CSE 20 Spring 2024
Practice for Test 2 Solutions

1. Predicates and Quantifiers Express each of the following statements symbolically using quantifiers, variables, propositional connectives, and predicates (make sure to define the domain and the meaning of any predicates you introduce). Then, express its negation as a logically equivalent compound proposition without a $\neg$ in front. Decide whether the statement or its negation is true, and prove it.

(a) If m is an integer, $m^2 + m + 1$ is odd.

Define the predicate $D(x)$ to mean "x is odd" with domain $\mathbb{Z}$.
The sentence is then $\forall m \in \mathbb{Z}(D(m^2 + m + 1))$.
Its negation is $\neg\forall m \in \mathbb{Z}(D(m^2 + m + 1)) \equiv \exists m(\neg D(m^2 + m + 1))$.
The original statement is true.

Proof Towards universal generalization, let m be an arbitrary integer. By definition of odd, **to show**: there is an integer k such that $m^2 + m + 1 = 2k + 1$. Factoring the LHS, $m^2 + m + 1 = m(m + 1) + 1$. We proceed in proof by cases, since m is either even or odd.

- Case 1 **to show**: (m is even) $\rightarrow D(m^2 + m + 1)$. In a direct proof of the conditional, assume m is even and choose c be an integer such that $m = 2c$. **To show**: there is an integer such that $m^2 + m + 1$ is twice that integer plus 1. To find a witness for this existential, we substitute $m = 2c$ and calculate, $m(m + 1) = 2c(2c + 1) = 2(2c^2 + c)$. Since $2c^2+c \in \mathbb{Z}$ (by properties of integer addition and multiplication), it is the witness we need, because $2(2c^2 + c) + 1 = 2c(2c + 1) + 1 = m(m + 1) + 1 = m^2 + m + 1$.
- Case 2 **to show**: (m is odd) $\rightarrow D(m^2 + m + 1)$. In a direct proof of the conditional, assume m is odd and choose c be an integer such that $m = 2c + 1$. **To show**: there is an integer such that $m^2 + m + 1$ is twice that integer plus 1. To find a witness for this existential, we substitute $m = 2c+1$ and calculate, $m(m+1) = (2c+1)(2c+2) = 2(2c+1)(c+1)$. Since $(2c+1)(c+1) \in \mathbb{Z}$ (by properties of integer addition and multiplication), it is the witness we need, because $2(2c + 1)(c + 1) + 1 = m(m + 1) + 1 = m^2 + m + 1$.

Alternate proof Let m be an integer. WTS that $m^2 + m + 1 \mod 2 = 1$. Case 1: $m \mod 2 = 0$. Then by modular arithmetic (Corollary 2 on page 242),

$$\begin{aligned}(m^2 + m + 1) \mod 2 &= ((m \mod 2)(m \mod 2) + (m \mod 2) + 1) \mod 2\\ &= (0 \cdot 0 + 0 + 1) \mod 2 = 1\end{aligned}$$

as required. Case 2: $m \mod 2 = 1$. Then by modular arithmetic (Corollary 2 on page 242),

$$\begin{aligned}(m^2 + m + 1) \mod 2 &= ((m \mod 2)(m \mod 2) + (m \mod 2) + 1) \mod 2\\ &= (1 \cdot 1 + 1 + 1) \mod 2 = 3 \mod 2 = 1\end{aligned}$$

as required.

(b) For all integers n greater than 5, $2^n - 1$ is not prime.

Define the predicate: $P(x)$ to mean "x is prime" with domain: $\mathbb{Z}$.
The sentence translates to $\forall n(n > 5 \to \neg P(2^n - 1))$.
Its negation is $\neg\forall n \in \mathbb{Z}(n > 5 \to \neg P(2^n - 1)) \equiv \exists n \in \mathbb{Z}(n > 5 \wedge P(2^n - 1))$.
The negation is true: to prove it, we need a witness integer greater than 5 where $2^n - 1$ is prime. Consider $n = 7$, an integer so in the domain. Calculating, $2^7 - 1 = 128 - 1 = 127$. This is a prime number because all integers greater than 1 and less than $\sqrt{127}$ (namely $2, 3, 4, 5, 6, 7, 8, 9, 10, 11$) are not factors of 127.

(c) If n and m are even integers, then $n - m$ is even.

Define the predicate $E(x)$ to mean "x is even" with domain $\mathbb{Z}$. Note: this predicate could be expressed symbolically as $\exists k \in \mathbb{Z}(x = 2k)$.
The sentence translates to $\forall x \in \mathbb{Z}\ \forall y \in \mathbb{Z}\ ((E(x) \wedge E(y)) \to E(x - y))$.
Its negation is

$$\neg\forall x \in \mathbb{Z}\ \forall y \in \mathbb{Z}\ ((E(x) \wedge E(y)) \to E(x - y)) \equiv \exists x \in \mathbb{Z}\exists y \in \mathbb{Z}(E(x) \wedge E(y) \wedge \neg E(x - y))$$

The original statement is true: Towards universal generalization let x, y be arbitrary integers and to show is $(E(x) \wedge E(y)) \to E(x - y)$. Assume, towards a direct proof of the conditional that x and y are both even. We WTS that $x - y$ is even. By definition of even, there are integers j, h such that $x = 2j$ and $y = 2h$. Then $x - y = (2j) - (2h) = 2(j - h)$. Choosing the witness $k = j - h$ (an integer), we have proved that $x - y$ is even.

