Basic Logic

1. Simplify each of the following propositions as much as possible (P, Q, R represent propositions):

   (a) $P \vee \text{true}$ (b) $P \vee \text{false}$ (c) $P \wedge \text{true}$ (d) $P \wedge \text{false}$

   (e) $P \vee \neg P$ (f) $P \wedge \neg P$

2. Compare the truth tables or Venn diagrams for the propositions $P \wedge Q$ and P, and use this to explain why $(P \wedge Q) \Rightarrow P$. In words, what does this tell us?
   Does $P \Rightarrow (P \wedge Q)$?

3. Complete, and briefly explain using a Venn diagram, the distributive laws for $\vee$ and $\wedge$:

   (a) $P \vee (Q \wedge R) =$

   (b) $P \wedge (Q \vee R) =$

4. Use the fact that $P \Rightarrow Q$ is equivalent to $Q \vee \neg P$, along with the results above, to simplify each of the following as far as possible:

   (a) $(P \Rightarrow Q) \wedge P$

   (b) $(P \Rightarrow Q) \wedge \neg Q$

   In words, what does each of these tell us?

5. Rewrite the following proposition in *twelve* different ways, using only the symbols given:
$$P \vee (Q \vee R).$$

*6. Using only P, Q, R, $\vee$, and $\wedge$ (along with () as necessary), how many ways can you find to rewrite:
$$P \wedge (Q \vee R).$$

7. Two of the following propositions are true, and two are false; determine which are which and briefly explain, then express each one in words.

   (a) $\forall\, y \in \mathbb{R}, \exists\, x \in \mathbb{R}$ such that $x^3 = y$.

   (b) $\forall\, y \in \mathbb{Z}, \exists\, x \in \mathbb{Z}$ such that $x^3 = y$.

   (c) $\exists\, b \in \mathbb{R}$ such that $\forall\, a \in \mathbb{R},\ b > a$.

   (d) $\forall\, a \in \mathbb{Z}, \exists\, b \in \mathbb{Z}$ such that $b > a$.

8. Use the logical laws of negation to pass the $\neg$ symbol in each proposition below past every other logical connective and quantifier:

   (a) $\neg\big[Q \vee (P \wedge \neg R)\big]$.

   (b) $\neg\big[\forall\, M, \exists\, N \text{ such that } \forall\, x,\ (x > N \Rightarrow 2x > M)\big]$.

Logic and sets

9. Which two of our logical connectives and/or constants don't have analogues in the world of sets?
   Briefly, why not?

10. Use the logical definitions of our set operations to unravel each of the following into propositions with *no set operations* other than $\in$; simplify the resulting propositions, if possible:

    (a) $a \in A \cup (B \setminus C)$ (b) $y \in K \triangle (M \cap N)$ (c) $x \in (X \setminus Y) \cup (X \cap Y)$

    (d) $(A \cup B) \subset C$ (e) $(A \setminus C) = C \cap B$

Direct Proofs

Define the following subsets of $\mathbb{Z}$:

$E = \{ x \in \mathbb{Z} : \exists\, k \in \mathbb{Z} \text{ such that } x = 2k \}$ *even*

$O = \{ x \in \mathbb{Z} : \exists\, k \in \mathbb{Z} \text{ such that } x = 2k+1 \}$ *odd*

$T = \{ 3k : k \in \mathbb{Z} \}$ *threven*

$U = \{ 6k+1 : k \in \mathbb{Z} \}$ *unty*

$Q = \{ 6k+5 : k \in \mathbb{Z} \}$ *quinty*

Remember that the variables used in the set notations above are not significant—don't ever use the same variable to represent more than one thing in a given context!

Prove the following propositions:

1. $x, y \in O \;\Rightarrow\; x + y \in E$.
2. $x \in E \wedge y \in O \;\Rightarrow\; xy \in E$.
3. $x \in O \;\Rightarrow\; \frac{x+1}{2} \in \mathbb{Z}$.
4. $\mathbb{Z} = E \cup O$. [Hint: Use (*) with $b = 2$ and split cases.]
5. $U \subset O$
6. $x \in T \;\Rightarrow\; 2x - 5 \in U$
7. $x \in Q \;\Rightarrow\; x^2 \in U$
8. $x \in Q \wedge y \in U \;\Rightarrow\; x - y \in E$
9. $\mathbb{Z} = E \cup T \cup U \cup Q$ [Hint: Use (*) with $b = 6$ and split cases.]
10. $x \in Q \;\wedge\; y \in U \;\Rightarrow\; x + y \in E \cap T$

* Keep in mind the *division algorithm:* Given any $a, b \in \mathbb{Z}$ with $b > 0$, we can write $a = bq + r$ for some $q, r \in \mathbb{Z}$ with $0 \leq r < b$.

$\frac{a}{b} = q + \frac{r}{b}$;
q is the quotient and
r is the remainder.

Proof by Induction

The principle of mathematical induction: to prove a proposition of the form $\forall\, n \geq 0,\ P(n)$, we can instead prove: ① $P(0)$

and ② $\forall\, n \geq 0,\ P(n) \Rightarrow P(n+1)$.

Take care substituting $n \rightsquigarrow (n+1)$

1. Use induction on n to prove that $\forall\, n \geq 0,\ 1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{1}{6}(2n^3 + 3n^2 + n)$.

2. Suppose that $r \neq 1$; use induction on n to prove that $\forall\, n \geq 0,\ 1 + r + r^2 + \cdots + r^n = \dfrac{1 - r^{n+1}}{1 - r}$.

3. Use induction on n to prove the following inequalities:
   (a) $\forall\, n \geq 10,\ 100n < 2^n$
   (b) $\forall\, n \geq 4,\ 2^n < n!$ Recall that $n! = 1 \cdot 2 \cdot 3 \cdots n$
   (c) $\forall\, n \geq 0,\ (1+x)^n \geq 1 + nx$ Suppose that $x > -1$ for this part!

. . . and the Fibonacci sequence

The *Fibonacci sequence* $F_0, F_1, F_2, F_3, \ldots$ is defined by $F_0 = 0$,

$F_1 = 1$, and

$$\forall\, n \geq 1,\ \boxed{F_{n+1} = F_n + F_{n-1}}.$$

4. Use the definition above to compute the first eleven terms of the Fibonacci sequence, $F_0, F_1, \ldots, F_{10}$.

5. Prove the following Fibonacci identities using induction on n:
   (a) $\forall\, n \geq 0,\ F_0 + F_1 + F_2 + \cdots + F_n = F_{n+2} - 1$
   (b) $\forall\, n \geq 0,\ F_0^2 + F_1^2 + F_2^2 + \cdots + F_n^2 = F_n F_{n+1}$
   (c) $\forall\, n \geq 0,\ F_{2n+1} = F_{n+1}^2 + F_n^2 \ \wedge\ F_{2n+2} = F_{n+2}^2 - F_n^2$
   (*d) $\forall\, n \geq 1,\ F_{n+1}F_{n-1} = F_n^2 + (-1)^n$

6. Note that the results of problem 2(c) allow us to compute later Fibonacci numbers (with both odd and even indices) in terms of *much* earlier ones.
   (a) Use your answers to problem 1 to double-check what they say about F_9 and F_{10}.

