why $P \to Q$ is only false when P is true and Q is false.

- $\neg Q \to \neg P$ is called the **contrapositive** of $P \to Q$; the two are logically equivalent. As a result, we have a useful proof technique: to prove the statement "if p, then q" we can instead prove "if not q, then not p."
- $Q \to P$ is called the **converse** of $P \to Q$. It is **not** logically equivalent to $P \to Q$. If both a statement and its converse are true, then the biconditional $P \leftrightarrow Q$ is true.

2.3 Normal Forms

We say a proposition is in **DNF (disjunctive normal form)** when it is the disjunction (clauses ORed together ($\lor$)) of conjunctions (literals ANDed together ($\land$)).

We say a proposition is in **CNF (conjunctive normal form)** when it is the conjunction (clauses ANDed together ($\land$)) of disjunctions (literals ORed together ($\lor$)).

Here's a truth table, and propositions in DNF and CNF that represent it:

P	Q	R	?
T	T	T	F
T	T	F	T
T	F	T	F
T	F	F	T
F	T	T	F
F	T	F	F
F	F	T	T
F	F	F	T

DNF: $(P \land Q \land \neg R) \lor (P \land \neg Q \land \neg R) \lor (\neg P \land \neg Q \land R) \lor (\neg P \land \neg Q \land \neg R)$

CNF: $(\neg P \lor \neg Q \lor \neg R) \land (\neg P \lor Q \lor \neg R) \land (P \lor \neg Q \lor \neg R) \land (P \lor \neg Q \lor R)$

If we have an arbitrary truth table, here are two ways we can think about describing it:

- Listing the true rows.
- Listing the false rows.

Since every row must be either true or false, both of these ways will uniquely describe our truth table.

These two ways correspond to DNF and CNF, respectively. To write a proposition in DNF, we can think about it like this: we find all rows where our proposition should evaluate to true, and we say that we must be in one of those rows. On the other hand, to write a proposition in CNF, we find all rows where our proposition should evaluate to false, and say we are not in any of those rows.

For DNF, we $\land$ the true variables and negations of the false variables (to be in the row, the inputs must exactly correspond to the row). For CNF, we $\lor$ the false variables and the negations of the true variables (to not be in the row, we just need at least one variable to be different).

In this way, we can represent any truth table in DNF or CNF. We can also rewrite any logical expression to be in DNF or CNF.

Practice Problem(s)

1. Suppose we define a new operation $\star$ on logical propositions such that

$$x \star y \equiv \neg(x \land y)$$

Create a truth table for each of the following expressions, and state which logical operator the expression is equivalent to.

- $x \star x$
- $(x \star y) \star (x \star y)$

- $(x \star x) \star (y \star y)$

- $(x \star (x \star y)) \star (y \star (y \star x))$

2. Write two propositions corresponding to the following truth table: one in DNF and one in CNF.

P	Q	R	?
T	T	T	T
T	T	F	T
T	F	T	F
T	F	F	F
F	T	T	F
F	T	F	T
F	F	T	T
F	F	F	T

2.4 First-Order Logic

In propositional logic, we only consider "atomic" propositions, represented by propositional variables like p and q. First-order logic is more expressive: it allows us to write propositions *about* particular data (like numbers).

In particular, first-order logic lets us write expressions that make assertions about particular entities. Such expressions are called **predicates**; you can think of these a bit like functions from the data in question to the values "true" and "false." For instance, we might define a predicate $\text{Odd}(n)$ that holds of a natural number $n : \mathbb{N}$ if and only if n is odd. Predicates are syntactically represented by *predicate variables* (like Odd), with the value of which they are being asserted written in parentheses after the predicate variable.

A given predicate can only make assertions about a certain *kind* of object: for instance, it wouldn't really make sense to apply the predicate Odd above to an irrational number like $\sqrt{2}$. We therefore define for each predicate a **domain**, the collection of all the possible values of which the predicate can be asserted. (The domain of Odd would be $\mathbb{N}$.)

Given some predicate, we may wish to make claims about whether it holds of *any*, or of *all*, elements in its domain. **Quantifiers** allow us to express such claims in first-order logic. There are two quantifiers of note:

1. Universal quantifier: denoted by the $\forall$ symbol, it represents that a predicate holds for *every* element in its domain.

2. Existential quantifier: denoted by the $\exists$ symbol, it represents that a predicate holds for *some* element in its domain.

For instance, the formula $\forall n : \mathbb{N}, \text{Odd}(n)$ asserts that every natural number is odd (this is false!). On the other hand, $\exists n : \mathbb{N}, \text{Odd}(n)$ asserts that at least one odd natural number exists (this is true!).

Note that a universal quantification over an empty domain is always true, while an existential quantification over an empty domain is always false.

We can also chain quantifiers in sequence to represent a more complex proposition.

Example

Problem: Render *Goldbach's conjecture*, that every integer greater than 2 is the sum of two primes, in first-order logic.[a]

We first can reformulate this in English in a way that better matches our first-order syntax: "For every even integer n greater than 2, there exist primes p and q such that $n = p + q$". *Note: the TAs find this to be especially helpful!*

We can then define some predicates. Let G be the predicate on natural numbers defined by $G(n) := n \geq 2$; that is, $G(n)$ is true just in case $n \geq 2$. Let P be the predicate on natural numbers such that $P(n)$ holds just in case n is prime.

We can thus describe the conjecture in first-order logic as follows:

$$\forall n : \mathbb{N}, G(n) \to \exists p, q : \mathbb{N}, P(p) \wedge P(q) \wedge n = p + q$$

Note that the order of quantifiers is essential. If we switched the order of the quantifiers, we would essentially assert that there are two prime numbers whose sum is equal to every number greater than 2. (This is clearly false!)

Note also the different ways we "restrict" universal and existential quantifiers (so that we are only considering n satisfying G and so that the witnesses p and q must have property P). If we switched the $\to$ and $\wedge$ symbols in the above, our formula would be incorrect! (Think about why.)

[a] *Mathematics for Computer Science* Eric Lehman. 3.6.

Practice Problem(s)

1. For following questions, assume these definitions:

   - Sets:
     - T: The set of CS22 TAs.
   - Predicates:
     - $D(x)$ "x double majors at Brown"
     - $M(x, y)$ "x and y share a major in common."
     - $F(x, y)$ "x and y are friends."

- Functions
  - $n(x)$: The number of majors x studies.
- Constants
  - r: Rob, cs22 professor (also known as the Last Logician)!

Using the above definitions, translate these sentences into First-Order logic.

a) Every TA on the 22 staff double majors.

b) A TA on staff is friends with every other TA on staff.

c) Every TA on staff is friends with Rob. (<3).

d) TAs that study the same major are friends.

e) Every TA has a friend who studies more majors than them.

f) There is a TA who doesn't have the same major as anyone else on staff.

g) There is a TA that studies more majors than any other TA on staff.

2. Translate the following sentences into first-order logic. You may only use $\mathbb{N}$ as a domain of quantification. You may use the relations $=$, $<$, $\leq$, $>$, and $\geq$; functions $+$, $-$, and $\times$; and one-place predicate Prime. You may not use any quantifiers or connectives other than those we have discussed in this guide.

   a) The difference of any two natural numbers is no greater than their sum.

   b) No prime number is square.

   c) (Challenge) There is a *unique*[a] natural number that is less than every other natural number.

[a] I.e., it is the only natural number with this property.

3 Formal Proofs

Remember that the Lean documentation website has additional materials covering these topics!