2. Proof strategies Prove each of the following claims. Use only basic definitions and general proof strategies in your proofs; do not use any results proved in class / the textbook as lemmas. You may use basic properties of real numbers that we stated (without proof) in class, e.g. that the difference of two integers is an integer.

(a) The difference of any rational number and any irrational number is irrational.

Proof: Symbolically, to show is:

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} ((x \in \mathbb{Q} \wedge y \notin \mathbb{Q}) \to (x - y \notin \mathbb{Q}))$$

. Towards universal generalization, let x, y be arbitrary real numbers. Proceed in a proof by contradiction with

$$p = (x \in \mathbb{Q} \wedge y \notin \mathbb{Q}) \to (x - y \notin \mathbb{Q}) \qquad r = y \in \mathbb{Q}$$

. To show: $\neg p \to (r \wedge \neg r)$. Assume $\neg p$, that is $x \in \mathbb{Q} \wedge y \notin \mathbb{Q} \wedge x - y \in \mathbb{Q}$. To show: $r \wedge \neg r$. By definition of conjunction, each conjunct is true and $y \notin \mathbb{Q}$ being true means $\neg(y \in \mathbb{Q}$ is true. Thus, $\neg r$ is true, and to show is: r. By definition of $\mathbb{Q}$, there are integers a, b, p, q with $b \neq 0$ and $q \neq 0$ such that $x = \frac{a}{b}, x - y = \frac{p}{q}$. Then

$$y = x - (x - y) = \frac{a}{b} - \frac{p}{q} = \frac{aq - bp}{bq}.$$

By properties of integer multiplication and subtraction, $aq - bp \in \mathbb{Z}$ and $bq \in \mathbb{Z}$, and since b, q are both nonzero $bq \neq 0$. Thus, by definition of $\mathbb{Q}$, $y \in \mathbb{Q}$, namely, r. Thus, we proved $\neg p \to (r \wedge \neg r)$ and so we conclude that p holds.

(b) If a, b, c are integers and $a^2 + b^2 = c^2$, then at least one of a and b is even.

We prove two helper claims (lemmas) before the proof.
Lemma 1: For any integer x, if x^2 is even , then so is x.
Proof of Lemma 1: Let x be an arbitrary integer and assume (towards proof by contrapositive) that x is not even (hence, odd). WTS that x^2 is also not even. By assumption, $x = 2k + 1$ for some integer k. Squaring: $x^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$, and by properties of integer addition and multiplication $2k^2+2k \in \mathbb{Z}$, so x^2 is odd, hence not even.

Lemma 2: For any odd integer x, $x^2 \textbf{ mod } 4 = 1$.
Proof of Lemma 2: Let x be an arbitrary odd integer. By definition, this means there is an integer, call it k, such that $x = 2k + 1$. Squaring: $x^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1$. Since $k^2 + k \in \mathbb{Z}$, $x^2 \textbf{ mod } 4 = 1$.

Proof: Let a, b, c be arbitrary integers. To show: $a^2 + b^2 = c^2 \to (E(a) \vee E(b))$. Towards a contradiction, consider

$$p = (a^2 + b^2 = c^2) \to (E(a) \vee E(b)) \qquad r = (c^2 \textbf{ mod } 4 = 0)$$

To show: $\neg p \to (r \wedge \neg r)$. Assume $\neg p$, that is $a^2 + b^2 = c^2$ and a and b are both not even. By properties of even/odd, a and b are both odd. Thus, by Lemma 2, $a^2 \textbf{ mod } 4 = b^2 \textbf{ mod } 4 = 1$. By modular arithmetic, $c^2 \textbf{ mod } 4 = (a^2 + b^2) \textbf{ mod } 4 = (1 + 1) \textbf{ mod } 4 = 2$. Since $2 \neq 0$, this proves $\neg r$. To show: r. Since $c^2 \textbf{ mod } 4 = 2$, there is an integer, call it q, such that $c^2 = 4q + 2 = 2(2q + 1)$. Thus, c^2 is even. Therefore, applying Lemma 1, c is even. By definition, this gives an integer k such that $c = 2k$. Thus, $c^2 = 4k^2 = 4k^2{+}0$ so $c^2 \textbf{ mod } 4 = 0$, namely r. Thus, we proved $\neg p \to (r \wedge \neg r)$ and so we conclude that p holds.

(c) For any integer n, $n^2 + 5$ is not divisible by 4.

Let n be an arbitrary integer. Towards proof by cases, notice that $(n \textbf{ mod } 4 = 0) \vee (n \textbf{ mod } 4 = 1) \vee (n \textbf{ mod } 4 = 2) \vee (n \textbf{ mod } 4 = 3)$

- **Case 1 to show**: $(n \textbf{ mod } 4 = 0) \to (n^2{+}5$ is not divisible by $4)$. Assume $(n \textbf{ mod } 4 = 0)$. Then, by modular arithmetic,

  $$(n^2+5) \textbf{ mod } 4 = ((n \textbf{ mod } 4)^2 + 5 \textbf{ mod } 4) \textbf{ mod } 4 = (0^2+1) \textbf{ mod } 4 = 1 \textbf{ mod } 4.$$

  Since $(n^2+5) \textbf{ mod } 4 \neq 0$, n^2+5 is not divisible by 4, as required.