   Feel free to use a calculator or computer to help with the arithmetic in parts (b,c,d)!

   (b) Use them to find the values of F_{15}, F_{16}, and F_{17}.
   (c) Now use them again to find the values of F_{31}, F_{32}, and F_{33}.
   (d) How do your three values in part (b) relate to one another?
   How about those in part (c)?

Graphs: Cycles and Trees

0. Keep playing with the online exercises to solidify the basic concepts and terminology! Your intuition for these concepts will be built largely visually, and that be essential in helping to you work through formal details without getting lost in symbols.

> I've transitioned to using words in the propositions below: *if-then*, *every*, *there is*, *has a/an*, etc. Practice reading these as the appropriate logical constructions $\Rightarrow$, $\forall$, and $\exists$, as you should already know exactly how direct proofs of such statements proceed!
>
> Recall that the ***contrapositive*** of $P \Rightarrow Q$ is $\neg Q \Rightarrow \neg P$ (these two propositions are logically equivalent), and that we can also prove $P \Rightarrow Q$ by ***contradiction***, showing that $\neg[P \Rightarrow Q]$ (i.e., $P \wedge \neg Q$) is *false* by supposing this and deducing a false statement.

1. Prove the following basic results about cycles and acyclic graphs:
   (a) Prove that for any graph G, if G has a loop or a parallel edge, then it is *not* acyclic.
   (b) Change the blue bit above to its contrapositive—what does this tell us?
   (c) What is the minimum number of vertices in a *simple* cycle? Why?
   (d) Prove that for every simple cycle C, there exist vertices v_0, v_1 of C with two distinct paths connecting them (a picture will suffice).

2. Prove the following fundamental results regarding trees, using induction on the number $n \geq 1$ of vertices:
   (a) Every tree T is a simple graph.
   (**b) For every tree T and each pair of vertices v_0, v_1 of T, there is a *unique* path in T connecting v_0 and v_1.
   (c) If T is a tree, then $n(T) = e(T) + 1$.
   $n(T)$ and $e(T)$ represent the numbers of vertices and edges of T, respectively.

3. Deduce the following results about trees:
   (a) Every tree T is nonempty. Use the definition of a tree.
   (b) Every tree T is acyclic.
   Strategy: see 2(b) & 1(d)—suppose that T contains a cycle C to obtain a ***contradiction***!
   (c) Every tree T is connected. See 2(b).
   (d) If a graph G has a spanning tree T, then G is connected. See previous part.

4. Prove that if G is nonempty and connected, and if G is *not* a tree, then G contains a cycle.

   Strategy: start by taking a spanning tree T for G,
   and deduce that G has an edge not in T;
   use 2(b) and this edge to form a cycle in G.

   Change the blue bit to its contrapositive; how does the resulting statement relate to problem 3(a,b,c)?

5. Suppose that we add one edge to a tree T (without changing its set of vertices) to obtain a new graph G.
   (a) Explain why (e.g., using problem 2(c)) G cannot be a tree; what does problem 4 then tell us?
   (b) Precisely what could we *remove* from G to get back to having a [possibly different] tree?

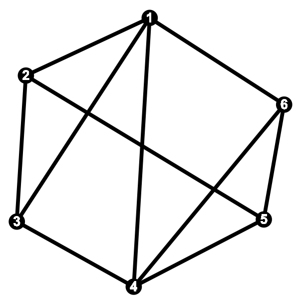
6. Suppose that we number the vertices of a simple graph G as $1, 2, 3, \ldots n$,
   and that we have a list of pairs of integers $\{ v_0, v_1 \}$ representing the edges of G.
   Outline algorithms for doing each of the following:
   (a) Finding the components of a graph G.
   (b) Finding a spanning tree for a connected graph G.
   (c) Finding a spanning forest for a [possibly disconnected] graph G.

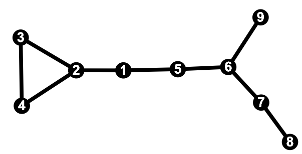
Graph Implementation

1. How do we represent a graph via an ***adjacency matrix***?
   (a) How does this provide us with convenient storage of node- and/or edge-weights?
   (b) In general, what affects whether an adjacency matrix a better or worse choice?

2. How do we represent a graph via ***adjacency lists***?
   (a) In general, what affects whether an adjacency list a better or worse choice?

3. Write out matrix and adjacency-list representations for each of the following graphs, and compare:

   (a) 

   (b) 

Graph Applications

4. What is a ***trie***, and what purpose does a trie serve?
   (a) Form a trie for the (spelled-out) numbers from zero to thirteen in a tree, and envision what text-completion would look like with this trie.
   (b) Think about tries next time you get auto-completion from a device!

5. How does ***Huffman encoding*** work for data, and what makes it useful?
   (a) Given the frequency data to the right for the letters in American English, find the associated Huffman encoding.
   (b) `Count up the total bits used to encode these letters via your encoding above.`
   How did it do, relative to, say, 5 bits per letter?
   (c) Will this encoding work well for all sets of words?
   Why or why not?

E	12.02	D	4.32	P	1.82
T	9.10	L	3.9	B	1.49
A	8.12	U	2.88	V	1.11
O	7.68	C	2.71	K	0.69
I	7.31	M	2.61	X	0.17
N	6.95	F	2.30	Q	0.11
S	6.28	Y	2.11	J	0.10
R	6.02	W	2.09	Z	0.07
H	5.92	G	2.03		

6. What is a ***Markov chain***?
   (a) How can we represent a Markov chain by an edge-weighted digraph?
   What are the rules for the edge-weights, and why?
   (b) Briefly explain MCMC (Markov Chain Monte Carlo) as a computational means of helping to understand a large Markov chain.
   (c) Draw an edge-weighted digraph for the following super-simplified flow of a basketball [or soccer/hockey/lacrosse] game:

   States: ① Team A has the ball ② Team B has the ball
   ③ Team A scores ④ Team B scores

   Transitions: When either team has the ball, 40% of the time they score, 20% of the time they miss but recover the ball, and 40% of the time, the other team gets the ball via a rebound or turnover.

   After a team scores, the other team gets the ball 100% of the time.

