CS 0220 — Spring 2023 **Final Review**

1 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. Additionally, the example proofs we provide in this review guide may be informal, or stripped to their bare bones. They strive to convey an idea, but are not necessarily paragons of perfect proofs. We suggest looking at the website or the homework solutions for completely polished proofs.

2 Proof Techniques

2.1 Direct Proof

We directly use our statements to imply ($\Rightarrow$) that our conclusion is correct.

Example: Prove the claim that the product of two odd numbers is odd.

2.2 Contradiction

Say we have some proposition T that we are trying to prove. Here is how we can prove it by contradiction:

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

Often T is of the form "If p then q." In this case, assume that p is true and q is false to reach a contradiction. Often, this contradiction will be of the form "If p is true and q is false then p is false. This is a contradiction as p cannot both be true and false.

Example: Prove the claim that if n^2 is even, then n is even.

Example: 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 $X_1 = X_2$.

2.3 Proof by Cases

Example: Prove the claim that there exists x, y irrational such that x^y is rational.

2.4 Counterexample

Counterexamples help us prove that something is not true.

For example, suppose Kristy makes the claim that if xy is rational then x and y are rational.

Kareem can disprove her claim by coming up with a counterexample. For example, if $x = \sqrt{2}$ and $y = \sqrt{2}$, then $xy = 2$, which is rational.

However, you **cannot** prove a claim by showing one example of it. Kareem has not proven that x and y are irrational, he has just shown that they are not always rational.

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

Example: Prove the claim that $\mathcal{P}(A \cup B) \neq \mathcal{P}(A) \cup \mathcal{P}(B)$.

2.5 Proof by Element Method

How do you prove that $A = B$? 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.
Example: Prove the claim that $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.

2.6 Bijective Proof

Example provided later. (Put together proofs for injectivity and surjectivity further down in this sheet, and then add conclusion.)

2.7 Bidirectional Proof

If a claim is of the form "A if and only if B," you must prove both "if A, then B" and "if B, then A"

2.8 Inductive Proof

More detail later.

3 Logic

Here's information about logic.

3.1 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 OR NOT 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 AND NOT p.

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

P	Q	NOT P	P AND Q	P OR Q	P XOR Q	P IMPLIES Q	P IFF 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

3.2 Implication

In the formula P IMPLIES Q, we call P the **hypothesis** and Q the **conclusion**. P IMPLIES Q is logically equivalent to NOT P OR Q. In words, this means that for P IMPLIES 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 IMPLIES 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 that where the statement doesn't fit is the second, which is why P IMPLIES Q is only false when P is true and Q is false.

- NOT Q IMPLIES NOT P is called the **contrapositive** of P IMPLIES Q and is 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 IMPLIES P is called the **converse** of P IMPLIES Q. It is **not** logically equivalent to $P \Rightarrow Q$. If both a statement and its converse are true, then the biconditional P IFF Q is true.

3.3 Normal Forms

A literal is a variable or its negation.

We say a proposition is in **DNF (disjunctive normal form)** when it is the disjuction (clauses ORed together) of conjuctions (literals ANDed together).

We say a proposition is in **CNF (conjunctive normal form)** when it is the conjuction (clauses ANDed together) of disjuctions (literals ORed together).

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 AND Q AND NOT R) OR (P AND NOT Q AND NOT R)
OR (NOT P AND NOT Q AND R) OR (NOT P AND NOT Q AND NOT R)
CNF: (NOT P OR NOT Q OR NOT R) AND (NOT P OR Q OR NOT R)
AND (P OR NOT Q OR NOT R) AND (P OR NOT Q OR 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 AND 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 OR 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 CNF or DNF. We can also rewrite any logical expression to be in CNF or DNF.

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.

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

   i. $A \subseteq A$

   ii. $\{\} \subseteq A$

   iii. $\{\} \in A$

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

c. 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 $B - A$ 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!

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).

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 $a \in A$.

5 Relations

5.1 Relation on $A \times B$ vs. Relation on A

A *relation* R on the sets A and B is a subset of the Cartesian product $A \times B$. A relation R on the set A is a subset of the Cartesian product $A \times A$.