3.1 Proof Structure

When we write a proof, we are implicitly (or, in the case of Lean, explicitly) manipulating a **proof state**. Our proof state consists of:

1. One or more **goals**, the propositions we are trying to prove; and

2. Attached to each goal, a **context** consisting of **hypotheses**. These are things that we know or have assumed *for that particular goal.*

Our proof begins with no hypotheses and one goal (the theorem statement we're trying to prove). A proof proceeds by applying **proof rules** to our proof state. A proof rule may modify our goal, close (i.e., fully prove) our goal, add new goals, modify a hypothesis, or add new hypotheses.

When a new goal is created (e.g., as a result of a proof rule like disjunction elimination), it gets it own, distinct context. Changes to the context of one goal do *not* appear in the context of another! Proof rules apply only to one goal at a time! We "focus" one goal to work on, moving onto the next after closing the previous one.

3.2 Propositional Proof Rules

We divide our proof rules into two main categories: introduction rules and elimination rules.

1. **Introduction rules** tell you what is required in order to *prove* a *goal* of a given form.

2. **Elimination rules** tell you what is required in order to *extract information from* a *hypothesis* of a certain form.

In other words, if your goal contains some connective, you can use an introduction rule to "simplify" your goal into the propositions required in order to obtain the larger formula with the connective. Note that the vocabulary is a little confusing here: once you apply an introduction rule, the relevant connective "disappears" from your goal. This is because applying an introduction rule works *backward*: we're saying that in order to conclude a goal with the relevant connective, it suffices to show some simpler goal(s) that don't include that connective.

On the other hand, elimination rules "extract data from" *hypotheses* that you already know. They tell you what "simpler" propositions you're allowed to conclude based on the knowledge that you know a proposition containing a certain connective. Here are the proof rules for each propositional connective:

Connective	Intro Rule(s)	Elim Rule(s)
Conjunction ($\wedge$)	To show $p \wedge q$, show p, then show q (Lean: `split_goal`).	If you know $p \wedge q$, conclude p and q (Lean: `eliminate`).
Disjunction ($\vee$)	- To show $p \vee q$, show p (Lean: `left`). - To show $p \vee q$, show q (Lean: `right`).	(Proof by cases) If you know $p \vee q$, and you can show r assuming p, and you can show r assuming q, then conclude r (Lean: `eliminate`).
Implication ($\to$)	To show $p \to q$, assume p and show q (Lean: `assume`).	(*Modus ponens*) If you know $p \to q$ and you know p, conclude q (Lean: `have`/`apply`).
Bi-implication ($\leftrightarrow$)	To show $p \leftrightarrow q$, show $p \to q$, then show $q \to p$ (Lean: `split_goal`).	If you know $p \leftrightarrow q$, conclude $p \to q$ and $q \to p$ (Lean: `eliminate`).
Negation ($\neg$)	(Proof by contradiction) To show $\neg p$, assume p and derive a contradiction (Lean: `assume`).	(Explosion) If you know p and you know $\neg p$, conclude anything (Lean: `contradiction`).

We also have the "structural" rule of **atom introduction**: if you know p, then you can conclude p (Lean: `assumption`).

3.3 First-Order Proof Rules

In first-order logic, we enrich our notion of a context to allow for *data.* In addition to hypotheses, our context tracks variables that correspond to things like numbers or sets.

Just as for connectives, we have introduction and elimination rules for quantifiers:

Quantifier	Intro Rule	Elim Rule
Universal ($\forall$)	To show $\forall x : T,\ P(x)$, fix an arbitrary $t : T$ in your context, then show $P(t)$. (Lean : `fix`)	If you know $\forall x : T,\ P(x)$ and you have a piece of data $t : T$, then conclude $P(t)$. (Lean: `have`)
Existential ($\exists$)	To show $\exists x : T,\ P(x)$, choose a *specific* witness value $k : T$, then show that $P(k)$ holds. (Lean: `existsi`)	If you know $\exists x : T,\ P(x)$, add a piece of data $t : T$ to your context and conclude $P(t)$. (Lean: `eliminate`)

We have also seen two other kinds of rules for first-order logic:

1. **Rewriting** allows us to substitute equals for equals in a goal or hypothesis. For instance, if we have a hypothesis that says that $x = y$ and our goal is $P(x)$, rewriting lets us turn our goal into $P(y)$. (Lean: `rewrite`)

2. **Arithmetical reasoning** lets us close goals that are solvable using straightforward arithmetic or high-school algebra, like $15 + 151 > 22$ or (assuming $x, y : \mathbb{N}$) $x + y > x$. (Lean: `numbers`, `linarith`, `polyrith`)

Practice Problem(s)

1. Prove the proposition $(p \lor q) \to (\neg p \to q)$ using proof rules. Carefully describe the proof rule you are applying at each step and how it updates your proof state. If multiple goals arise, clearly indicate which goal you are focusing on and when you move from one goal to the next.

2. Suppose that we are in the middle of a proof, and our proof state looks as follows (assume that this is the only goal):

$$\begin{array}{ll} p\ q\ r & : \mathsf{Prop} \\ hp & : p \\ hq & : q \\ hpr & : p \to r \\ hr & : r \\ \vdash q \land r & \end{array}$$

What theorem might we be trying to prove? (In other words, what could our original goal at the start of the proof have been?) Assume that our initial goal contained *at most three → symbols*. Based on this, what proof rules must we have applied to get to our current state? What proof rules are required to finish this proof, and how will each change our proof state?

3. Suppose we've defined certain natural numbers to be *prehistoric*. (We won't actually specify what we've defined this term to mean.) Translate the sentence "for every prehistoric natural number, there's some some non-prehistoric natural number that is greater than it." The only domain of quantification you should use is $\mathbb{N}$; the only predicate symbols you should use are $\mathsf{Prehistoric}$ and $>$.

   Explain what proof rules you would use to prove this proposition. What would your proof state look like just after applying the introduction rule for the second quantifier in your formula?

4. Suppose we decided to change our formal proof system by getting rid of the usual elimination rule for disjunction and replacing it with the following: "If you know $p \vee q$, conclude p." Give an intuitive explanation of why this would be a bad idea. Then show (by exhibiting an appropriate proof) that we can use this formal proof system to prove False.

5. Why do we require that every goal have a distinct context? Give an example of a connective that highlights the importance of this property.

   Suppose, contrarily, that our proof system were such that all goals shared the same context (i.e., if a hypothesis is added to one goal's context, it gets added to all other goals' contexts, too). Give a proof that $0 = 1$ using this system.

4 Sets and Notation

A set is a collection of objects without order or repetition.

4.1 Membership vs. Subsets

If an object s is a member of a set S, we say $s \in S$. If a set T is a subset of a set S, we write $T \subseteq S$. This means that every member of T is also a member of S.

Practice Problem(s)

1. A is any set. Which of the following is **always true**?

   i. $A \subseteq A$

   ii. $\{\} \subseteq A$

   iii. $\{\} \in A$

2. A is any set and $\mathcal{P}(A)$ is the set of all subsets of A. Which of the following is **always true**?

   i. $A \in \mathcal{P}(A)$

   ii. $A \subseteq \mathcal{P}(A)$

   iii. $\emptyset \in \mathcal{P}(A)$

   iv. $\emptyset \subseteq \mathcal{P}(A)$

   v. $\{A, \emptyset\} \subseteq \mathcal{P}(A)$

3. S is the set of students in CS22. B is the set of students at Brown. Duncan is a student in CS22. Which of the following is **always true**?

   i. $S \subseteq B$

   ii. Duncan $\subseteq S$

   iii. Duncan $\in S$

   iv. $\{\text{Duncan}\} \subseteq B$

4.2 Set Operations

- The union $A \cup B$ of two sets A and B is the set of all elements that are in A or B.

- The intersection $A \cap B$ of two sets A and B is the set of all elements that are in A *and* B.