Relations on sets: in principle

1. If A is a set, what do we mean by a [binary] ***relation*** on the set A?
   (a) Conceptually, what *logical* concept we've already encountered is the same thing as a relation?
   (b) Give some examples of relations...
      (i) on the set $\mathbb{Z}$
      (ii) on the set of vertices in a graph G
      (iii) on the set of subgraphs of a graph G

2. Properties of relations:
   (a) What does it mean for a relation to be ***reflexive***?
   Give examples of reflexive and non-reflexive relations from those you listed above.
   (b) What does it mean for a relation to be ***symmetric***?
   Give examples of symmetric and non-symmetric relations.
   (c) What does it mean for a relation to be ***transitive***?
   Give examples of transitive and non-transitive relations.

3. Equivalence relations:
   (a) What is an ***equivalence relation*** on a set A?
   (b) How does an A give us an ***equivalence class*** (often written $[a]$, $\overline{a}$, or $\mathbf{a}$) for each element $a \in A$?
   (c) What properties make these equivalence classes a ***partition*** of the set A?
   How do we typically denote the collection of *all* equivalence classes of A under some relation $\sim$?
   (d) How does *any* relation R "generate" an equivalence relation, both intuitively and formally?

Integer relations: divisibility, congruence, and modular arithmetic

4. Define a relation $\mid$ on $\mathbb{Z}$ via: $a \mid b$ (read "a divides b") just when $\underline{\exists\, k \in \mathbb{Z} \text{ such that } b = ka}$.
   (a) Prove that $4 \mid 12$, $6 \mid 12$, and $12 \mid 12$, but $24 \nmid 12$ and $5 \nmid 12$.
   (b) Which of the above properties does this relation possess?

5. Fix a positive integer n.
   We say that $a, b \in \mathbb{Z}$ are "congruent modulo n", written $a \equiv b \pmod{n}$, just when $\underline{n \mid (b - a)}$.
   (a) Prove that $6 \equiv 1 \pmod{5}$ but that $4 \not\equiv 1 \pmod{5}$.
   (b) Prove that congruence modulo n is an equivalence relation on $\mathbb{Z}$.
   What are its equivalence classes, for $n = 1$, $n = 2$, and $n = 5$?
   In general, what are this relation's equivalence classes?

6. Again fix a positive integer n. Prove that addition, subtraction, and multiplication respect equivalence classes modulo n (i.e., equivalent operands give equivalent results) as below, giving us ***modular arithmetic***).
   Specifically, supposing that $a, a', b, b' \in \mathbb{Z}$ satisfy $a \equiv a'$ and $b \equiv b' \pmod{n}$, prove that:
   (a) $a + b \equiv a' + b' \pmod{n}$;
   (b) $a - b \equiv a' - b' \pmod{n}$; and
   (c) $a \cdot b \equiv a' \cdot b' \pmod{n}$;

   These properties allow us to think of performing arithmetic operations on the set of *equivalence classes* $\left\{\overline{0}, \overline{1}, \ldots, \overline{n-1}\right\}$ mod n—e.g., mod 6, we have $\overline{2} \cdot \overline{3} = \overline{6} = \overline{0}$.

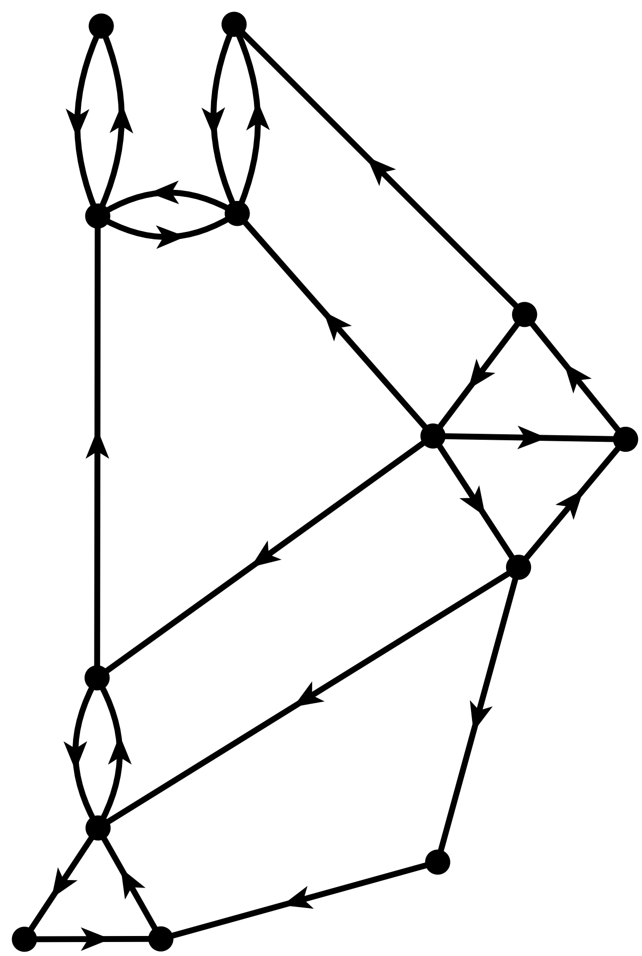
Some relations on graphs

7. Suppose that G is a simple [undirected] graph. Consider the relations given below on the vertices of G.
   (a) $v_0 \ A \ v_1$ if there is an edge joining the vertices v_0 and v_1. (What word might the "A" stand for?)
   (b) $v_0 \ C \ v_1$ if there is a path joining the vertices v_0 and v_1. (What word might the "C" stand for?)

   Which of these two relations is an equivalence relation? What are its equivalence classes? How does the other relation relate to it?

8. Suppose that G is a digraph, and define a relation on the vertices of G by $v_0 \ SC \ v_1$ just when there are *directed* paths both from v_0 to v_1 *and* from v_1 to v_0 (such vertices are called ***strongly connected***).
   (a) Show that SC is an equivalence relation on the vertices of G (the figure below gives you a digraph to think about).
   (b) Circle the equivalence classes for SC for the digraph below.
   (c) If we make a new graph $\overline{G}$ with a vertex for each *equivalence class* of SC and add all edges of G that connect vertices from different equivalence classes, explain why this new directed graph must be a DAG (directed acyclic graph).

   [Hint: what would a cycle on this new graph mean for G and the equivalence classes?]



Relations and Functions: definitions

1. If A and B are *two* sets, what is a ***relation*** R *from* A *to* B, and how does this differ from the single-set case? What are the ***domain*** and ***codomain*** of such a relation R?

2. What is the defining property of a ***function*** $f : A \to B$ (best written as a relation $\overset{f}{\mapsto}$)?
   How should we conceptualize what a function *does*?
   If $a \in A$, what do we mean by the expression $f(a)$ (what is less than ideal about the phrase "the function $f(x)$")?

