Problem 1: Set Theory Evaluation

Evaluate each statement as `true` or `false`. Provide brief justifications. Consider a and b to be distinct ($a \neq b$) *urelements* (atomic objects that are not sets).

1. $a \in \{\{a\}, b\}$
2. $\{a\} \in \{a, \{a\}\}$
3. $\{a\} \subseteq \{\{a\}, \{b\}\}$
4. $\{a, b\} \in \{a, b\}$
5. $\{\{a\}, b\} \subseteq \{a, \{a, b\}, \{b\}\}$
6. $\{\{a\}\} \subset \{\{a\}, \{a\}\}$
7. $\{a, a, a\} \setminus \{a\} = \{a, a\}$
8. $\varnothing \in \varnothing$
9. $\varnothing \in \{\varnothing\}$
10. $\varnothing \in \{\{\varnothing\}\}$
11. $\varnothing \subseteq \varnothing$
12. $\varnothing \subset \varnothing$
13. $\varnothing \subseteq \{\{\varnothing\}\}$
14. $\{\varnothing, \varnothing\} \subset \{\varnothing\}$
15. $\mathcal{P}(\varnothing) = \{\varnothing\}$
16. $a \in \mathcal{P}(\{a\})$
17. $\mathcal{P}(\{a, \varnothing\}) \subset \mathcal{P}(\{a, b, \varnothing\})$
18. $\{a, b\} \subseteq \mathcal{P}(\{a, b\})$
19. $\{a, a\} \in \mathcal{P}(\{a, a\})$
20. $\{\{a\}, \varnothing\} \subseteq \mathcal{P}(\{a, a\})$
21. $\mathcal{P}(\{a, b\}) \supseteq \{\{a\}, \{\varnothing\}\}$

Problem 2: Set Operations

A cybersecurity team monitors different types of *network threats*. They classify threats into sets based on their attack vectors:

- $A = \{\text{malware}, \text{phishing}, \text{ddos}, \text{ransomware}, \text{botnet}\}$ (currently *actively* detected threats)
- $P = \{\text{phishing}, \text{social-eng}, \text{ddos}, \text{insider}, \text{malware}\}$ (threats targeting *humans*)
- $N = \{\text{ransomware}, \text{cryptojack}, \text{ddos}, \text{botnet}, \text{worm}\}$ (threats requiring *network access*)

The universal set $T = A \cup P \cup N$ contains all distinct threat types mentioned above.

Note: All complements ($\overline{X}$) are taken relative to the universal set T.

Part (a): Compute the following and interpret each result in cybersecurity context:

1. $A \cap P$
2. $A \cup N$
3. $(A \cap P) \setminus N$
4. $\overline{A} \cap P$
5. $A \triangle P$
6. $P \setminus (A \cup N)$
7. $\mathcal{P}(\{A, P, N\})$
8. $\left|(P \cup N) \cap \overline{A}\right|$

Part (b): The security team needs to triage threats effectively. Define *priority levels*:

- Critical: $C = A \cap P \cap N$ (active, human-targeted, network-based)
- High: $H = (A \cap P) \setminus N$ (active and human-targeted, but not network-based)
- Medium: $M = A \setminus (C \cup H)$ (remaining active threats)

1. Compute C, H, M.
2. Determine whether $\{C, H, M\}$ is a partition of A.

Part (c): Draw a Venn diagram showing sets A, P, and N with all threat types labeled in their appropriate regions. Use colors to annotate the priority categories.

Problem 3: Similarity and Distance Metrics

Streaming services use similarity measures to recommend content.

Consider *user preferences* as sets of genres they enjoy. For example, if Anna loves mind-bending plots, her preference set is $A = \{\text{sci-fi}, \text{thriller}\}$.

	Sci-fi	**Thriller**	**Drama**	**Romance**	**Horror**	**Comedy**	**Action**	**Fantasy**
Anna	✓	✓	✗	✗	✗	✗	✗	✗
Boris	✗	✓	✓	✓	✗	✗	✗	✗
Clara	✗	✗	✓	✓	✓	✓	✗	✗
Diana	✓	✗	✗	✗	✗	✓	✓	✓

Part (a): The *Jaccard similarity* measures how much two users' tastes overlap:

$$\mathcal{J}(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{\text{shared preferences}}{\text{total unique preferences}}$$

with the convention that $\mathcal{J}(\varnothing, \varnothing) = 1$.