- The set difference $B \setminus A$ of two sets A and B is the set of all elements that are in B, but that are not in A.

- The complement $\overline{A}$ of a set A is the set of all elements that are *not* in A (where "all elements" refers to all elements in some universal set U.)

- The cardinality $|A|$ of a set A is the number of elements of A. Remember that sets have no duplicates!

Practice Problem(s)

1. For this problem, let

$$
\begin{aligned}
A &= \{-3, -1, 0, 6\} \\
B &= \{x : \mathbb{Z} \mid x^2 \leq 5\} \\
C &= \{x : \mathbb{Z} \mid \exists\, y \in \mathbb{Z} \text{ s.t. } 3y = x\} \\
D &= \{x : \mathbb{Z} \mid \exists\, y \in \mathbb{Z} \text{ s.t. } y^2 <= x\}
\end{aligned}
$$

Find the following sets (not all sets are finite):

i. $A \cup B$

ii. $A \cap \overline{C}$

iii. $B \cap D$

iv. $A \setminus C$

v. $B \setminus (C \cup D)$

vi. $C \cap D$

vii. $C \cup D$

Find the cardinalities of the followings sets:

i. A

ii. $(B \setminus C) \setminus D$

iii. $(A \cup B) \cap (C \cap D)$

4.3 Power Sets

The *power set* of a set S, denoted $\mathcal{P}(S)$, is the set of all subsets of S. The power set of S has cardinality $2^{|S|}$. We proved this last result by noticing that there are the same number of subsets of a set of size n as there are binary strings of length n (see the sample bijective proof on the website).

Practice Problem(s)

1. Let A, B, C, D be the sets from above, find the following sets:

   i. $\mathcal{P}(A \cap B)$

   ii. $\mathcal{P}(A) \cap \mathcal{P}(D)$

   iii. $\mathcal{P}(B) \setminus \mathcal{P}(C)$

   iv. $\mathcal{P}(\mathcal{P}(\varnothing))$

2. Prove that $|A| < |\mathcal{P}(A)|$ for any arbitrary finite set A.

4.4 Product

The product of two sets A and B, denoted $A \times B$, is the set of all ordered pairs (a, b) for $a \in A$, $b \in B$. The product of a single set, A, is the set of all ordered pairs (a, a) where $a \in A$.

Practice Problem(s)

1. Let A, B, C, D be the sets from above, find the following sets:

   i. $A \times B$

   ii. $B \times \{\emptyset\}$

   iii. $C \times \emptyset$

   iv. $\emptyset \times \emptyset$

2. Prove that $(\mathbb{Z} \times \mathbb{N}) \cap (\mathbb{N} \times \mathbb{Z}) = \mathbb{N} \times \mathbb{N}$.

3. Disprove the following claim: $|A \times A| < |\mathcal{P}(A)|$ for any arbitrary finite set A.

4. Disprove the following claim: for any two finite sets A and B, $|\mathcal{P}(A \times B)| = |\mathcal{P}(A) \times \mathcal{P}(B)|$.

4.5 Proof by Set-Element Method

How do you prove that some set A equals some set B? First show that $A \subseteq B$ and then you show that $B \subseteq A$. If every element in A is also an element in B and every element in B is also an element of A, then A must equal B.

To show that $A \subseteq B$ you consider an arbitrary element in A and show it is also in B. In use, this looks like the following:

1. Let x be an element of set A.

2. Prove that x is also an element of B.

3. Conclude that $A \subseteq B$.

Example

Claim: $A \cap (A \cup B) = A$.

Proof. We show that both $A \cap (A \cup B) \subseteq A$ and $A \subseteq A \cap (A \cup B)$.

We first prove the subclaim that $A \cap (A \cup B) \subseteq A$.

Consider any $x \in A \cap (A \cup B)$. By definition of intersection, this means that $x \in A$ and $x \in A \cup B$. Because every arbitrary x in $x \in A \cap (A \cup B)$ is in A, we can conclude that $A \cap (A \cup B) \subseteq A$.

We then prove the subclaim that $A \subseteq A \cap (A \cup B)$.

Consider any $x \in A$. By definition of union, we can reason that $x \in A \cup B$. Because we know that $x \in A \land x \in (A \cup B)$, by definition of intersection, we conclude $x \in A \cap (A \cup B)$. It follows that because every arbitrary x in A is in $A \cap (A \cup B)$, $A \subseteq A \cap (A \cup B)$.

Therefore, by the set element method we have proved that $A \cap (A \cup B) = A$. □

Practice Problem(s)

1. Prove the claim that $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$ via the set element method.

2. Prove the following properties using the set-element method.

   a) $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$

   b) $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$

3. Prove or disprove each of the following. To prove equality, use the set-element method.

   a) $A \cup (B \cap C) = (A \cup B) \cap C$

   b) $\overline{A} \cup (A \cap B) = \overline{A} \cup B$

   c) $A \cap (B \setminus C) = (A \setminus B) \cap (A \setminus C)$

4.6 Set Algebra

1. Conversion of one side of the equation to the other (or conversion of both sides to an identical expression) using stated laws of set algebra. (See list of set identities on course website!)

2. Conclusion based on the biconditionality of the steps taken.

Note: Do not assume equality before applying set identities! Either rewrite one side to look like the other or rewrite both sides separately to look like the same expression.

Example

Claim: $(A \cap B) \cup (A \setminus B) = A \cap (B \cup (A \setminus B))$.

Proof.

$$\begin{aligned}
&(A \cap B) \cup (A \setminus B) && \\
&= (A \cap B) \cup (A \cap \overline{B}) && \text{(Set Difference Law)} \\
&= A \cap (B \cup \overline{B}) && \text{(Distributive Law)} \\
&= A \cap U && \text{(Complement Law)} \\
&= A && \text{(Identity Law)} \\
&= A \cap (A \cup B) && \text{(Absorption)} \\
&= A \cap (B \cup A) && \text{(Commutativity)} \\
&= A \cap ((B \cup A) \cap U) && \text{(Identity Law)} \\
&= A \cap ((B \cup A) \cap (B \cup \overline{B})) && \text{(Complement Law)} \\
&= A \cap (B \cup (A \cap \overline{B})) && \text{(Distributive Law)} \\
&= A \cap (B \cup (A \setminus B)) && \text{(Set Difference Law)}
\end{aligned}$$

□

Practice Problem(s)

1. $(A \cup B) \cap \overline{(A \cap B)} = (B \setminus A) \cup (A \setminus B)$.

2. $(A \cap \overline{B}) \cup B = A \cup B$.

3. $(A \setminus B) \setminus (B \setminus C) = (A \cup B) \setminus (A \cap B)$.

5 Relations

5.1 Cartesian Products

A *binary* (or *two-place*) *relation* R consists of a set A, called the *domain*; a set B, called the *codomain*; and a subset of the Cartesian product $A \times B$ called the *graph*. If we say that a relation R is *on* a set A, we mean that both its domain and codomain are A.

Always remember to specify the set(s) on which the relation is defined!

We write aRb or $(a, b) \in R$ to mean that a is related to b by R.

5.2 Reflexivity

A relation R on A is *reflexive* if for all $a \in A, (a, a) \in R$. In other words, a relation is reflexive if *every element* in the set A is related to itself in R. This is why it's important to specify a set when talking about a relation: you can't tell if a relation is reflexive if you don't know which elements have to be related to themselves (and every element must be!).

5.3 Symmetry and Transitivity

A relation R is *symmetric* if for all a, b in its domain, the following holds: **if** $(a, b) \in R$, **then** $(b, a) \in R$.