3. What does it mean for two functions $f, f' : A \to B$ to be equal?

4. If $f : A \to B$ is a function, discuss each of the following terms, both *intuititively* and *symbolically*:
   (a) f is ***injective*** (or ***one-to-one***)—how does this relate to the definition of f being a function?
   (b) f is ***surjective*** (or ***onto***)—how does this relate to the ***range*** of f?
   (c) f is ***bijective*** (or ***one-to-one & onto***)—what is special about bijective functions?

5. Given functions $f : A \to B$ and $g : B \to C$, define their ***composition*** $g \circ f$—what are this function's domain and codomain, and how is the function defined?
   [Be very careful to note that the functions of a composition are read right-to-left!]

6. For any set X, how do we define the ***identity function***? Which of the above function properties does it possess? What happens when an identity function appears in a composition?

7. What does it mean for two functions $f : A \to B$ and $g : B \to A$ to be ***inverses*** of one another (formally, intutively, and symbolically)? How does this relate to identity functions?

. . . and key results and applications

8. Supposing $f : A \to B$ and $g : B \to C$, prove each of the following propositions (the proofs will easily click into place if you're careful with your new definitions, use of variables, and basic proof techniques!):
   (a) If f and g are both injective, then $g \circ f$ is injective.
   (b) If $g \circ f$ is injective, then f is injective.
   (c) If f and g are both surjective, then $g \circ f$ is surjective.
   (d) If $g \circ f$ is surjective, then g is surjective.

9. Using the function formulation of inverses, what do 8(b) and 8(d) tell us if $f : A \to B$ and $g : B \to A$ are inverses?

10. While it might seem "obvious*", each bijective function $f : A \to B$ has just one inverse.
    Use the definition of inverse functions to show that if $g, g' : B \to A$ are both inverses for f, then $g = g'$.
    [Hint: Consider the composition $g \circ f \circ g'$ grouped with parentheses in two different ways.]
    How do we denote this unique inverse of a bijective function f?

11. Suppose that $f : A \to B$ and $g : B \to C$ are both invertible. How do we know that $g \circ f$ is invertible, and what is the formula for its inverse (be careful with order!)?

12. Consider the function mapping graphs to $\mathbb{Z}$ via $G \overset{v}{\mapsto}$ number of vertices in G.
    (a) Explain why it is that if G and H are isomorphic graphs, then $v(G) = v(H)$.
    (b) What does this tell us if two graphs have *different* numbers of vertices?
    (c) What does this tell us if two graphs have the *same* number of vertices?
    (d) In the above contexts, we are using f as an ***invariant*** of isomorphism-classes of graphs (isomorphism is an equivalence relation!). Try to construct some other invariants of isomorphism classes of graphs.

Counting: set preliminaries

1. Set preliminaries:
   (a) What does it mean for a set A to be ***finite***?
   If A is finite, what do we mean by the ***cardinality*** of A, written $|A|$?
   (b) If A and B are sets, how do we define the ***Cartesian product*** $A \times B$ of A and B?
   (c) If $k \geq 1$, what do we mean by A^k?

2. Suppose that A and B are finite sets. Compute each of the following in terms of $|A|$, $|B|$, and $|A \cap B|$:
   (a) $|A \cup B|$ (b) $|A \smallsetminus B|$ (c) $|A \times B|$

3. Extend your answer for problem 2(a) to the case of the union of *three* sets A, B, and C, and briefly explain.

Counting choices: basic principles

4. What "choice words" are common signals for each of the set operations in problem 2?

5. Briefly explain the difference in meaning between choosing multiple elements from a set A *with replacement* and *without replacement*. If we choose k elements, in order, from a set of size n, how many choices do we have in each case?

6. How does the number of choices for k objects change when order matters versus when order doesn't matter?
   What do we need to be careful about when using this principle?
   What keywords might indicate whether or not we care about orderings?

7. If $0 \leq k \leq n$, define each of the following, and briefly describe what they count:
   (a) the ***factorial*** $n!$ of n;
   (b) the ***permutation function*** $P(n, k)$;
   (c) the ***combination function*** $C(n, k)$ [also known as the ***binomial coefficient*** $\binom{n}{k}$];
   (d) the ***Stirling number*** $S(n, k)$ [assuming $n, k \geq 1$].

8. Suppose that $n \geq 0$ and $k_1, k_2, \ldots, k_m$ are positive integers whose sum is k.
   (a) Define the ***multinomial*** $\binom{n}{k_1\ k_2\ \cdots\ k_m}$.
   (b) Briefly explain what this counts, and what the given conditions on $k_1, k_2, \ldots, k_n$ have to do with that.
   (c) Briefly explain how we can derive the formula for $\binom{n}{k_1\ k_2\ \cdots\ k_m}$ one "k" at a time.

9. Briefly explain the ***pigeonhole principle*** and the related notions of *excess* and *shortfall*.

10. Briefly describe the method of ***stars and bars*** for splitting a list of n entries into k segments (both with and without the condition of the segments being nonempty).

Basic probability: in principle

1. Set preliminaries:
   (a) If S is a set, what do we mean by its ***power set*** $\mathscr{P}(S)$?
   (b) What, then, is the simple meaning of the statement that $A \in \mathscr{P}(S)$?
   (c) If S is finite, what is $|\mathscr{P}(S)|$?

2. When we use probability to study a system, briefly explain the difference between the ***sample space*** and the ***event space*** (illustrate with the examples of a coin-flip or rolling of a die).
   What is an ***event*** simply another name for?

3. Suppose that Ω is a finite sample space, with $|\Omega| = n \geq 1$.
   (a) What is a ***probability distribution*** on Ω, and how does this allow us to compute probabilities of *events*?
   (b) What is a ***uniform probability distribution*** on Ω, and what does this mean about:
      (i) The probability attached to each *sample* in $x \in \Omega$?
      (ii) The probability attached to each *event* in $A \subset \Omega$?

4. How can we approach any basic probability question in the case of a uniform probability distribution?

. . . and in practice

5. Consider the sample space of rolls of a pair of fair 6-sided dice (for simplicity, consider the dice to be distinct).
   How many possible pairs (m, n) of rolls are there? What is the probability of each one?
   Make a chart of all of the possible rolls, as an aid in counting for the parts below.
   Determine the probabilities of each of the following events:
   (a) Both dice show the same number.
   (b) The first die shows a strictly greater number than the second.
   (c) S_n, in which the *sum* of the numbers the dice show is n (compute these for $n = 1, 2, \ldots, 12$).

6. Consider the sample space of 10 consecutive flips of a fair coin (i.e., all sequences of `H` & `T` of length 10).
   What is the size of this sample space? What is the size of the corresponding event space?
   Determine the probabilities of each of the following events:
   (a) The first flip is `H`.
   (b) The first and second flips have identical results.
   (c) The first and second flips have different results.
   (d) All flips are `T`.
   (e) The first 9 flips are all `T`, and the tenth is `H`.
   (f) The flips are `HTHTHTHTHT`.
   (g) H_k, in which there are *exactly* k `H`'s in the list of rolls (compute these for $k = 0, 1, \ldots, 10$).
   (h) There are exactly 3 `H`'s in the first five rolls, then exactly 2 `H`'s in the next five rolls.
   How does this compare with the value of H_5 in the previous part?