- **Case 2 to show**: $(n \textbf{ mod } 4 = 1) \to (n^2{+}5$ is not divisible by $4)$. Assume $(n \textbf{ mod } 4 = 1)$. Then, by modular arithmetic,

  $$(n^2+5) \textbf{ mod } 4 = ((n \textbf{ mod } 4)^2 + 5 \textbf{ mod } 4) \textbf{ mod } 4 = (1^2+1) \textbf{ mod } 4 = 2 \textbf{ mod } 4.$$

  Since $(n^2+5) \textbf{ mod } 4 \neq 0$, n^2+5 is not divisible by 4, as required.

- **Case 3 to show**: $(n \textbf{ mod } 4 = 2) \to (n^2{+}5$ is not divisible by $4)$. Assume $(n \textbf{ mod } 4 = 2)$. Then, by modular arithmetic,

  $$(n^2+5) \textbf{ mod } 4 = ((n \textbf{ mod } 4)^2 + 5 \textbf{ mod } 4) \textbf{ mod } 4 = (2^2+1) \textbf{ mod } 4 = 5 \textbf{ mod } 4 = 1$$

  Since $(n^2+5) \textbf{ mod } 4 \neq 0$, n^2+5 is not divisible by 4, as required.

- **Case 4 to show**: $(n \textbf{ mod } 4 = 3) \to (n^2+5$ is not divisible by $4)$.

  $$(n^2+5) \textbf{ mod } 4 = ((n \textbf{ mod } 4)^2 + 5 \textbf{ mod } 4) \textbf{ mod } 4 = (3^2+1) \textbf{ mod } 4 = 10 \textbf{ mod } 4 = 2$$

  Since $(n^2+5) \textbf{ mod } 4 \neq 0$, n^2+5 is not divisible by 4, as required.

The proof by cases is now complete for the arbitrary integer n.

(d) For all positive integers a, b, c, if $a \nmid bc$ then $a \nmid b$. (The notation $a \nmid b$ means "a does not divide b").

Let a, b, c be arbitrary positive integers. Assume, towards a proof by contrapositive, that $a|b$. To show: $a|bc$. By definition of divisibility, there is an integer k such that

$$b = ak.$$

Multiplying both sides by c,

$$bc = (ak)c = a(kc).$$

By properties of integer multiplication, $kc \in \mathbb{Z}$ and so it witnesses $a|bc$, as required.

Bonus: Watch the video `https://www.youtube.com/watch?v=MhJN9sByRS0` and write down the statement and one or both of the proofs described using the notation, definitions, and proof strategies from CSE 20.

3. Sets Prove or disprove each of the following statements.

(a) For all sets A and B, $\mathcal{P}(A) \cup \mathcal{P}(B) = \mathcal{P}(A \cup B)$.

False. A counterexample would be $A = \{1\}, B = \{2\}$. Then

$$\mathcal{P}(A) \cup \mathcal{P}(B) = \{\emptyset, \{1\}\} \cup \{\emptyset, \{2\}\} = \{\emptyset, \{1\}, \{2\}\}$$

but

$$\mathcal{P}(A \cup B) = \mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}.$$

These sets are not equal because they disagree about membership of $\{1,2\}$.

(b) For all sets A and B, $\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B)$.

True. Proof: let A, B be arbitrary sets. We WTS that $\mathcal{P}(A) \cap \mathcal{P}(B) \subseteq \mathcal{P}(A \cap B)$ and $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A) \cap \mathcal{P}(B)$.

(c) For the first subset inclusion, consider arbitrary $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$. By definition of intersection, $X \in \mathcal{P}(A)$ and $X \in \mathcal{P}(B)$. By definition of power set, this means that $X \subseteq A$ and $X \subseteq B$. WTS that $X \subseteq A \cap B$: let $x \in X$. Since $X \subseteq A$, $x \in A$. Since $X \subseteq B$, $x \in B$. Thus, $x \in A$ and $x \in B$ so $x \in A \cap B$. Since x was arbitrary, $\forall x(x \in X \to x \in A \cap B)$ so $X \subseteq A \cap B$. Therefore, by definition of power set $X \in \mathcal{P}(A \cap B)$, as required.

(d) For the second subset inclusion, consider arbitrary $X \in \mathcal{P}(A \cap B)$. By definition of power set, $X \subseteq (A \cap B)$. WTS that $X \subseteq A$ and $X \subseteq B$. Let $x \in X$. Then by definition of subsets, $x \in A \cap B$. By definition of intersection, $x \in A$ and $x \in B$. Thus $X \subseteq A$ and $X \subseteq B$ and so by definition of power set, $X \in \mathcal{P}(A)$ and $X \in \mathcal{P}(B)$. By definition of intersection, $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$, as required.

(e) For any sets A, B, C, D, if the Cartesian products $A \times B$ and $C \times D$ are disjoint then either A and C are disjoint or B and D are disjoint (or both).

True. Proof: Let A, B, C, D be arbitrary sets. Towards a proof by contrapositive, we assume that $A \cap C \neq \emptyset$ and $B \cap D \neq \emptyset$ and WTS that $(A \times B) \cap (C \times D) \neq \emptyset$. Since $A \cap C \neq \emptyset$, let $x \in A \cap C$. Since $B \cap D \neq \emptyset$, let $y \in B \cap D$. Let's consider (x, y). By definition of intersection, since $x \in A \cap C$ and $y \in B \cap D$, $x \in A$ and $y \in B$. Thus, by definition of Cartesian product, $(x, y) \in A \times B$. Similarly, $(x, y) \in C \times D$. Thus, by definition of intersection, $(x, y) \in (A \times B) \cap (C \times D)$. In particular, this means that $(A \times B) \cap (C \times D) \neq \emptyset$, as required.