A relation R is *transitive* if for all a, b, c in its domain, the following holds: **if** $(a, b) \in R$ and $(b, c) \in R$, **then** $(a, c) \in R$. Remember that a, b, and c do not need to be different elements.

It's important to note that the definitions of symmetry and transitivity are phrased as if-then statements. A relation is symmetric/transitive *unless* it violates the appropriate if-then condition. To violate the condition, you must simultaneously satisfy the if-clause, and violate the then-clause.

Consider the following example of a relation that is not transitive: the ordered pairs $(1, 2)$ and $(2, 1)$ are in the relation (this satisfies the if-clause of the transitivity definition) but there is no pair $(1, 1)$ in the relation (this violates the then-clause).

As another illustrative example: any empty relation is reflexive, symmetric, and transitive, as there are no ordered pairs in the empty relation to satisfy the if-clause of the definitions.

5.4 Equivalence Relation

An *equivalence relation* is a relation that is reflexive, symmetric, and transitive.

5.5 Equivalence Classes

Let R be an equivalence relation on A. Then the *equivalence class* of $a \in A$ is defined as

$$[a]_R := \{x \mid x \in A, (x, a) \in R\}.$$

Note that a is not unique (unless it is the only element in its equivalence class.) Rather, any element in the same equivalent class can serve equally well as the representative for the class. An equivalence relation splits a set into equivalences classes. In other words, it forms partitions.

A *partition* of a set A is a collection of nonempty subsets $B_1, \ldots, B_k$ of A such that

1. $B_1 \cup \cdots \cup B_k = A$, and
2. $B_i \cap B_j = \emptyset \ \ \forall i, j$ where $i \neq j$.

Practice Problem(s)

1. Consider the set B of all students at Brown. For each of the following relations on B, state if they are reflexive, symmetric, or transitive. If they are an equivalence relation, then list the equivalence classes.

   i. Two students are related if they are the same age (e.g. 21).

   ii. s_1 and s_2 are students and $(s_1, s_2) \in R$ if s_1 is younger than s_2.

   iii. Two students are related if they are studying anthropology.

   iv. Two students are related if they go to Brown.

2. Let $A = \{1, 2, 3\}$. Consider the following relations on $\mathcal{P}(A)$. State if they are reflexive, symmetric, or transitive. If they are an equivalence relation, then list the equivalence classes.

   i. $(S_1, S_2) \in R$ if $|S_1| = |S_2|$.

   ii. $(S_1, S_2) \in R$ if $S_1 \subseteq S_2$.

   iii. $(S_1, S_2) \in R$ if S_1 and S_2 share an element.

6 Functions

6.1 Formal Definition

A *function* $f : A \to B$ is a relation on A and B with the following property: for every $a \in A$ there exists exactly one pair (a, b) in the relation, where $b \in B$.

We call A the domain and B the codomain.

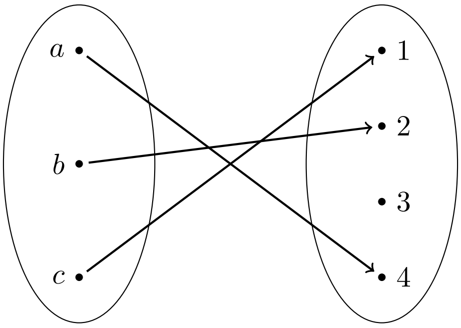
It's important to note that a function is characterized not only by the "rule" that maps inputs to outputs, but also by the domain and codomain.

Additionally, we call the set of all $b \in B$ such that there exists $a \in A$ where $f(a) = b$ the *image* of f. In other words, the image is the set of all elements mapped to by f.

6.2 Injectivity

A function is injective if for all $b \in B$, there exists at most one $a \in A$ such that $f(a) = b$. In other words, no two distinct elements map to the same thing!

If a function $f : A \to B$ is injective, we know that $|A| \leq |B|$. This is because every element in A needs some unmatched element in B, so B needs to have at least as many elements as A!



There are two ways to prove that a function is injective:

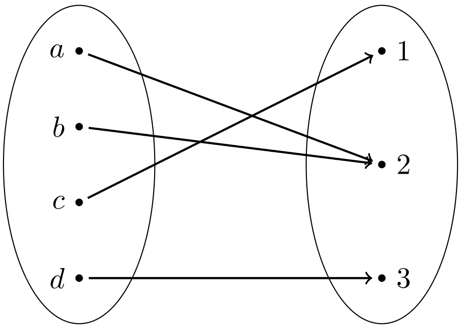
1. Consider two arbitrary elements a and b of the domain, and show that if $f(a) = f(b)$, then we must have $a = b$.

2. Consider two arbitrary distinct elements $a \neq b$ in the domain. Show that they must map to distinct outputs $f(a) \neq f(b)$.

6.3 Surjectivity

A function is surjective if for all $b \in B$, there exists at *least* one $a \in A$ such that $f(a) = b$. In other words, no element in the codomain gets left behind: there is always some element that maps to it. Equivalently, a function is surjective if the image of the function is the entire codomain.

If a function $f : A \to B$ is surjective, we know that $|A| \geq |B|$. This is because every element in B needs some element in A to map to it, so A needs to have at least as many elements as B.



To prove that a function is surjective, consider an arbitrary element in the codomain, and construct the specific element in the domain that maps to it.

Practice Problem(s)

1. Let $S = \{0, 1\}$, $T = \{t \mid t \subseteq S \times S\}$, and R be the set of all possible functions from S to S. [a]

   i. Can an injection from T to R exist? If so, give one such injection and prove that this mapping is indeed injective. If not, prove why such a mapping cannot exist.

   ii. Can a surjection from T to R exist? If so, give one such surjection and prove that this mapping is indeed surjective. If not, prove why such a mapping cannot exist.

[a] *HW3 Problem 1* CSCI0220 2022 Spring

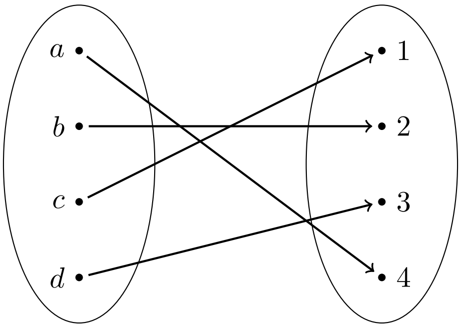
6.4 Bijectivity

A bijection is a function that is both injective and surjective. Thus, to prove that a function is a bijection, prove that it is injective and surjective.

If we combine our results from injectivity and surjectivity, we know that the cardinality of the domain must be less than or equal to that of the codomain (by injectivity), and that the cardinality of the domain must be greater than or equal to that of the domain (by surjectivity.) Thus, the cardinalities of the two sets must be equal. This is a powerful result:

There exists a bijection between two sets if and only if they have equal cardinality.

Thus, to prove that the sizes of two sets are equal, it suffices to prove that there exists a biejction between them.



Practice Problem(s)

1. For each of the following, state if it is a function- if it is a function, conclude if it is injective and/or surjective.

   a) $f : \mathbb{Z} \to \mathbb{Z}$ where $f(x) = x^2$

   b) $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ where $f(x) = x^2$. $\mathbb{Z}^+$ denotes the positive integers.

c) $f : \mathbb{Z} \to \mathbb{Z}$ where $f(x) = \sqrt{x}$.

d) $f : A \to B$, where $f(\text{student})$ = the dorm that the student lives in, A represents the set of first year students at Brown, and B represents the set of first year dorms at brown.

e) $f : A \to B$, where $f(\text{student})$ = the banner ID of student, A represents the set of students at Brown, and B represents the set containing the Banner IDS of all current students at brown.

f) f : People in the World $\to \{0, 1\}$ where $f(\text{person}) = 1$ if they are Prof. Lewis and 0 otherwise.

g) $f : A \to \mathbb{Z}$, where A represents the set of libraries at Brown and $f(\text{Library}) =$ number of books in the library.

h) $f : S \to \mathcal{P}(S)$ where $f(S) = \{S\}$.

i) $f : \mathcal{P}(\{1, 2, 4\}) \to \{0, 1, 2, 3\}$ where $f(X) = |X|$.