Probability spaces, independence, and conditional probability: in principle

1. What are the two ingredients of a ***finite probability space***?

2. Joint probability distributions: suppose that $(X, \mathbb{P}_X)$ and $(Y, \mathbb{P}_Y)$ are finite probability spaces.
   How do we define a probability distribution on the product $X \times Y$?
   How do we interpret the meaning of this new probability space? Illustrate with an example.

3. Subspaces: suppose that B is an event in some probability space $(\Omega, \mathbb{P})$, and that $\mathbb{P}(B) > 0$.
   (a) How do the events of Ω give us a set of events for B?
   (b) How does the probability distribution $\mathbb{P}$ give us a probability distribution $\mathbb{P}_B$ on B?
   What must we be careful about when doing this, and why is it so important that B had $\mathbb{P}(B) > 0$?

4. Suppose that A and B are events in some probability space $(\Omega, \mathbb{P})$.
   (a) What does it mean for A and B to be ***mutually exclusive***? Illustrate with an example.
   (b) What does it mean for A and B to be ***independent***? Illustrate with an example.

5. Suppose that A and B are events in some probability space, and that $\mathbb{P}(B) > 0$.
   (a) How do we define the ***conditional probability*** $\mathbb{P}(A \mid B)$?
   (b) In words, what does this tell us?
   (c) How does this relate to our notion of *subspaces*?
   (d) What can be determined about this conditional probability if A and B are *independent*?

. . . and in practice

6. Let $(\Omega, \mathbb{P})$ be the probability space for two fair six-sided dice being rolled.
   (a) How can this probability space be considered as a product?
   (b) Compute the probabilities of each of the following, and determine which pairs of the events are independent:
   (i) The first roll is even.
   (ii) The first roll is a 1, 2, or 3.
   (iii) The second roll is a multiple of 3.
   (iv) The second roll is a 1, 2, or 3.
   (c) Let B be the event that the sum of the dice is between 3 and 6, inclusive.
   (i) Find the probability distribution for the the subspace given by B;
   what adjective describes this distribution, and why?
   (ii) Determine $\mathbb{P}(A_k \mid B)$ for each $k = 1, 2, 3, 4, 5, 6$, where A_k is the event that the first roll is k.
   Are any of these event A_k mututally exclusive of B? Are any independent of B?
   (iii) Determine $\mathbb{P}(A \mid B)$ and $\mathbb{P}(B \mid A)$, if A is the event that both rolls are equal.

7. Let $(\Omega, \mathbb{P})$ be the probability space for ten flips of a fair coin,
   and let B be the event that exactly 5 of the flips are `H`.
   (a) Compute $P(A \mid B)$ and $P(B \mid A)$, if A is the event that the first five flips are all `H`.
   (b) Consider the subspace of $(\Omega, \mathbb{P})$ given by B,
   Compute the probabilities of the three events below in this subspace:
   (i) The first flip is `T`. (ii) The second flip is `T`. (iii) The first two flips are both `T`.
   Are the events (i) and (ii) independent events in this subspace? Discuss.
   (c) Again consider the subspace of $(\Omega, \mathbb{P})$ given by B.
   Compute the probabilities of the two events below and that of their intersection.
   (i) The first three flips are `H`. (ii) The last three flips are `H`.
   What describes the relationship between these two events in this subspace?

Random variables and expectation: in principle

1. Random variables:
   (a) What is a ***random variable*** on a sample space Ω?
   (b) If $A \subset \Omega$, what is the ***indicator variable*** $\mathbb{1}_A$? How are indicator variables useful?
   (c) If X and Y are random variables on Ω and $c \in \mathbb{R}$, define each of the following:
      (i) cX (ii) $X + Y$ (iii) $X \cdot Y$ (iv) $\min(X, Y)$ (v) $\max(X, Y)$
   (d) What do we mean by the *probability distribution* of a random variable?
2. If X is a random variable on a probability space $(\Omega, \mathbb{P})$, how do we define its ***expected value*** $E(X)$? What properties does expected value have?

. . . and in practice

3. If two fair six-sided dice are rolled, let X be the value of the first roll and Y the value of the second.
   (a) For each combination of the random variables X and Y below, find the distribution for its values and compute its expected value:
      (i) $X + Y$ (ii) $X - Y$ (iii) $\min(X, Y)$ (iv) $\max(X, Y)$
   (b) Determine the distribution for the sum of the random variables in subparts (i) and (ii) above, and briefly explain.
   (c) Determine the distribution for the sum of the random variables in subparts (iii) and (iv) above, and briefly explain.
4. Consider a pair of fair six-sided dice having the following numbers on their faces: Die X: $1, 2, 2, 3, 3, 4$
   Die Y: $1, 3, 4, 5, 6, 8$
   For the event space of rolling both dice (as in the above problem), find the probability distributions for the values X, Y, and $X + Y$; compare with your answer to 3(a)(i).
5. Suppose that a fair coin is flipped 9 times.
   (a) What is the probability of event A_{123}, that the first three flips all have the same result?
   (b) What is the probability of event A_{234}, that the second, third, and fourth flips all have the same result?
   (c) What is the probability $\mathbb{P}(A_{123} \mid A_{234})$? Are these two events independent?
   (d) Represent each of the events $A_{123}, A_{234}, \ldots, A_{789}$ by an indicator variable; what is the expected value of each of these?
   (e) What is the *expected* count of subsequences of three consecutive flips being identical?*
6. Suppose that a permutation of the list $(1, 2, 3, \ldots, n)$ is chosen at random (with each of the $n!$ permutations having equal probability).
   (a) What is the probability that the permutation's first number listed is 1?
   What is the probability that the permutation's second number listed is 2?
   Are these two events independent?
   (b) In general, what is the probability that the permutation's k^{th} number is k?
   Represent each of the above by an indicator variable; what is the expected value of each of these functions?
   (c) What is the *expected* count of numbers k that end up in the k^{th} spot?*

* Hint: Add up the indicator variables, and remember that expectation is additive!

Complexity: in principle

In all problems below, functions will be from $\mathbb{N}$ to $[0, \infty)$.

1. ***Big-O notation***: if f and g are functions, consider the statement that $f(n) = O(g(n))$.
   (a) What is the logical *definition* of this statement?
      (i) Briefly explain the purpose of each quantifier and predicate in the definition.
      (ii) How can we interpret the meaning of this concept graphically (in the sense of classical graphs of functions, not digraphs)?
   (b) What is the *point* of Big-O notation?
   (c) Semantic flaws:
      (i) In what way is the "=" in this expression misleading?
      (ii) What is a little shady about the n's in this expression?
   (d) How does Big-O notation work with respect to addition and subtraction?
   (e) How does Big-O notation work with respect to multiplication, and how is this significantly different than the previous case?