(f) There are sets A, B such that $A \in B$ and $A \subseteq B$.

True. Proof: Consider the example $A = \{1, 2\}$ and $B = \{1, 2, \{1, 2\}\}$. Then $A \in B$ because it ($\{1, 2\}$) shows up in the list of elements of B. Moreover, since each of the elements of A (the numbers 1 and 2) are also elements of B (they each show up in B's list of elements), A is a subset of B.

(g) For all sets A, B, C: $A \cap B = \emptyset$ and $B \cap C = \emptyset$ if and only if $(A \cap B) \cap C = \emptyset$.

False. Proof: Consider the counterexample $A = \{1, 2\}$ and $B = \{2, 3\}$ and $C = \{3\}$. Then it is not the case that "$A \cap B = \emptyset$ and $B \cap C = \emptyset$" because $A \cap B = \{2\}$. However, it is the case that $(A \cap B) \cap C = \emptyset$ because

$$(A \cap B) \cap C = (\{1, 2\} \cap \{2, 3\}) \cap \{3\} = \{2\} \cap \{3\} = \emptyset.$$

The biconditional statement is false because one of its arguments is false while the other is true.

4. Induction and Recursion Prove the following statements:

(a) $\forall s \in S\, (\, rnalen(s) = basecount(s, \texttt{A}) + basecount(s, \texttt{C}) + basecount(s, \texttt{U}) + basecount(s, \texttt{G})\,)$ where S RNA strands and the functions *rnalen* and *basecount* are define recursively as:

$$\begin{array}{llll}
 & & rnalen : S & \to \mathbb{Z}^+ \\
\text{Basis Step:} & \text{If } b \in B \text{ then} & rnalen(b) & = 1 \\
\text{Recursive Step:} & \text{If } s \in S \text{ and } b \in B\text{, then} & rnalen(sb) & = 1 + rnalen(s)
\end{array}$$

$$\begin{array}{llll}
 & & basecount : S \times B & \to \mathbb{N} \\
\text{Basis Step:} & \text{If } b_1 \in B, b_2 \in B & basecount(b_1, b_2) & = \begin{cases} 1 & \text{when } b_1 = b_2 \\ 0 & \text{when } b_1 \neq b_2 \end{cases} \\
\text{Recursive Step:} & \text{If } s \in S, b_1 \in B, b_2 \in B & basecount(sb_1, b_2) & = \begin{cases} 1 + basecount(s, b_2) \\ \qquad \text{when } b_1 = b_2 \\ basecount(s, b_2) \\ \qquad \text{when } b_1 \neq b_2 \end{cases}
\end{array}$$

We proceed by structural induction:

- Basis step: We consider the four cases of $b \in \{\mathtt{A}, \mathtt{C}, \mathtt{U}, \mathtt{G}\}$, in each to show is that $rnalen(b) = basecount(b, \mathtt{A}) + basecount(b, \mathtt{C}) + basecount(b, \mathtt{U}) + basecount(b, \mathtt{G})$.

  $$rnalen(\mathtt{A}) = 1 \qquad \text{by the basis step in the recursive definition of } rnalen$$

  $$basecount(\mathtt{A}, \mathtt{A}) + basecount(\mathtt{A}, \mathtt{C}) + basecount(\mathtt{A}, \mathtt{U}) + basecount(\mathtt{A}, \mathtt{G}) = 1 + 0 + 0 + 0 = 1$$

  by the basis step in the recursive definition of *basecount*, where the first term is the case where $b_1 = b_2 = \mathtt{A}$ and the other three terms are the case where $b_1 \neq b_2$. Thus, $rnalen(\mathtt{A}) = 1 = basecount(\mathtt{A}, \mathtt{A}) + basecount(\mathtt{A}, \mathtt{C}) + basecount(\mathtt{A}, \mathtt{U}) + basecount(\mathtt{A}, \mathtt{G})$.

  Similarly,

  $$rnalen(\mathtt{C}) = 1 \qquad \text{by the basis step in the recursive definition of } rnalen$$

  $$basecount(\mathtt{C}, \mathtt{A}) + basecount(\mathtt{C}, \mathtt{C}) + basecount(\mathtt{C}, \mathtt{U}) + basecount(\mathtt{C}, \mathtt{G}) = 0 + 1 + 0 + 0 = 1$$

  by the basis step in the recursive definition of *basecount*, where the second term is the case where $b_1 = b_2 = \mathtt{C}$ and the other three terms are the case where $b_1 \neq b_2$. Thus, $rnalen(\mathtt{C}) = 1 = basecount(\mathtt{C}, \mathtt{A}) + basecount(\mathtt{C}, \mathtt{C}) + basecount(\mathtt{C}, \mathtt{U}) + basecount(\mathtt{C}, \mathtt{G})$;

  $$rnalen(\mathtt{U}) = 1 \qquad \text{by the basis step in the recursive definition of } rnalen$$

  $$basecount(\mathtt{U}, \mathtt{A}) + basecount(\mathtt{U}, \mathtt{C}) + basecount(\mathtt{U}, \mathtt{U}) + basecount(\mathtt{U}, \mathtt{G}) = 0 + 0 + 1 + 0 = 1$$