The *Jaccard distance* measures how different users are:

$$d_{\mathcal{J}}(X, Y) = 1 - \mathcal{J}(X, Y)$$

1. Calculate $\mathcal{J}(X, Y)$ and $d_{\mathcal{J}}(X, Y)$ for all pairs among users.
2. Determine which pair is most similar and which is most dissimilar.
3. Draw a social network graph with users as nodes and edges weighted by Jaccard similarity, excluding edges with weight 0.
4. Build $G_{0.25}$: the graph with edges where Jaccard similarity ≥ 0.25. List all connected components.

Part (b): The *Cosine similarity* for sets can be defined[1] as:

$$\mathcal{C}(X, Y) = \frac{|X \cap Y|}{\sqrt{|X| \cdot |Y|}}$$

The *Cosine distance* is $d_{\mathcal{C}}(X, Y) = 1 - \mathcal{C}(X, Y)$.

1. Calculate $\mathcal{C}(X, Y)$ and $d_{\mathcal{C}}(X, Y)$ for all user pairs.
2. Determine which pair is most similar and which is most dissimilar.
3. Draw a graph with users as nodes and edges weighted by Cosine similarity.
4. Show that $\mathcal{J}(X, Y) \leq \mathcal{C}(X, Y)$ for all nonempty finite sets X, Y. When the equality holds?

Part (c): Prove that Jaccard distance satisfies the triangle inequality:

$$d_{\mathcal{J}}(A, C) \leq d_{\mathcal{J}}(A, B) + d_{\mathcal{J}}(B, C)$$

for arbitrary finite sets A, B, and C.

[1] Otsuka–Ochiai coefficient

Part (d): Show that cosine distance does NOT satisfy the triangle inequality by providing a specific counterexample. Find three non-empty sets X, Y, and Z such that:

$$d_{\mathcal{C}}(X, Z) > d_{\mathcal{C}}(X, Y) + d_{\mathcal{C}}(Y, Z)$$

Part (e): A new user joins with preferences $U = \{\text{thriller}, \text{horror}\}$. Using Jaccard similarity, find existing users with similarity ≥ 0.25 to recommend as "users with similar taste."

Challenge: Design your own similarity metric that you think would work better than Jaccard for movie recommendations. Explain your reasoning.

Problem 4: Logic and Set Identities

This problem bridges set theory and logical reasoning, preparing for formal proofs.

Part (a): Translate each statement to first-order logic with quantifiers over a universal set U:
1. $A \subseteq B$
2. $A = B$
3. $A \subseteq B \leftrightarrow \mathcal{P}(A) \subseteq \mathcal{P}(B)$

Part (b): Prove the following identities using both Venn diagrams and symbolic reasoning:
1. $(A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$
2. De Morgan's laws: $\overline{A \cup B} = \overline{A} \cap \overline{B}$ and $\overline{A \cap B} = \overline{A} \cup \overline{B}$
3. $A \subseteq B$ if and only if $A \cap B = A$ if and only if $A \cup B = B$
4. Distributive law: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Part (c): For any universe U and a set $X \subseteq U$, prove that the complement operator is:
1. An *involution*: $\overline{\overline{X}} = X$
2. Order-reversing (*anti-monotonic*): if $X \subseteq Y$, then $\overline{Y} \subseteq \overline{X}$

Problem 5: Coordinate Systems

A game developer is designing a 2D puzzle game with different gameplay zones. Each zone is defined by specific coordinate regions in $\mathbb{R}^2$.