2. Discuss the meanings of the two related notations below; how does each relate to Big-O notation?
   (a) ***Big-*Ω notation $f(n) = \Omega(g(n))$
   (b) ***Big-*Θ notation $f(n) = \Theta(g(n))$

3. What is the hierarchy of sizes for basic expressions in n when n is large, and how does it guide our usage of the above asymptotics?

. . . and in practice

4. Asymptotically simplify the following expressions using Big-O notation.
   (a) $5n^3 + 3n + 1000$ (b) $6\log(n^3) + 2n + 5$ (c) $8n^{100} + 2^n + \log n$ (d) $100^n + 4n^{50} + n!$
   (e) $(100n + n^4)(n^2 + 2^n)$ (f) $(\log\log n + 10000)(\log n + n + \sqrt{n})$
   (g) $(n! + n^n + 1000^n)(10 + n^3 + 300n^2)$

5. Prove the following statements (you may use trial and error and/or a calculator to help with the $\exists$'s!):
   (a) $1000 = O(\log n)$
   (b) $100n = O(n^2)$
   (c) $3^n = \Omega(100 \cdot 2^n)$
   (d) $100n^2 = \Theta(n^2)$

[Problem set continues on back side.]

6. Suppose that you have an array of $n = 1023$ numbers, $a_0, a_1, a_2, \dots, a_{1022}$,
   and that you want to search for a given number A in that array.
   Suppose, for the sake of this problem, that A does *not* match any element of the array.
   (a) If you perform a *linear search*, in which you compare A to each element of your array, how many comparisons will be performed during the search?
   (b) Suppose now that the elements of your array are sorted, and you do a *binary search*, comparing A to the middle number remaining at each step. How many comparisons will be performed? How does this relate to the number 1023?
      (i) What are the specific elements in the array you'll compare A to if A ends up being larger than all of them?
      (ii) What are the specific elements in the array you'll compare A to if A ends up being smaller than all of them?
   (c) Now suppose that your array has $n = 2^k - 1$ elements.
   How many comparisons will each of these algorithms use?
   What are the Big-O bounds for the complexities of each of these algorithms?

7. Suppose that you have numbers arranged on a 5×5 grid. Starting from the top-left cell, you can choose to go down or right at each step (staying on the grid) until you end at the bottom-right, and you'd like to find the path for which the sum of the numbers you pass through is as large as possible.
   (a) The brute-force search:
      (i) How many such paths are there?
      (You will take $4 + 4 = 8$ steps; you just need to *choose* which four of these are "down.")
      (ii) Along each path, how many times will you have to add two numbers?
      (iii) How many additions will you perform in total?
   (b) Dynamic programming:
      (i) Suppose that instead of the brute-force search, you work down and right through the chart, starting at the top-left and, at each stage, updating just each next cell reachable with the sum of it and the larger of the numbers above and/or left of it (sort of like Dijkstra!).
      (Choose a 5×5 grid and actually *do* this to see how it works!)
      (ii) Count the total number of additions necessary in this algorithm.
   (c) Which of the two algorithms above is more efficient?
   (What makes the more efficient algorithm possible?)
   (d) For each of these algorithms, what would the total number of additions be for a general $n \times n$ grid?
   (e) Call a solution *infeasible* if the number of additions is more than one billion.
   At what specific n does each of the above algorithms become infeasible?
   (Feel free to use a calculator or online tool to compute the values!)

Recurrence relations: in principle

1. What, in general, is a ***recurrence relation*** for a function or sequence?

2. Linear recurrence relations:

   (a) What does it mean for a recurrence relation to be ***linear***?
   Give some examples of linear and non-linear recurrence relations.

   (b) What is the ***characteristic polynomial*** for a linear recurrence relation?

   (c) Given roughly what happens to the value of the terms in a linear recursion A_n when we increment n, why might we expect exponentials to be involved in the *closed* formulæ for their terms?

   (d) Briefly outline the procedure for finding a closed formula for the terms of a linear recurrence relation using its characteristic polynomial:

      (i) What do the roots of the characteristic polynomial tell us?
      What do we do when the characteristic polynomial has a repeated root?

      (ii) What role is played by *initial values*?

. . . and in practice

3. Find a closed formula for the terms of the sequence given by $A_0 = 1$ and, for $n \geq 0$, $A_{n+1} = (n+1)A_n$.
[Note that this recurrence relation is *not* linear!]

4. Find closed formulæ for the terms of each of the following recursively-defined sequences:

   (a) $B_0 = 5$ and for $n \geq 0$, $B_{n+1} = B_n$.

   (b) $C_0 = 3$ and for $n \geq 0$, $C_{n+1} = 10C_n$.

   (c) $D_0 = 0$, $D_1 = 1$, and for $n \geq 0$, $D_{n+2} = 3D_{n+1} - 2D_n$.

   (d) $E_0 = 1$, $E_1 = 8$, and for $n \geq 0$, $E_{n+2} = 4E_{n+1} - 4E_n$.

   (*e) $F_0 = 0$, $F_1 = 1$, and for $n \geq 0$, $F_{n+2} = F_{n+1} + F_n$.

Please do these for next class!

5. Consider the recurrence relation $T(n) = T\left(\frac{n}{2}\right) + 1$, and suppose that $T(1) = 1$.

   (a) Find $T(1024)$ and, in general, $T(2^k)$.

   (b) If you change variables via $n = 2^k$ (so $k = \log n$), what is the resulting formula for $T(n)$?

6. Consider the recurrence relation $T(n) = 2T\left(\frac{n}{2}\right) + n$, and suppose $T(1) = 0$.

   (a) Show all steps in computing $T(64)$; along the way, you'll have found $T(2^k)$ for $k = 2, 3, \ldots, 6$.

   (b) What is the general formula for $T(2^k)$?
   If you change variables via $n = 2^k$ (so $k = \log n$), what is the resulting formula for $T(n)$?

7. Carefully do the same as in the previous problem, but now for $T(n) = 2T\left(\frac{n}{2}\right) + n^2$.

Set equivalence and infinite sets

1. Define the relation $\sim$ of set equivalence—what properties does it have, and why? If A and B are two sets with $A \sim B$, what phrase do we use to describe this situation, and what other symbols can we use to denote this?

2. Define and briefly explain the concepts of ***finite set*** and ***infinite set***.

3. What set is our model for the ***countably infinite*** sets, and how can we conceptualize what it means to be countably infinite? Give some examples of countably infinite sets among the sets of numbers we've discussed.