  by the basis step in the recursive definition of *basecount*, where the third term is the case where $b_1 = b_2 = \mathtt{U}$ and the other three terms are the case where $b_1 \neq b_2$. Thus, $rnalen(\mathtt{U}) = 1 = basecount(\mathtt{U}, \mathtt{A}) + basecount(\mathtt{U}, \mathtt{C}) + basecount(\mathtt{U}, \mathtt{U}) + basecount(\mathtt{U}, \mathtt{G})$;

  $$rnalen(\mathtt{G}) = 1 \qquad \text{by the basis step in the recursive definition of } rnalen$$

  $$basecount(\mathtt{G}, \mathtt{A}) + basecount(\mathtt{G}, \mathtt{C}) + basecount(\mathtt{G}, \mathtt{U}) + basecount(\mathtt{G}, \mathtt{G}) = 0 + 0 + 0 + 1 = 1$$

  by the basis step in the recursive definition of *basecount*, where the last term is the case where $b_1 = b_2 = \mathtt{G}$ and the other three terms are the case where $b_1 \neq b_2$. Thus, $rnalen(\mathtt{G}) = 1 = basecount(\mathtt{G}, \mathtt{A}) + basecount(\mathtt{G}, \mathtt{C}) + basecount(\mathtt{G}, \mathtt{U}) + basecount(\mathtt{G}, \mathtt{G})$. The basis step is now complete.

- Recursive step: Consider an arbitrary strand s and an arbitrary base b. Assume, as the **induction hypothesis** that

  $$rnalen(s) = basecount(s, \mathtt{A}) + basecount(s, \mathtt{C}) + basecount(s, \mathtt{U}) + basecount(s, \mathtt{G})$$

  (continued next page)

- (Recursive step, continued). We need to show that

  $$rnalen(sb) = basecount(sb, \texttt{A}) + basecount(sb, \texttt{C}) + basecount(sb, \texttt{U}) + basecount(sb, \texttt{G})$$

  We consider the four cases of $b \in \{\texttt{A}, \texttt{C}, \texttt{U}, \texttt{G}\}$. In each case, the LHS of the to show is

  $$rnalen(sb) = 1 + rnalen(s) \qquad \text{by the recursive step in the definition of } rnalen$$

  For the RHS, when $b = \texttt{A}$:

  $$\begin{aligned} RHS =& basecount(s\texttt{A}, \texttt{A}) + basecount(s\texttt{A}, \texttt{C}) + basecount(s\texttt{A}, \texttt{U}) + basecount(s\texttt{A}, \texttt{G}) \\ =& (1 + basecount(s, \texttt{A})) + basecount(s, \texttt{C}) + basecount(s, \texttt{U}) + basecount(s, \texttt{G}) \\ =& 1 + (basecount(s, \texttt{A})) + basecount(s, \texttt{C}) + basecount(s, \texttt{U}) + basecount(s, \texttt{G})) \\ =& 1 + rnalen(s) \qquad \text{by the induction hypothesis} \end{aligned}$$

  (where the first equation listed is by the recursive step in the definition of basecount, the first term from the case $b_1 = b_2 = \texttt{A}$ and the rest of the terms from the case $b_1 \neq b_2$)). Thus, $LHS = rnalen(s\texttt{A}) = 1 + rnalen(s) = RHS$.

  Similarly, when $b = \texttt{C}$:

  $$\begin{aligned} RHS =& basecount(s\texttt{C}, \texttt{A}) + basecount(s\texttt{C}, \texttt{C}) + basecount(s\texttt{C}, \texttt{U}) + basecount(s\texttt{C}, \texttt{G}) \\ =& basecount(s, \texttt{A}) + (1 + basecount(s, \texttt{C})) + basecount(s, \texttt{U}) + basecount(s, \texttt{G}) \\ =& 1 + rnalen(s) \qquad \text{by the induction hypothesis} \end{aligned}$$

  (where the first equation listed is by the recursive step in the definition of basecount, the second term from the case $b_1 = b_2 = \texttt{C}$ and the rest of the terms from the case $b_1 \neq b_2$)). Thus, $LHS = rnalen(s\texttt{C}) = 1 + rnalen(s) = RHS$. When $b = \texttt{U}$:

  $$\begin{aligned} RHS =& basecount(s\texttt{U}, \texttt{A}) + basecount(s\texttt{U}, \texttt{C}) + basecount(s\texttt{U}, \texttt{U}) + basecount(s\texttt{U}, \texttt{G}) \\ =& basecount(s, \texttt{A}) + basecount(s, \texttt{C}) + (1 + basecount(s, \texttt{U})) + basecount(s, \texttt{G}) \\ =& 1 + rnalen(s) \qquad \text{by the induction hypothesis} \end{aligned}$$

  (where the first equation listed is by the recursive step in the definition of basecount, the third term from the case $b_1 = b_2 = \texttt{U}$ and the rest of the terms from the case $b_1 \neq b_2$)). Thus, $LHS = rnalen(s\texttt{U}) = 1 + rnalen(s) = RHS$. Finally, for $b = \texttt{G}$:

  $$\begin{aligned} RHS =& basecount(s\texttt{G}, \texttt{A}) + basecount(s\texttt{G}, \texttt{C}) + basecount(s\texttt{G}, \texttt{U}) + basecount(s\texttt{G}, \texttt{G}) \\ =& basecount(s, \texttt{A}) + basecount(s, \texttt{C}) + basecount(s, \texttt{U}) + (1 + basecount(s, \texttt{G})) \\ =& 1 + rnalen(s) \qquad \text{by the induction hypothesis} \end{aligned}$$

  (where the first equation listed is by the recursive step in the definition of basecount, the last term from the case $b_1 = b_2 = \texttt{G}$ and the rest of the terms from the case $b_1 \neq b_2$)). Thus, $LHS = rnalen(s\texttt{G}) = 1 + rnalen(s) = RHS$.