2. Let X be a set with n elements. Let B be the set of bit strings of length n. B can be expressed as $\{0, 1\}^n$. Prove that there is a bijection between $\mathcal{P}(X)$ and B, then conclude that $|\mathcal{P}(X)| = 2^n$.

3. (Challenge) Let A be a set with n elements. Let T be the set of all ordered pairs (X, Y) where X and Y are subsets of A. Let S be the set of 0/1/2/3 strings of length n. That is, elements of S are strings of length n where each character is 0, 1, 2, or 3. Prove that T and S must be the same size by defining a bijection between T and S. [a]

4. Let A, B, and C be sets, and let $f : B \to C$ and $g : A \to B$ be functions. Let $h : A \to C$ be the composition, $f \circ g$, that is, $h(x) = f(g(x))$ for $x \in A$. We want to prove or disprove the following claims:

   (a) If h is surjective, then f must be surjective.

   (b) If h is surjective, then g must be surjective.

   (c) If h is injective, then f must be injective.

   (d) If h is injective and f is total, then g must be injective.

[a] *HW3 Problem 2* CSCI0220 2023 Spring

7 Induction

7.1 Template and Weak Induction

Induction is a proof method for which we can assume some n case, and prove that every $n+1$ case holds. If we can prove that the $n+1$ case holds, we can confirm that our original claim holds for all values of n in the desired domain..

Idea: If you are stuck on an induction problem on the exam, start by writing out the inductive hypothesis and the structure of the proof. You will receive partial credit for this and it will also help you think of how to proceed.

Idea: Often the inductive step is a direct proof using the inductive hypothesis. This is not always the case; sometimes you might have to use ***proof by cases*** *or even* ***contradiction****.*

We will first provide a review of the template for an inductive proof and provide an example.

Example

For example, say we are trying to prove that $\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$ is true for all $n \in \mathbb{N}$.

1. Define the predicate $P(n)$.

   Let $P(n)$ be the predicate that $\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$.

2. Show that the base case is true.

   We will first show $P(0)$ is true. $\sum_{i=0}^{0} i = 0$ and $\frac{0(0+1)}{2} = 0$ so they are equal as needed.

3. Assume the inductive hypothesis is true. If you are using stardard induction then you will assume $P(k)$ is true for some integer k. If you are using strong induction then you will assume $P(i)$ is true for all $i \leq k$. Either way, you should specify that k is some integer greater than or equal to your greatest base case.

   Assume $P(k)$ is true for some arbitrary integer $k \geq 0$.

4. Show that $P(k+1)$ is true given the inductive hypothesis.

   We will now show that $\sum_{i=0}^{k+1} i = \frac{(k+1)(k+2)}{2}$.

   We know that $\sum_{i=0}^{k+1} i = \left(\sum_{i=0}^{k} i\right) + (k+1)$.

   By our inductive hypothesis $\sum_{i=0}^{k} i = \frac{k(k+1)}{2}$.

   Therefore

$$\begin{aligned}\sum_{i=0}^{k+1} i &= \left(\sum_{i=0}^{k} i\right) + (k+1)\\ &= \frac{k(k+1)}{2} + (k+1)\\ &= \frac{k(k+1) + 2(k+1)}{2}\\ &= \frac{(k+1)(k+2)}{2}\end{aligned}$$

   as needed. □

5. Conclude the proof.

   Therefore, as $P(0)$ is true and $P(k)$ implies $P(k+1)$ for all $k \in \mathbb{Z}$, $k \geq 0$, $P(n)$ is true for all nonnegative integers n.

Practice Problem(s)

1. Show that $6 \mid (n^3 - n)$ for all $n \in \mathbb{N}$.

2. Recall that the Fibonacci numbers are defined by $f_0 = 0, f_1 = f_2 = 1$ and the recursive relation $f_{n+1} = f_n + f_{n-1}$ for all $n \geq 1$. Challenging exercise:

   Show that f_n and f_{n+1} are 'relatively prime' for all $n \geq 1$. That is, they share no factor in common other than the number 1.

3. Use induction to show that $12^n + 2(5^{n-1})$ is divisible by 7 for all $n \geq 1$.

7.2 Strong Induction

The difference in approach between weak and strong induction comes in the induction hypothesis! In weak induction, we only assume that the predicate holds for some arbitrary step k, while in strong induction, we assume that the predicate holds at all steps from the base case to some arbitrary step k. Your inductive step may differ depending on whether you approach a problem using weak or strong induction, but they are equivalent!

Practice Problem(s)

1. Define the sequence S as follows: $S_1 = 1$, $S_2 = 3$, $S_n = S_{n-1} * S_{n-2}$ for integers $n \geq 2$. Prove that S_n is odd for all positive integers n.

2. Prove that every positive integer greater than one can be factored as a product of primes, using strong induction.

3. You begin with a stack of n boxes. Then you make a sequence of moves. In each move, you divide one stack of boxes into two nonempty stacks. The game ends when you have n stacks, each containing a single box. You earn points for each move; in particular, if you divide one stack of height a C b into two stacks with heights a and b, then you score ab points for that move. Your overall score is the sum of the points that you earn for each move. What strategy should you use to maximize your total score?

8 Number Theory

8.1 Definitions

Definition 1: We say that a *divides* b, denoted $a \mid b$, when $b = ka$ for some $k \in \mathbb{Z}$.

Definition 2: We say that a is *congruent* to b mod m, denoted $a \equiv b \pmod{m}$, if $m \mid (b - a)$. Another way to say this is that $a = b + km$ for some $k \in \mathbb{Z}$. Yet another way to say this: a and b have the same remainder upon division by m. Take a moment to convince yourself that these statements are equivalent.

8.2 Properties of Congruence Relations:

For $a, b \in \mathbb{Z}$, if $a \equiv b \pmod{m}$,

1. $a + c \equiv b + c \pmod{m}$ for any $c \in \mathbb{Z}$
2. $ac \equiv bc \pmod{m}$ for any $c \in \mathbb{Z}$
3. $a^n \equiv b^n \pmod{m}$ for $n \in \mathbb{Z}^+$

If we also have $c \equiv d \pmod{m}$,

1. $a + c \equiv b + d \pmod{m}$
2. $ac \equiv bd \pmod{m}$

8.3 GCD

The greatest common denominator of a and b is the largest positive integer which divides both a and b. To find the gcd of two numbers, we can run the Euclidean algorithm.

Theorem 1: For any $a, b \in \mathbb{Z}$ there exists $u, v \in \mathbb{Z}$ such that $au + bv = \gcd(a, b)$. In words, we say that a and b can be written as a linear combination of their gcd.

Theorem 2: An integer is a linear combination of a and b if and only if it is a multiple of their gcd.

Practice Problem(s)

1. Calculate the GCD(1147,899)
2. Find the gcd(556,148) using the Euclidean algorithm, and express it as a linear combination of 556 and 148.
3. A bug is standing on a grid and can take four possible kinds of steps: $(9, 2)$, $(-12, 3)$, $(3, -6)$, $(-9, -12)$. Prove that if the bug starts at $(0, 0)$ then it can never reach $(1, 1)$.