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

5.2 Notation

aRb and $(a, b) \in R$ are both compact ways of saying the same thing: a is related to b in R.

Remember that a relation is a set of ordered pairs, not a description of the way elements are linked. For example, the following is a relation on $\mathbb{Z}$:

$$R = \{(x, y) \mid x \leq y\}.$$

However, "$\leq$" is not a relation.

5.3 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!)

5.4 Symmetry and Transitivity

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

A relation R on A is *transitive* if for all $a, b, c \in A$, 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 order 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 both symmetric and transitive, as there are no ordered pairs in the empty relation to satisfy the if-clause of either definition.

5.5 Equivalence Relation

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

5.6 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 a partition of the set.

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. $B_i \cap B_j = \emptyset \quad \forall i, j$ where $i \neq j$.

5.7 Examples

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.

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.

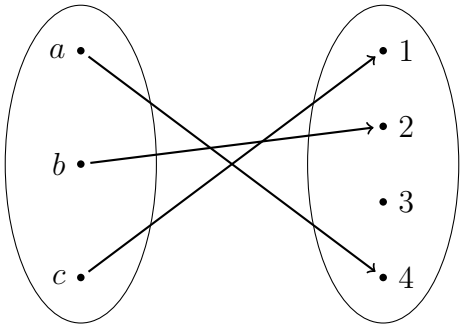
It's important to note that a function is characterized not only by the "rule" that maps inputs to outputs, but also by the domain and codomain.

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! Another way to think about this: if you give me an element in the image of the function, I can tell you without a doubt which element mapped to it. Why? Because there won't be more than one element that maps to it.

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:

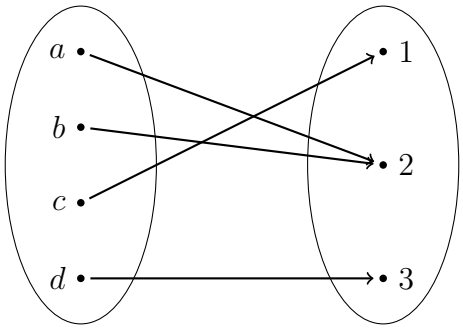
1. Consider two arbitrary distinct elements in the domain. Show that they must map to distinct outputs.

2. Consider two equal elements in the image of f (say, $f(a)$ and $f(b)$.) Show that $a = b$.

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.

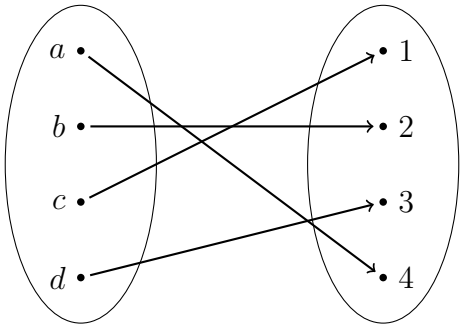
6.4 Bijectivity

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.



6.5 Intuition and Examples

For each of the following, state if it is a function, injections, surjection, or neither.

(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 : First Year Students at Brown $\to$ First Year Dorms at Brown where f(student) = the dorm that the student lives in.

(e) f : Students at Brown $\to$ Banner IDs of current Students where f(student) = the banner ID of student.

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

(g) f : Library at Brown $\to \mathbb{Z}$ where f(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 $f(X) = |X|$.

7 Induction

7.1 Example 1: Template and Weak Induction

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 ***different cases*** *or even* ***contradiction***

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

7.2 Example 2: Strong Induction

Example: 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.

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 $a \equiv b \mod m$, then

- $a + c \equiv b + c \mod m$ for $c \in \mathbb{Z}$
- $ac \equiv bc \pmod{m}$ for $c \in \mathbb{Z}$
- $a^n \equiv b^n \mod m$ for $n \in \mathbb{Z}^+$

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

- $a + c \equiv b + d \pmod{m}$
- $ac \equiv bd \pmod{m}$

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.

Theorom 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.

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.) For something to divide 1, it must itself be 1. 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 $a \bmod p$.

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.

8.7 Example Problems

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.