4. What do we mean by a ***countable*** set, and thus what is meant by an *uncountable* set?
   Draw a Venn diagram including all of the types of sets listed above.

Cardinality and set operations/relations

5. The operations of $\cup$ and $\times$ can be used to combine nonempty sets into larger ones.
   (a) What can be said about the union or Cartesian product of two *finite* sets?
   (b) What can be said about the union or Cartesian product of two *countably infinite* sets?
   (c) What can be said about the union or Cartesian product of two *uncountable* sets?
   What about unions and products of mixed combinations of nonempty sets? What about *empty* sets?

6. Suppose that A and B are two sets, and that $A \subset B$.
   (a) What can be concluded if A is: (i) finite; (ii) countably infinite; (iii) uncountable?
   (b) What can be concluded if B is: (i) finite; (ii) countably infinite; (iii) uncountable?
   What does this tell us about the cardinality of $C \cap D$ for two sets C and D, and why must we be careful?

7. We proved that for any set S, we have $S \not\sim \mathscr{P}(S)$.
   What does this tell us about $|\mathscr{P}(S)|$ if S is: (i) finite; (ii) countably infinite; (iii) uncountable?
   Use this to give an example of an uncountable set.

Cardinality problems

8. Suppose that U is uncountable and that C is countable. Use contradiction to prove that $U \setminus C$ is still uncountable.

9. Explain why the union of a countably infinite collection of finite sets is countable.
   What about the union of a countable collection of countable sets, or a finite union of finite sets?

10. Suppose that S is a set. Find a bijection between $\mathscr{P}(S)$ and the set $2^S = \{f : S \to \{0, 1\}\}$,
    i.e., the set of functions from S into the set $\{0, 1\}$. [Hint: think of the set $\{0, 1\}$ as $\{\text{false}, \text{true}\}$.]

11. Determine the cardinality (finite of some size, countably infinite, or uncountable) of each of the following:
    (a) The set of all sequences of 0's and 1's of length n. [Hint: count choices.]
    (b) The set of *all* finite sequences of 0's and 1's. [Hint: use the previous part and problem 9.]
    (c) The set of all *infinite* sequences of 0's and 1's. [Hint: consider functions from $\mathbb{N} \to \{0, 1\}$.]

12. Prove that the set S of real numbers between 0 and 1 written $0.d_1d_2d_3d_4\cdots$ using only the decimal digits 0 and 1 is uncountable in two ways:
    (a) by using problem 11(c); and
    (b) by contradiction, using ***Cantor's diagonal argument***.
    What does this tell us about the set $\mathbb{R}$?

Formally encoding relations and functions via sets

1. We can formally encode a ***relation*** R on a set X as a *subset* $R \subset X \times X$.

   Use this to translate each of the following assertions into set language:

   (a) $a \, R \, b$

   (b) R is ***reflexive***

   (c) R is ***symmetric***

   (d) R is ***transitive***

   (e) The relation S on X ***subsumes*** R

2. In a similar manner, we can formally encode a ***relation*** from A to B as a subset of $A \times B$.

   Use this to translate each of the following assertions into set language.

   (a) $f : A \to B$ (considered as $f \subset A \times B$) is a ***function***

   (b) $f : A \to B$ is ***injective***

   (c) $f : A \to B$ is ***surjective***

   (d) The ***range*** of $f : A \to B$

Formally encoding graphs via sets

We can formally encode various sort of ***graphs*** in a number of ways—in some way or another, these formal encodings must tell us the ***vertices*** and ***edges*** of our graph, as well as any extra structures put on the graph. One way is via an ordered pair (V, E).

3. Suppose that V is a finite set, and that E is a set consisting of sets $\{v, w\}$ with $v, w \in V$ and $v \neq w$. What type of graph does this encode?

4. Again suppose that V is a finite set, and that E is a set consisting of *ordered pairs* (v, w) of elements of V. What type of graph does this encode?

5. In either of the above cases, we could also take E to be a ***multiset***, which may contain a given element multiple times—how does this affect the type of graph we've encoded?

6. In each of the above cases, an ***isomorphism*** ϕ between graphs $G = (V, E)$ and $G' = (V', E')$ consists of a pair of *bijections* $\phi_v : V \to V'$ and $\phi_e : E \to E'$.

   (a) To be a graph isomorphism, the bijection between the vertex sets must relate to that of the edge sets (we can't pair up the edges in any way we like, because they connect vertices!).

   Determine what extra conditions we need to get a graph isomorphism in each of the cases above.

   (b) Once we use functions, the fact that graph isomorphism is an equivalence relation boils down to function compositions and what we know about bijective functions.

   What exactly needs to be shown to prove this?

Sets, sets, sets

7. Technically, we can define the ordered pair (x, y) as the set $\big\{\{x\}, \{x, y\}\big\}$.

   Supposing that $a \neq b$, what sets do (a, b), (b, a), (a, a), and (b, b) correspond to?

   Explain how this set encoding allows us to distinguish all *ordered* pairs from one another.

*8. It appears that *every* object we've discussed since basic logic and sets could be formally encoded using sets.

   While it's important to point this out, we more or less *never* go that far—why not? Why do we nonetheless need to know that these formalizations exist?

String algebra: in principle

Suppose that Σ is a finite "***alphabet***" (set) of ***characters*** or ***symbols***.

1. Define what we mean by a ***string*** (or ***word***) over the alphabet Σ.
   (a) Compare the abstract concept of a string with those used in programs.
   (b) What do we mean by the string ε?
   (c) Given a string x, what does $|x|$ represent? What is $|\varepsilon|$?

2. If x and y are strings over Σ, define their ***product*** xy, i.e., $x \cdot y$.
   (a) What other term describes this operation?
   (b) What properties does this operation have, and—crucially—what key property does it *not* have?
   (c) How does the product let us define nonnegative whole ***powers*** z^n of a string z?
   (d) If x and y are strings, how can we simplify $|xy|$ and $|x^n|$?

3. If W is a set of strings, define its ***asterate*** W^*.
   (a) What properties does this operation have, and what string does this set *always* include?
   (b) If x is a single character or string over Σ, we can define $x^* = \{x\}^*$.
   What strings does the set x^* contain?
   (c) In words, what is Σ^*, and what does the statement "$x \in \Sigma^*$" mean?

4. Blurring the line between strings and sets of strings a bit, if x and y are strings, what is meant by their formal ***sum*** $x + y$, and what set operation does this correpond to?
   What properties does this operation have, and how does it relate to our other operation "$\cdot$"?

5. What is a ***regular expression*** over Σ?

. . . and in practice

In the problems below, characters in Σ will be written in `typewriter font`, with variables representing strings written as usual via *italic letters*.