8.4 Multiplicative Inverse

Consider the particular congruence

$$ax \equiv 1 \pmod{m}.$$

If this equation has a solution, then we know we can find some integer x which, when multiplied by a, yields $1 \pmod{m}$. We define this integer to be the *multiplicative inverse* of $a \pmod{m}$, and we denote it a^{-1}. If a multiplicative inverse exists $\pmod{m}$, then when working $\pmod{m}$, we can "divide" by a—that is, we can multiply two sides of a congruence by a^{-1}, cancelling a from both sides.

When does a multiplicative inverse exist? According to the above Theorem 2: a^{-1} exists if and only if $\gcd(a, m)$ divides 1 (which is c in this particular congruence.) Thus, a^{-1} exists $\pmod{m}$ if and only if $\gcd(a, m) = 1$, that is, if and only if a and m are relatively prime. How do we find the multiplicative inverse? We can run the Euclidean algorithm and then backtrack to obtain the multiplicative inverse (gcdcombo).

8.5 Fermat's little Theorem

If p is prime and does not divide $a \in \mathbb{Z}$ then

$$a^{p-1} \equiv 1 \pmod{p}.$$

This means a^{p-2} is a multiplicative inverse for a mod p.

8.6 Euler's Totient Function

The totient function of n is a count of how many positive integers less than or equal to n are relatively prime to it. For any prime p, $\phi(p) = p - 1$. If m and a are relatively prime, then

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

This means $a^{\phi(m)-1}$ is a multiplicative inverse for a mod m. Fermat's little theorem is just a special case of this rule.

Practice Problem(s)

1. Prove that if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$ then $a + c \equiv b + d \pmod{m}$
2. Compute the multiplicative inverse of 8 mod 27 with the Euclidean algorithm and the Euler-Fermat method.
3. For all positive (non-zero) integers a, b, and k, prove that $\gcd(ka, kb) = k \cdot \gcd(a, b)$.
4. Use Fermat's Little Theorem or Euler's Theorem to find the inverse of these numbers. If the inverse does not exist, show why it does not exist.

   i. $14^2 \pmod{5}$

   ii. $1452 \pmod{9}$

   iii. $4^6 \pmod{15}$

9 Counting

9.1 Product Rule and Permutations

The **product rule** states that for finite sets $S_1, ..., S_n$, $|S_1 \times ... \times S_n| = |S_1| * ... * |S_n|$. This can be useful in representing how many ways we could make a series of n independent choices. If we know how many options we have for each choice, we can find the number of ways we could make all of the choices by multiplying all the numbers of options together.

If the choices are instead dependent on each other, so what we choose from S_1 affects what we can choose from S_2 but not the *number* of things we could choose from S_2, we can use the **generalized product rule**. The generalized product rule tells us that if we are making a sequence of length k and we have $n_1, ..., n_k$ options for each position, then there are $n_1 * ... * n_k$ total sequences we can form.

A **permutation** of a set A is an ordered list of the elements of A. The number of permutations of n elements is $n!$, which we can prove with the generalized product rule.

Practice Problem(s)

1. How many strings of eight English letters are there

   a) that contain no vowels, if letters can be repeated?

   b) that contain no vowels, if letters cannot be repeated?

   c) that start with a vowel, if letters can be repeated?

   d) that start with a vowel, if letters cannot be repeated?

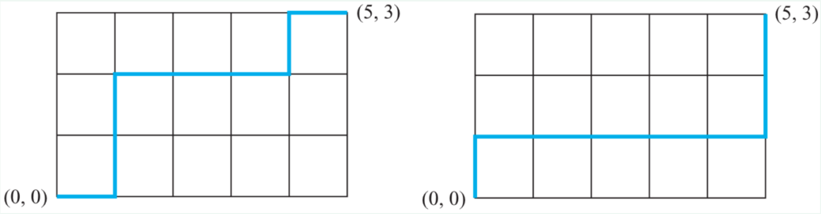
   e) that contain at least one vowel, if letters can be repeated?

   f) that contain exactly one vowel, if letters can be repeated?

   g) that start with X and contain at least one vowel, if letters can be repeated?

   h) that start and end with X and contain at least one vowel, if letters can be repeated?

2. Count the number of paths in the xy plane between the origin $(0, 0)$ and point $(5, 3)$. Each path consists of a series of steps, where each step is a move one unit to the right or a move one unit upward. Two such paths are illustrated below:



3. (This was a scrapped homework problem that one of the TAs thought was fun - an equivalent exam question would be one part of this problem)

   Contrary to popular belief, the dinosaurs invented cards. Consider a standard deck of 52 cards. The dinosaurs played a game where each dino received 4 cards. For this problem, the order of the cards dealt does not matter, but the suits do.

a) How many combinations of four card hands would a dino have **exactly** three of the same number (three of a kind)?

b) How many combinations of four card hands would a dino have two pairs of different numbered cards (two-pair)?

c) The dinos got bored and decided to invent the joker, a 53rd card that acts as a wild card, and can represent any number, but not a suit. (Ex: the joker can be a 2, but not the 2 of spades) **Assume that if the joker can be used to make a two-pair, it will.** Additionally, assume this is the only way the joker can be used, and otherwise it is just its own separate card with no suit.

   i) Repeat a) with the joker and the above assumption

   ii) Repeat b) with the joker and the above assumption

d) After a million years, the dinos got bored again. **Now, assume that if the joker can be used to make three of a kind, it will.**

   i) Repeat a) with the joker and the above assumption

   ii) Repeat b) with the joker and the above assumption

e) The dinos wanted to determine a winner for their game, and wanted the hand that was rarer, or had less combinations, to win. Lets say Dan has two pairs, and Brendan has three of a kind.

   i) List the winners for the three scenarios (a & b, c, d)

   ii) The dinos have decided they want freedom in how they use the joker. Given that the winner is the one with the rarest hand, how should the dinos use the joker? Briefly explain how/if a winner would be decided.

9.2 Binomial Coefficients and Theorem

The **binomial coefficient**, also called n *choose* k, is defined to be

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

for $n \geq k$ and $n, k \in \mathbb{Z}^+$.

The binomial coefficient $\binom{n}{k}$ counts the number of ways to choose k objects from n objects. Equivalently, its counts the number of subsets of size k of a set of size n.

Binomial Theorem: The coefficients of the terms in the polynomial $(x+y)^n$ are binomial coefficients, i.e.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Practice Problem(s)

1. What is the hundreds-place digit of 11^{2024}?
2. Find the middle term of $(\frac{p}{x} + \frac{x}{p})^9$

9.3 Counting Arguments

A **counting argument** shows that the LHS (lefthand side) and the RHS (righthand side) of some equation count the same thing. Instead of using algebraic manipulation, we explain why both sides ultimately count the elements of some set, just in different ways.

Importantly, if a question asks you to use a counting argument, you cannot use the definition of $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ or other algebraic arguments.

For instance, consider the following identity.

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Let S be a set with n elements. The LHS counts the number of ways to form a subset of S size k. Let x be some element of S. Each subset of S of size k either includes x or does not include x. If the subset includes x, then we need to pick $k-1$ other elements for the subset from the remaining $n-1$ elements, which we can do in $\binom{n-1}{k-1}$ ways. If the subset does not include x, then we still need to pick all k elements, and can do so from the remaining $n-1$ elements since we can't pick x, which we can do in $\binom{n-1}{k}$ ways. So, adding these together to get the RHS, this also counts the number of subsets of S of size k.

Practice Problem(s)

1. Prove via counting argument that $1 + 2 + \cdots + n = \binom{n+1}{2}$.
2. Prove via counting argument that $\binom{2n}{2} = 2\binom{n}{2} + n^2$.