Since the basis step and recursive steps are complete, we have proved that the equation holds for all RNA strands.

(b) $\forall l_1 \in L \, \forall l_2 \in L \, (\mathit{sum}(\mathit{concat}(l_1, l_2)) = \mathit{sum}(l_1) + \mathit{sum}(l_2))$, where the function *concat* is defined as:

		$\mathit{concat} : L \times L$	$\to L$
Basis Step:	If $l \in L$	$\mathit{concat}([], l)$	$= l$
Recursive Step:	If $l, l' \in L$ and $n \in \mathbb{N}$, then	$\mathit{concat}((n, l), l')$	$= (n, \mathit{concat}(l, l'))$

and the function $\mathit{sum} : L \to \mathbb{N}$ that sums all the elements of a list and is defined by:

		$\mathit{sum} : L$	$\to \mathbb{N}$
Basis Step:		$\mathit{sum}([])$	$= 0$
Recursive Step:	If $l \in L, n \in \mathbb{N}$	$\mathit{sum}((n, l))$	$= n + \mathit{sum}(l)$

We proceed by structural induction on l_1 by definition of L.

- Basis step: To show is that $\forall l_2 \in L \, (\mathit{sum}(\mathit{concat}([], l_2)) = \mathit{sum}([]) + \mathit{sum}(l_2))$. We proceed by universal generalization and let l_2 be an arbitrary element of L. By definition of *concat*, we can rewrite the To Show as $\mathit{sum}(l_2) = \mathit{sum}([]) + \mathit{sum}(l_2)$. By definition of *sum*, we can rewrite the right-hand side as $0 + \mathit{sum}(l_2)$. This shows that our goal can be rewritten to $\mathit{sum}(l_2) = \mathit{sum}(l_2)$, which is true.

- Recursive step: To show is that $\forall l_2 \in L \, (\mathit{sum}(\mathit{concat}((n, l_1'), l_2)) = \mathit{sum}((n, l_1')) + \mathit{sum}(l_2))$, where $n \in \mathbb{N}$ and $l_1' \in L$. We assume as the inductive hypothesis that $\forall l_2 \in L \, (\mathit{sum}(\mathit{concat}(l_1', l_2)) = \mathit{sum}(l_1') + \mathit{sum}(l_2))$.

  We proceed by universal generalization and assume that l_2 is an arbitrary element of L. Applying the definition of *concat*, we can rewrite To Show as

  $$\mathit{sum}((n, \mathit{concat}(l_1', l_2))) = \mathit{sum}((n, l_1')) + \mathit{sum}(l_2)$$

  . We can further apply the definition of *sum* on both sides to rewrite To Show as:

  $$n + \mathit{sum}(\mathit{concat}(l_1', l_2))) = n + \mathit{sum}(l_1') + \mathit{sum}(l_2)$$

  . By the rules of arithmetic on +, this is true when

  $$\mathit{sum}(\mathit{concat}(l_1', l_2))) = \mathit{sum}(l_1') + \mathit{sum}(l_2)$$

  . Since $l_2 \in L$, we can apply the inductive hypothesis, whose body exactly matches this goal.

5. **Induction and Recursion** At time 0, a particle resides at the point 0 on the real line. Within 1 second, it divides into 2 particles that fly in opposite directions and stop at distance 1 from the original particle. Within the next second, each of these particles again divides into 2 particles flying in opposite directions and stopping at distance 1 from the point of division, and so on. Whenever particles meet they annihilate (leaving nothing

behind). How many particles will there be at time $2^{11}-1$? You do not need to justify your answer. Hint: Derive a formula for the number/ locations of particles at time 2^n-1 for arbitrary positive integer n, prove the formula using induction, and apply it when $n=11$.

2^{11} **particles**

Proof: We will prove by mathematical induction that,

> For each positive integer n, a particle that started at time 0 in position 0 has the property that between time 0 and time 2^n-1, it and all of its descendants have stayed in the interval $[-2^n+1, 2^n-1]$ and at time 2^n-1, its descendants are located exactly on each odd location in $[-2^n+1, 2^n-1]$ (inclusive).

How to use this lemma? Substituting $n=11$, this lemma says that at time $2^{11}-1$, the descendants of the original particle are located at odd each location in the interval $[-2^{11}+1, 2^{11}-1]$ (inclusive). Therefore, the number of particles at this time is the number of odd integers in the interval. There are $\lceil\frac{2^{11}-1}{2}\rceil = \frac{2^{11}}{2} = 2^{10}$ many such negative odd integers, and $\lceil\frac{2^{11}-1}{2}\rceil = \frac{2^{11}}{2} = 2^{10}$ many such positive odd integers, so a total of $2^{10}+2^{10}=2^{11}$ particles.

To prove the lemma, we proceed by mathematical induction.

- **Basis step** For $n=1$, we WTS that between time 0 and time 2^1-1, the particle and all its descendants have stayed in the interval $[-2^1+1, 2^1-1]$, and at time 2^1-1, there are particles at each odd location this interval. Plugging in $n=1$, we see that the interval is $[-1,1]$ and we WTS that at time 1 there are particles exactly at the endpoints of this interval (these are the odd locations). Applying the definition of the process, at time 1, the original particle divides into two and the two new particles reach distance 1 from the original. This means that the only locations at which particles resided in this interval are $0, -1, 1$ (all within the required interval) and that at time 1, the particles are at -1 and 1, as required.