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.

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.

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

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.

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.

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

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.

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 $\Pr : S \Rightarrow \mathbb{R}$ 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, $\Pr(E) = \frac{|E|}{|S|}$.

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 $E_1, ..., E_n$, $\Pr(E_1 \cup ... \cup E_n) \leq \Pr(E_1) + ... + \Pr(E_n)$

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

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

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!

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.

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:

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

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

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

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

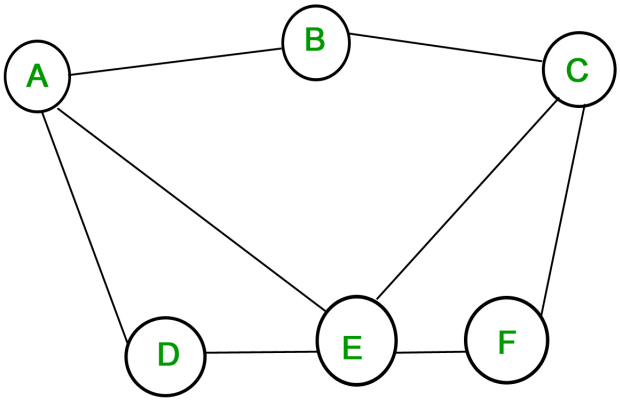
$$\mathbb{E}[a_1 R_1 + ... + a_n R_n] = a_1 \mathbb{E}[R_1] + ... + a_n \mathbb{E}[R_n]$$

11 Graph Theory

11.1 Basic Definitions

A **graph** is two related sets: $V(G)$, the vertex set, and $E(G)$, the edge set. Each element of $E(G)$ is a set containing exactly two elements of $V(G)$.

We often visualize graphs by drawing the vertices as dots and the edges as lines between them. Here is an example of a graph with vertex set $\{A, B, C, D, E, F\}$ and edge set $\{\{A, B\}, \{B, C\}, \{A, E\}, \{C, E\}, \{A, D\}, \{C, F\}, \{D, E\}, \{E, F\}\}$



Note that we have defined an edge as a set of cardinality two: this means that a vertex cannot have an edge to itself, and there is no sense of direction in edges. Additionally, since edges are contained in a set, there can be at most one edge between any two vertices.

- If $u, v \in V(G)$, u is **adjacent to** v if $\{u, v\} \in E(G)$. (Note that by this definition a vertex is not adjacent to itself)
- The **degree** of a vertex is a count of the number of vertices it is adjacent to. Formally, for a vertex v, $deg(v) = |\{u|\{v, u\} \in E(G)\}|$.
- The **empty graph** on n vertices has an empty edge set.
- The **complete graph** on n vertices K_n has all possible edges between vertices, that is $E(G) = \{(u, v)|u, v \in E(G), u \neq v\}$. This means every pair of vertices is adjacent.
- A **path** is a sequence of vertices such that any two vertices that appear subsequently in the sequence are adjacent in the graph. For instance, in the graph above, (A, B, C, E) is a path from A to E. A path is simple if no vertices are repeated. We defined the length of a path by the number of vertices, not edges.
- Two vertices are **connected** if there exists a path between them (or if they are the same vertex).
- A **subgraph** G' of a graph G is a graph such that $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. Since G' is a graph, each edge in $E(G')$ must be between two vertices in G'.

11.2 Trees

- A **cycle** is a path that starts and ends with the same vertex. A cycle is **simple** if there are no other repeated vertices.
- A graph is **cyclic** if it contains a simple cycle and **acyclic** otherwise.
- A graph is **connected** if there is a simple path between each pair of vertices (that is, all vertices are connected).
- A **tree** is a connected, acyclic graph.
- A **forest** is an acyclic graph. In other words, a forest is a set of trees.
- A **leaf** of a tree is a vertex with degree 1.
- A **spanning tree** of a graph G is a subgraph T of G such that $V(T) = V(G)$ and T is a tree.