9.4 Inclusion/Exclusion Formula:

The inclusion/exclusion formula provides a way of counting the size of a union of sets, and it is especially helpful if the sets overlap (and thus merely summing the sizes would result in over-counting.)

For two sets A and B, the inclusion/exclusion formula says that

$$|A \cup B| = |A| + |B| - |A \cap B|$$

While the formula for three sets A, B, and C is

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|$$

Going further, we can repeat this process for any number of sets, alternating between adding and subtracting the sizes of sets.

Practice Problem(s)

1. How many 6-digit numbers are even or are divisible by 5?
2. How many positive integers less than 1000 are multiples of 3, 5, or 7?
3. How many elements are in $A_1 \cup A_2$ if there are 12 elements in A_1, 18 elements in A_2, and

   a) $A_1 \cap A_2 = \emptyset$?

   b) $|A_1 \cap A_2| = 1$?

   c) $|A_1 \cap A_2| = 6$?

   d) $A_1 \subseteq A_2$?

9.5 Counting Donuts

The number of ways to distribute m identical objects among n distinct groups is

$$\binom{m+n-1}{n-1}.$$

Why is this? We can uniquely represent such a distribution with a 0/1 string of length $m+n-1$ that has exactly m 0's:

Let the m 0's represent the objects. The remaining $n-1$ bits in the string are 1's. Let all of the 0's to the left of the first 1 belong to group 1. Then, let all the 0's between the first 1 and the second 1 belong to the second group. Continue determining group membership in this fashion. Below is a diagram illustrating this. Note that, since 1's two and three are adjacent, nothing is in group 3.

$$\underbrace{0\ldots0}_{\text{group 1}}1\underbrace{0\ldots0}_{\text{group 2}}11\underbrace{0\ldots0}_{\text{group 4}}10\ldots01\ \underbrace{000}_{\text{group } n}$$

What's important to note is that (1) any distribution we choose can be represented with some length $m+n-1$ binary string, and (2) any such binary string represents a valid distribution of m identical objects into n distinct groups under this interpretation. In other words, the distributions and binary strings are in bijection with each other, meaning we can count one by counting the other.

We know how to count such binary strings: it is simply the number of ways you can choose $n-1$ of the bits to be 1's, leaving the other m bits to be 0's: $\binom{m+n-1}{n-1}$.

Practice Problem(s)

1. Each of the counting problems below can be solved with stars and bars. For each, say

what outcome the diagram

$$* * * \mid * \mid\mid ** \mid$$

represents, if there are the correct number of stars and bars for the problem. Otherwise, say why the diagram does not represent any outcome, and what a correct diagram would look like:

a) How many ways are there to select a handful of 6 jellybeans from a jar that contains 5 different flavors?

b) How many ways can you distribute 5 identical lollipops to 6 kids?

c) How many 6-letter words can you make using the 5 vowels in alphabetical order?

d) How many solutions are there to the equation $x_1 + x_2 + x_3 + x_4 = 6$?

2. Using the digits 2 through 9 (inclusive), find the number of 4-digit numbers such that:

   a. Digits cannot be repeated and must be written in increasing order. (Increasing means strictly increasing. For example, the digits of 134 are increasing, but the digits of 133 are not.)

   b. Digits can be repeated and must be written in non-decreasing order. (Now the digits don't need to be strictly increasing; 133 has digits non-decreasing.)

3. How many integer solutions to $x_1+x_2+x_3+x_4 = 25$ are there for which $x_1 \geq 1, x_2 \geq 2, x_3 \geq 3$ and $x_4 \geq 4$?

4. Consider functions $f : \{1, 2, 3, 4, 5\} \to \{0, 1, 2, \cdots, 9\}$.

   a. How many of these functions are strictly increasing? Explain. (A function is strictly increasing provided if $a < b$, then $f(a) < f(b)$.)

   b. How many of the functions are non-decreasing? Explain. (A function is non-decreasing provided if $a < b$, then $f(a) \leq f(b)$.)

9.6 Pigeonhole Principle

Pigeonhole Principle: If we put $k + 1$ objects into k boxes, then some box has at least 2 objects. More generally, if we place n objects into k boxes, then some box must have at least $\lceil \frac{n}{k} \rceil$ objects.

Another way we can think about the Pigeonhole Principle is this. It tells us that if we have a function, $f : |X| \to |Y|$, such that the cardinality of X is n and the cardinality of Y is k, then there is some $y \in Y$ such that the number of $x \in X$ that map to y is greater than or equal to $\lceil \frac{n}{k} \rceil$.

Pigeonhole principle basically says that *some* box must have the average number of items per box (assume for the sake of contradiction that this were not the case—what would have to be true?) We get the ceiling function because we can't have fractional objects—objects must remain whole as they are placed into boxes.

Practice Problem(s)

1. What is the minimum number of times you must roll a six-sided dice before you can guarantee that 10 or more of the rolls resulted in the same number?
2. Prove that any set of seven distinct natural numbers contains a pair of numbers whose sum or difference is a multiple of 10.
3. A game comes with 40 six-sided dice (each numbered 1 to 6). Suppose you roll all 40 dice. Prove that there will be at least seven dice that land on the same number.

10 Probability

10.1 Definitions

- A countable **sample space** S is a countable nonempty set. Don't worry too much about the countable part. Usually, we'll work with finite sets. If you're curious about when an infinite set is considered "countable," see the end of recitation 2.
- An element $\omega \in S$ is called an **outcome**.
- A **probability function** on S is a function $\Pr : S \to \mathbb{R}$ with the following two properties:
  1. $\Pr(\omega) \geq 0 \; \forall \omega \in S$
  2. $\sum_{\omega \in S} \Pr(\omega) = 1$
- Together, a sample space and probability function are called a **probability space**.
- A subset $E \subseteq S$ is called an **event**. The probability of E is defined as $\Pr(E) = \sum_{\omega \in E} \Pr(\omega)$
- A probability space is **uniform** if all outcomes have equal probability, that is $\forall \omega \in S$, $\Pr(\omega) = \frac{1}{|S|}$. If this is true, for any event E, $\Pr(E) = \frac{|E|}{|S|}$.

Practice Problem(s)

1. Determine which of the following are valid probability spaces (as defined above):

   a) The tuple $(\{1, 2, 3, 4\}, \Pr)$ where $\Pr(n) = \frac{1}{2\text{rem}(n,2)+4}$ for all $n \in \{1, 2, 3, 4\}$.

   b) The tuple $(\mathbb{N}, \Pr)$ where $\Pr(n) = 2^{-n}$ for all $n \in \mathbb{N}$.

10.2 Rules

Here are some rules about the probabilities of events. You should be comfortable working with them. Some of them are very closely related to counting rules!

- **Sum Rule**: If $E_1, ..., E_n$ are disjoint events (that is, there are no outcomes which are members of more than one event) then $\Pr(E_1 \cup ... \cup E_n) = \sum_{i=1}^{n} \Pr(E_i)$
- **Complement Rule**: For any event E, $\Pr(\overline{E}) = 1 - \Pr(E)$
- **Difference Rule**: For events A and B, $\Pr(B - A) = \Pr(B) - \Pr(B \cap A)$
- **Inclusion-Exclusion**: For events A and B, $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
- **Union Bound**: For events $E_1, ..., E_n$, $\Pr(E_1 \cup ... \cup E_n) \leq \Pr(E_1) + ... + \Pr(E_n)$

Practice Problem(s)

1. Show that $P(A) = P(A \cap B) + P(A \cap \overline{B})$. (This should be very straightforward from one of the rules.)

10.3 Conditional Probability and Independence

The **conditional probability** $\Pr(A|B)$ is the probability that A happened given that we know B did. Essentially, we limit our set of possibilities to the outcomes in B and find how many of those are also in A. It is defined as

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Bayes' Rule is a useful rearrangement of the definition of conditional probability and tells us

$$\Pr(A|B) = \frac{\Pr(B|A) \cdot \Pr(A)}{\Pr(B)}$$

A is **independent** of B if knowing B occurred does not give us any additional information about whether A did. Mathematically, A is independent of B if $\Pr(A|B) = \Pr(A)$, or if $\Pr(B) = 0$.