(continued next page)

- **Induction step** Let k be an arbitrary positive integer. Assume, as the **Induction Hypothesis (IH)**, that

  > between time 0 and time $2^k - 1$, the particle and all of its descendants have stayed in the region on the number line $[-2^k + 1, 2^k - 1]$ and at time $2^k - 1$, its descendants are located exactly on each odd location in $[-2^k + 1, 2^k - 1]$ (inclusive).

  We WTS that

  > between time 0 and time $2^{k+1} - 1$, the particle and all of its descendants have stayed in the region on the number line $[-2^{k+1} + 1, 2^{k+1} - 1]$ and at time $2^{k+1} - 1$, its descendants are located exactly on each odd location in $[-2^{k+1} + 1, 2^{k+1} - 1]$ (inclusive).

  We'll consider the progress of the experiment through time. At time step 2^k, one unit of time has elapsed since the situation described in the IH. By the setup of the experiment, during this unit of time, each particle splits into two and these travel one unit in each axis direction. Since (by the IH) at time $2^k - 1$, there are particles at each odd location in the interval $[-2^k + 1, 2^k - 1]$, each of these particles are two units apart. Thus, the particles originating from successive odd locations in time $2^k - 1$ will meet and annihilate at time 2^k. The only particles that will remain at time 2^k are (1) the particle that originates at time $2^k - 1$ from location $-2^k + 1$ and goes left and (2) the particle that originates at time $2^k - 1$ from location $2^k - 1$ and goes right. Thus, at 2^k there are two particles, one at -2^k and the other at 2^k. For definiteness, let's call the particle at position -2^k, p_L, and the one at position 2^k, p_R. Recalibrating the experiment with position 2^k as the new "origin" and particle p_R as our original particle, the IH guarantees that between the current time (2^k) and $2^k - 1$ steps in the future, p_R will divide into particles such that all of its descendants will stay in the interval ["origin" $- 2^k + 1$, "origin" $+ 2^k - 1$] and its descendants will end up exactly on each odd location in ["origin" $- 2^k + 1$, "origin" $+ 2^k - 1$] (inclusive). That is, at time $2^k + 2^k - 1 = 2^{k+1} - 1$, the process starting p_R will have generated particles at all odd locations in $[0, 2^{k+1} - 1]$. Symmetrically, the IH guarantees that at time $2^{k+1} - 1$, all descendants of the particle p_L will be at the odd locations in $[-2^{k+1} + 1, 0]$, and will never have gone beyond that interval at any time since time $2^k - 1$. Since the intervals in which the descendants of p_L and p_R land between these timestamps only overlap at 0, the particles generated by these two experiments never collide and so the two experiments run independently. Thus, at time $2^{k+1} - 1$, there are $2^k + 2^k = 2^{k+1}$ many particles, occupying each odd location in the interval $[-2^{k+1} + 1, 0] \cup [0, 2^{k+1} - 1] = [-2^{k+1} + 1, 2^{k+1} - 1]$, and these particles (and their predecessors) have not left this interval at any point up to this time step. In particular, this means that the lemma is proved.

6. **Induction and Recursion** Prove that every positive integer has a base 3 expansion. *Hint: use strong induction.*

By definition of base expansion, we need to prove that for every positive integer n, there is a positive integer k and nonnegative integers $a_0, a_1, \ldots, a_{k-1}$ such that each $a_i \in \{0,1,2\}$, $a_{k-1} \neq 0$, and

$$n = \sum_{i=0}^{k-1} a_i 3^i$$

We proceed by strong induction on $n \geq 0$.

- Basis steps (using the terminology for strong induction: we choose $b = 1, j = 1$): For the positive integer 1, we have witnesses $k = 1, a_0 = 1$ and since $1 \in \{0,1,2\}$, $1 \neq 0$, and $1 = \sum_{i=0}^{0} a_i 3^i$, the claim is proved. For the positive integer 2, we have witnesses $k = 1, a_0 = 2$ and since $2 \in \{0,1,2\}$, $2 \neq 0$, and $2 = \sum_{i=0}^{0} a_i 3^i$, the claim is proved.

- Recursive step: Consider an arbitrary positive integer n greater than or equal to 2. Assume, as the strong induction induction hypothesis that each positive number less than or equal to n has a base 3 expansion. We need to show that $n+1$ has a base 3 expansion. To build the required witnesses, consider the integer $c = (n+1)$ **div** 3. By properties of integers, since $n + 1 \geq 3$ (because $n \geq 2$), $1 \leq c \leq n$. Thus, the strong induction hypothesis applies to c and there is a positive integer k_c and nonnegative integers $x_0, x_1, \ldots, x_{k_c-1}$ such that each $x_i \in \{0,1,2\}$, $x_{k_c-1} \neq 0$, and

  $$c = \sum_{i=0}^{k_c-1} x_i 3^i$$

  Define $k = k_c + 1$, $a_i = x_{i-1}$ for each i from 1 to k, and $a_0 = (n+1)$ **mod** 3. Then k is positive (because k_c is), each $a_i \in \{0,1,2\}$ (because all the x_i are in this set, and the remainder upon division by 3 of any positive integer is also in this set), $a_{k-1} = a_{k_c+1-1} = x_{k_c-1} \neq 0$ (by choice of k and definition of x_is). It remains to prove that the sum gives the right value:

  $$\sum_{i=0}^{k-1} a_i 3^i = \left(\sum_{i=1}^{k-1} a_i 3^i\right) + a_0 3^0 = 3\left(\sum_{i=1}^{(k_c+1)-1} x_{i-1} 3^{i-1}\right) + a_0 3^0 = 3\left(\sum_{j=0}^{k_c-1} x_j 3^j\right) + a_0$$
  $$= 3c + a_0 = 3((n+1) \textbf{ div } 3) + ((n+1) \textbf{ mod } 3) = n+1$$

  as required.