A tree is a graph G which is connected and acyclic. There are many other ways of saying this:

- For all $u, v \in V(G)$, there exists a unique path from u to v (if there are two distinct paths, G has a cycle)
- G is maximally acyclic (adding another edge would create a cycle)
- G is minimally connected (removing any edge would disconnect G)

Given a tree, we can define any vertex r to be the **root**. For any vertex v, we define the **depth** of v, $dep_r(v)$ to be the length of the simple path from v to the root. The **parent** of v is the vertex adjacent to it which is closer to the root.

There are n^{n-2} distinct trees on n vertices. We found that this was true by defining the **Prufer code** of a tree and constructing a bijection between the set of possible Prufer codes (strings of length $n-2$ where each position was 1-n) and the set of trees.

11.3 Tours

Definition: An *Eulerian tour* is like a cycle which uses every edge in a graph. The length of an Eulerian tour, since we define length by the number of vertices, is $|E(G)|+1$.

If a graph has an Eulerian tour, then all vertices have even degree. If all vertices have even degree and the graph is connected, it has a Eulerian tour. The proof of this was done in lecture.

Definition: A *Hamiltonian tour* is a cycle that visits every vertex exactly once (with the exception of the start/end vertex, which appears twice to make it a cycle). The length of a Hamiltonian tour is $V(G) + 1$.

11.4 Coloring

Definition: A *coloring* of a graph G is an assignment of a label or "color" to each vertex in the graph.

A coloring is proper or valid if for all $\{u, v\} \in E$, u and v are different color.

We say a graph has a k-coloring if it can be given a proper coloring with k colors. Figuring out what the smallest value k can be for a graph can be challenging!

A graph is **bipartite** if its vertices can be partitioned into two sets, A and B, such that all edges contain one vertex from A and one vertex from B. A bipartite graph is a 2-colorable graph.

We proved in class that if $d \geq deg(v)\ \forall v \in V(G)$, G has a $d+1$-coloring.

11.5 Solving Graph Theory Questions

When doing graph theory questions, it is often helpful to draw an example graph! Graphs can be hard to think about abstractly, and seeing one in front of you can go a long way when trying to reason about them.

Many graph theory questions involve induction. When doing induction on graphs, we often use a technique sometimes referred to as "build-down" induction. This is just regular induction, with a slightly different perspective. To review, here are the steps of an inductive proof:

1. Define the predicate $P(n)$.
2. Show that the base case is true.
3. Assume the inductive hypothesis is true. If you are using standard 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 $i \leq k$.
4. Show that $P(k+1)$ is true given the inductive hypothesis.

   Here is where build down induction differs from our usual perspective. A lot of the time, we start by invoking our inductive hypothesis, and manipulating the k case to get the $k+1$ case. In build down induction, we do the following instead:

   (a) Consider an arbitrary $k+1$ case.

   (b) Alter the $k+1$ case to obtain an k case.

   (c) Invoke the induction hypothesis, demonstrating that your property holds for the k case.

   (d) Recover the original $k+1$ case by undoing your alteration.

(e) Prove that the property still holds despite building back up to the original $k+1$ case.

This technique is especially useful when it is not clear how to address all $k+1$ cases in general when starting from an k case. For example, if we started with a graph with k vertices, how do we add a vertex to obtain a general graph with $k+1$ vertices? There are a lot of ways to add this extra vertex that we have to consider. To avoid this, we can start with an arbitrary $k+1$ case. Then, we build-down to an k case, at which point we can invoke our inductive hypothesis. Then, we build back up to the *original* $k+1$ case.

Note that build down induction can be used outside of graph theory.

Exercises:

1. Use a counting argument to prove that

$$\sum_{k=1}^{n} k = \binom{n+1}{2}.$$

2. Consider a graph G on six vertices. Prove that at least one of the following is true:

   (a) There exists a 3-cycle.

   (b) There exists three vertices with no edges between any two of them.

3. Suppose a graph G on $n \geq 4$ vertices has $\lceil \frac{n}{2} \rceil^2 + 1$ edges. Prove that G is not 2-colorable.

   Hint: If G is 2-colorable, what sets can we partition the vertices into?

   Hint: If it helps, you may assume that the product nm for $n, m \in \mathbb{Z}$ is maximal when $n = m$.