A set of events $\{E_1, ..., E_n\}$ is **mutually independent** if for every subset S of the set of events, the probability of the intersection of the events is equal to the product of the probabilities of each event.

For any set, pairwise independence of the events does *not* guarantee mutual independence!

Practice Problem(s)

1. Let A, B, and C be mutually independent events. Show that $A \cap B$ and C are independent, and show that $A \cup B$ and C are independent.

2. It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

3. Suppose that we have two pints of Half Baked^{TM}, each containing a mix of cookie dough ice cream and fudge brownie ice cream. One pint contains three times as much cookie dough as fudge brownie. The other pint contains three times as many fudge brownie as cookie dough. Suppose we choose one of these pints at random. From this pint, we select five spoonfuls at random with replacement (i.e. each selected spoonful is returned to its respective pint, because we're looking for the perfect bite). Each spoonful necessarily has either one brownie or one cookie. The result is that we find 4 spoonfuls of cookie dough and one spoonful of fudge brownie. What is the probability that we were using the pint with mainly cookie dough?

4. Of the high blood pressure patients in a particular clinic, 62 % are treated with medication X, the remainder with medication Y. It is known that 1.4% of the patients using medication X suffer from fainting spells, as do 2.9% of the patients using medication Y. A patient known by the clinic to have high blood pressure suffers a fainting spell but does not remember which medication she is on. Which medication is she more likely to be taking?

5. In a best-of-three tournament, the local C-league hockey team wins the first game with probability 1/2. In subsequent games, their probability of winning is determined by the outcome of the previous game. If the local team won the previous game, then they are invigorated by victory and win the current game with probability 2/3. If they lost the previous game, then they are demoralized by defeat and win the current game with probability only 1/3. What is the probability that the local team wins the tournament, given that they win the first game? (solution 17.3 of textbook)

10.4 Random Variables

A **random variable** is a function from outcomes of a probability space. Usually, the codomain of the function is the real numbers or integers.

Some examples are a mapping from a sequence of coin flips to the number of heads that occur in the sequence or mapping from a person to the number of emails in their inbox.

An **indicator random variable** “indicates" whether an event occurs by mapping all outcomes to either 1 or 0. These are also referred to as Bernoulli variables.

Practice Problem(s)

1. A fair coin is tossed repeatedly until either it lands heads or a total of five tosses have been made, whichever comes first. Let X denote the number of tosses made. What is the probability distribution for X?

2. 100 homeworks are on the table, with two questions to be graded. Andy is in charge of grading question one and Patrick is in charge of grading question two. First, Andy grades some homeworks at random; each homework has probability 0.3 of being graded. Next, Patrick randomly grades half of the homeworks. That is, he grades 50 homeworks. Assume Andy and Patrick make their choices independently.

   a) Let N be the number of homeworks that Andy graded. What is $\mathbb{E}[N]$?

   b) Let M be the number of homeworks that Andy graded and Patrick did not grade. What is $\mathbb{E}[M]$?

   c) Let a "good pair" be a pair of adjacent homeworks such that both questions are graded. Let P be the number of good pairs. What is $\mathbb{E}[P]$?

3. A supervisor in a manufacturing plant has three men and three women working for him. He wants to choose two workers for a special job. He decides to select the two workers at random. Let Y denote the number of women in his selection. Find the probability distribution for Y.

10.5 Expected Value

The **expected value** (or just expectation) of a random variable is a probability-weighted average of its values. That is, if one value is far more likely to occur, we weight it higher in the average. The expected value of a random variable R is defined as

$$\mathbb{E}[R] = \sum_{\omega \in S} R(\omega) \Pr(\omega)$$

It can also be useful to think about summing over the output of R rather than the events in S. This is an equivalent definition of expected value:

$$\mathbb{E}[R] = \sum_{x \in \text{range } R} x \cdot \Pr(R = x)$$

The **conditional expectation** of a random variable R given an event A is defined as

$$\mathbb{E}[R|A] = \sum_{x \in \text{range} R} x \cdot \Pr(R = x|A)$$

Perhaps the most important property from this section is **linearity of expectation**. For random variables $R_1, .., R_n$ and real numbers $a_1, ..., a_n$,

$$\mathbb{E}[a_1 R_1 + ... + a_n R_n] = a_1 \mathbb{E}[R_1] + ... + a_n \mathbb{E}[R_n]$$

Practice Problem(s)

1. A coin is biased so that the probability a head comes up when it is flipped is 0.6. What is the expected number of heads that come up when it is flipped 10 times?

2. The final exam of a discrete mathematics course consists of 50 true/false questions, each worth two points, and 25 multiple-choice questions, each worth four points. The probability that Linda answers a true/false question correctly is 0.9, and the probability that she answers a multiple-choice question correctly is 0.8. What is her expected score on the final?

3. Suppose that we roll a pair of fair dice until the sum of the numbers on the dice is seven. Let X be the random variable denoting the number of rolls.

   a) What is $\mathbb{E}[X]$?

   b) What is $\mathbb{E}[X^2]$?

10.6 Variance

Sometimes measuring the mean (expectation) of a random variable doesn't give us enough information: it can be helpful to know how much we expect the variable to *stray* from its average.

Markov's inequality gives a generally coarse estimate of the probability that a random variable takes a value much larger than its mean.

Theorem (Markov). If R is a nonnegative random variable, then for all $x > 0$,

$$\Pr[R \geq x] \leq \frac{\mathbb{E}[R]}{x}.$$

Expressed differently:

Corollary. If R is a nonnegative random variable, then for all $c \geq 1$,

$$\Pr[R \geq c \cdot \mathbb{E}[R]] \leq \frac{1}{c}.$$

That is: the probability of R being more than c times its mean is at most $1/c$.

A related notion is that of *variance*:

Definition. The *variance* $\mathrm{Var}[R]$ of a random variable R is defined to be $\mathbb{E}[(R - \mathbb{E}[R])^2]$.

Unpacking this from the inside out: $R - \mathbb{E}[R]$ is a random variable measuring the distance between R and its mean at each outcome. Averaging the square of this gives us a sense of, overall, how far R tends to be from its mean.

There is an equivalent way to state this:

Lemma. For any random variable R,

$$\operatorname{Var}[R] = \mathbb{E}[R^2] - (\mathbb{E}[R])^2.$$

This leads us to state Chebyshev's theorem, an application of Markov's inequality:

Theorem (Chebyshev). Let R be a random variable and $x \in \mathbb{R}^+$. Then

$$\Pr\left[|R - \mathbb{E}\left[R\right]| \geq x\right] \leq \frac{\operatorname{Var}[R]}{x^2}.$$

Practice Problem(s)

1. Let X be the random variable that denotes the number that comes up when a fair die is rolled. What is the variance of X?

2. What is the variance of the number of heads that come up when a fair coin is flipped 10 times?

3. Suppose that the number of tin cans recycled in a day at a recycling center is a random variable with an expected value of 50,000 and a variance of 10,000.

   a) Use Markov's inequality to find an upper bound on the probability that the center will recycle more than 55,000 cans on a particular day.

   b) Use Chebyshev's inequality to provide a lower bound on the probability that the center will recycle 40,000 to 60,000 cans on a certain day.