Since the existence of base 3 expansions was proved for all positive integers (for $1, 2$ in the basis cases, and all integers greater than or equal to 3 in the recursive step), the proof by strong induction is complete.

7. **Functions & Cardinalities of sets** Prove each of the following claims.

   (a) For the "function" $f : \mathbb{Z} \to \{0,1,2,3\}$ given by $f(x) = x$ **mod** 5, it is not the case that every element of the domain maps to exactly one element of the codomain (that is, it is not a well-defined function).

This function is not well defined because $f(4)$ is not in the codomain even though 4 is in the domain: $f(4) = 4 \textbf{ mod } 5 = 4 \notin \{0, 1, 2, 3\}$.

(b) The "function" $f : \mathcal{P}(\mathbb{Z}^+) \to \mathbb{Z}$ given by

$$f(A) = \text{the maximum element in A}$$

is not well-defined.

This function is not well defined because $f(\mathbb{Z}^+)$ is not well-defined even though $\mathbb{Z}^+$ is in the domain: the set $\mathbb{Z}^+$ doesn't have a maximum element.

(c) There is a one-to-one function with domain $\{a, b, c\}$ and codomain $\mathbb{R}$.

Define the function $f : \{a, b, c\} \to \mathbb{R}$ by the piecewise definition

$$f(a) = 1, f(b) = 2, f(c) = 3$$

This function is well-defined because it maps every domain element to a unique output, and all outputs are in the specified codomain. Moreover, it is one-to-one: each of the three distinct domain elements have different images from one another.

(d) There is an onto function with domain R and codomain $\{\pi, \frac{1}{17}\}$, where R is the set of user ratings in a database with 5 movies.

Define the function $f : R \to \{\pi, \frac{1}{17}\}$ by the piecewise definition

$$f(w) = \begin{cases} \pi & \text{if } w = (0, 0, 0, 0, 0) \\ \frac{1}{17} & \text{otherwise} \end{cases}$$

This function is well-defined because it maps every ratings 5-tuple to a unique output, and all outputs are in the specified codomain. Moreover, it is onto: we need to confirm that each element in the codomain has a preimage. Consider, for example

$$f((0, 0, 0, 0, 0)) = \pi \qquad f((1, 1, 1, 1, 1)) = \frac{1}{17}.$$

(e) The Cartesian product $\mathbb{Z}^+ \times \{a, b, c\}$ is countable.

We rewrite

$$\mathbb{Z}^+ \times \{a, b, c\} = \big(\mathbb{Z}^+ \times \{a\}\big) \cup \big(\mathbb{Z}^+ \times \{b\}\big) \cup \big(\mathbb{Z}^+ \times \{c\}\big)$$

Using the function $f : \mathbb{Z}^+ \to \mathbb{Z}^+ \times \{a\}$ given by $f(n) = (n, a)$, we can prove that $\mathbb{Z}^+ \times \{a\}$ is countably infinite. Similarly, $\mathbb{Z}^+ \times \{b\}$ and $\mathbb{Z}^+ \times \{c\}$ are countably infinite. By Theorem 1 on page 174 in the book, $(\mathbb{Z}^+ \times \{a\}) \cup (\mathbb{Z}^+ \times \{b\})$ is countable because it is the union of two countably infinite (hence countable) sets. Applying this theorem again when taking the union of this resulting set with the set $(\mathbb{Z}^+ \times \{c\})$, we see that the set is countable.

(f) The interval of real numbers $\{x \in \mathbb{R} \mid 5 \leq x \leq 8\}$ is uncountable. *Hint: you may use the fact that the unit interval* $\{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ *is uncountable.*

For brevity, we use the interval notation $[5,8] = \{x \in \mathbb{R} \mid 5 \leq x \leq 8\}$ and $[0,1] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$. Define the function $f : [0,1] \to [5,8]$ by $f(x) = 5 + 3x$. Then f is a bijection:

- Well-defined? for each x in $[0,1]$, $0 \leq x \leq 1$ so $0 \leq 3x \leq 3$ and $5 \leq 5 + 3x \leq 8$. Thus, $f(x) \in [5,8]$, as required.
- Invertible? Consider the function $g : [5,8] \to [0,1]$ defined by $g(x) = \frac{x-5}{3}$. WTS that for each $x \in [0,1]$, $g(f(x)) = x$ and for each $x \in [5,8]$, $f(g(x)) = x$. Let $x \in [0,1]$. Then

  $$g(f(x)) = g(5+3x) = \frac{(5+3x)-5}{3} = \frac{3x}{3} = x,$$

  as required. Similarly, let $x \in [5,8]$. Then

  $$f(g(x)) = f(\frac{x-5}{3}) = 5 + 3 \cdot \frac{x-5}{3} = 5 + (x-5) = x,$$

  as required. Thus, f is invertible (with inverse g) and so is a bijection.