6. Constrast, in writing, the meanings of the following,
   including the context of whether each is an *element* of Σ^* or a *subset* of Σ^*: (a) ε (b) $\varnothing$ (c) $\{\varepsilon\}$

7. Suppose that $\Sigma_1 = \{\mathtt{a}\}$ and $\Sigma_2 = \{\mathtt{a}, \mathtt{b}\}$. How do the cardinalities of Σ_1^* and Σ_2^* compare? What is $\varnothing^*$?

8. For any set W of strings, describe in writing what $W^* \setminus \{\varepsilon\}$ means.
   [This comes up enough that we often denote it by W^+.]

9. Simplify the following string expressions:
   (a) $\mathtt{a}^3\varepsilon^5(\mathtt{bca})^2$ (b) $\mathtt{a}(\varepsilon + \mathtt{b} + \mathtt{c})w$ (c) $(\varepsilon + \mathtt{a})(\varepsilon + \mathtt{b})$ (d) $(\mathtt{a} + xy)^3$

10. Describe the sets of strings generated by the following regular expressions over $\Sigma = \{0, 1\}$, both in writing and using set-builder notation:
    (a) $\mathtt{1}^*$ (b) $\mathtt{01}^*$ (c) $\mathtt{0}^*\mathtt{1}$ (d) $(\mathtt{01})^*$ (e) $\mathtt{0}^*\mathtt{1}^*$

11. Find regular expressions expressing the following sets of strings over $\Sigma = \{0, 1\}$:
    (a) all strings starting with `101`;
    (b) all strings containing the substring `000`;
    (c) all strings that can be built as a products of some number of copies of the string `001`; and
    (d) all strings in which a `0` is not preceded by a `1`.

Finite Automata: in principle

Suppose that Σ is a finite alphabet.

1. Briefly describe how a ***discrete finite automaton*** (DFA) over Σ acts, in terms of an edge-labeled digraph:
   (a) What in the graph corresponds to the finite set of ***states*** of the DFA?
   (b) What in the graph corresponds to the ***transitions*** of the DFA, and what rules apply to those transitions?
   (c) How do we mark the unique ***start state*** of the graph?
   (d) What special designation can each state have (or not have)?
   (e) Given a *string* of characters from Σ, describe how the DFA operates one character at a time, and what it means for the DFA to ***accept*** or ***reject*** the string.

2. Taking a step back, as an input/output machine, what does a DFA over Σ take as input? What are its possible outputs? Viewing this as a *function*, what are its domain and codomain?

3. Briefly explain why each string $x \in \Sigma^*$ must be either accepted or rejected (but not both!) by a DFA.
   What do we mean by the ***language accepted by*** a DFA?

4. Taking a further step back, how can we think of this whole process as giving us a function from the set of all DFA's to $\mathscr{P}(\Sigma^*)$?

...and in practice

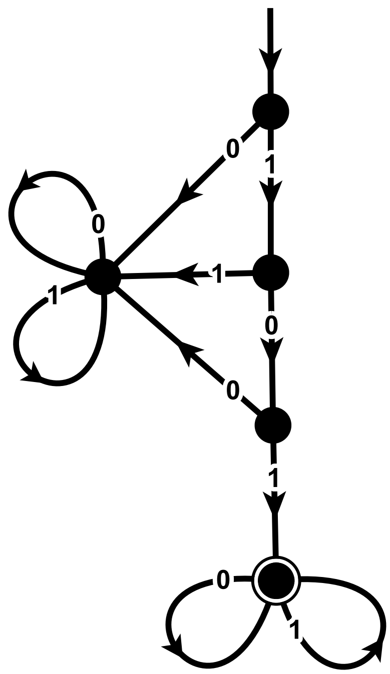
5. Consider the DFA labeled A on the back of this page.
   (a) Will this DFA accept or reject: (i) a string x that starts with `101`;
   (ii) a string y that starts with `100`;
   (iii) a string z that starts with `11`;
   (iv) a string w that starts with `0`?
   (b) Find a regular expression that produces the language accepted by A.

6. Consider the DFA labeled B on the back of this page.
   (a) Will this DFA accept or reject: (i) the string ε;
   (ii) the string `1000100`;
   (iii) the string `00100100`;
   (iv) the string `11100011`?
   (b) Find a regular expression that produces the language accepted by B.

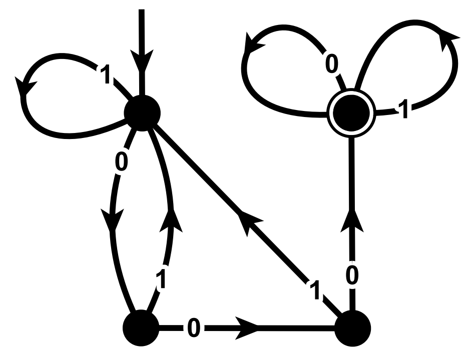
7. Consider the DFA labeled C on the back of this page.
   (a) Will this DFA accept or reject: (i) the string ε;
   (ii) the string `001`;
   (iii) the string `001001`;
   (iv) the string `00100100`?
   (b) Find a regular expression that produces the language accepted by C.

8. Consider the DFA labeled D on the back of this page.
   (a) Will this DFA accept or reject: (i) the string ε;
   (ii) the string `1111`;
   (iii) the string `00011`;
   (iv) the string `001110`?
   (b) Find a regular expression that produces the language accepted by D.

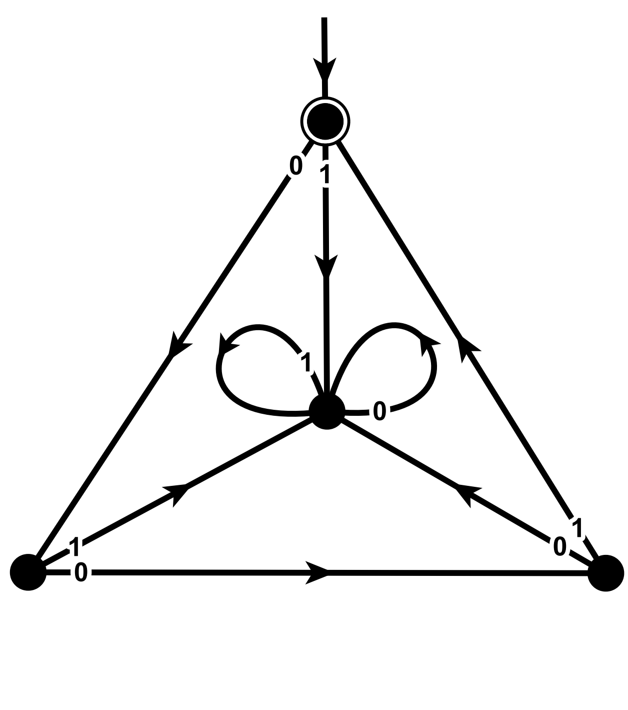
DFA "A"



DFA "B"



DFA "C"



DFA "D"