Part (a): Sketch all gameplay zones on the coordinate plane:
1. Game Area: $G = [0; 8] \times [0; 7]$
2. Safe Zone: $S = (1; 4) \times (5; 7]$
3. Impassible Wall: $W = \{\langle x, 4\rangle \mid 0 \leq x \leq 6\}$
4. Danger Zone: $D = \{\langle x, y\rangle \in G \mid y < x \text{ or } y < 4\}$
5. Treasure Zones: $T = \{\langle x, y\rangle \mid x \in \{1, 2, 3\}, 1 \leq y < 3\}$
6. Boss Arena: $B = \{\langle x, y\rangle \in G \mid 16(x-9)^2 + 25y^2 \leq 400\}$

Part (b): Power-ups spawn at lattice points (integer coordinates) within the Danger Zone D, excluding the wall W and borders of G. Count the number of such points.

Part (c): A player starts at position $P_0 = \langle 2, 6\rangle$ and makes exactly three moves according to vectors $v_1 = \langle 4, 0\rangle$, $v_2 = \langle 1, -2\rangle$, and $v_3 = \langle -4, -3\rangle$ (in that order).
1. Calculate the player's position after each move: $P_i = P_{i-1} + v_i$ for $i = 1, 2, 3$.
2. Determine which zones the player is in after each move.

3. Does the player ever enter the Boss Arena B?

Problem 6: Self-Referential Set Puzzles

In computer science, recursive data structures reference themselves. These mathematical puzzles explore similar self-referential concepts that appear in programming, logic, and even philosophy.

Part (a): Find all sets X and Y that satisfy this system:

$$X = \{1, 2, |Y|\}$$
$$Y = \{|X|, 3, 4\}$$

Start by determining possible values for $|X|$ and $|Y|$, then verify which combinations work.

Part (b): Consider a more complex system:

$$A = \{1, |B|, |C|\}$$
$$B = \{2, |A|, |C|\}$$
$$C = \{1, 2, |A|, |B|\}$$

Find all valid solutions (A, B, C). Explain why some potential solutions don't work.

Part (c): Design your own *non-trivial* self-referential set system involving 2–4 sets.

Problem 7: Fuzzy Logic

In the real world, boundaries aren't always crisp. Is a 180cm person tall? Is 10°C warm? *Fuzzy sets* model this uncertainty and are crucial in AI, machine learning, and control systems.

Unlike classical sets where membership is binary (in/out), fuzzy sets assign each element a *membership degree* $\mu(x) \in [0; 1] \subseteq \mathbb{R}$ representing how "strongly" the element belongs.

Consider two fuzzy sets over $X = \{a, b, c, d, e\}$:

$$F = \{a : 0.4, b : 0.8, c : 0.2, d : 0.9, e : 0.7\}$$
$$R = \{a : 0.6, b : 0.9, c : 0.4, d : 0.1, e : 0.5\}$$

Here, for example, $\mu_F(b) = 0.8$ means element b belongs to fuzzy set F with degree 0.8.

Part (a): Define the complement of a fuzzy set S to be $\mu_{\overline{S}}(x) = 1 - \mu_S(x)$. Compute $\overline{F}$ and $\overline{R}$.

Part (b): For the union, define $\mu_{S \cup T}(x) = \max\{\mu_S(x), \mu_T(x)\}$. Compute $F \cup R$.

Part (c): For the intersection, define $\mu_{S \cap T}(x) = \min\{\mu_S(x), \mu_T(x)\}$. Compute $F \cap R$.

Part (d): Propose and justify a definition for $S \setminus T$. Compute $F \setminus R$ and $R \setminus F$.

Part (e): One fuzzy analogue of Jaccard similarity is:

$$\tilde{\mathcal{J}}_f(F, R) = \frac{\sum_{x \in X} \min\{\mu_F(x), \mu_R(x)\}}{\sum_{x \in X} \max\{\mu_F(x), \mu_R(x)\}}$$

Compute $\tilde{\mathcal{J}}_f(F, R)$ and the corresponding distance $1 - \tilde{\mathcal{J}}_f(F, R)$.

Part (f): Defuzzification. Suppose a system triggers an alert if an element's membership in $F \cup R$ exceeds 0.75. List all triggered elements with their membership degrees, and briefly discuss how changing this threshold value would affect the system's sensitivity and the number of alerts.

Problem 8: Power Sets

Let A and B be finite sets. For each statement below, either provide a *rigorous proof* or find a *counterexample* that disproves the claim.

1. If $A \subseteq B$, then $\mathcal{P}(A) \subseteq \mathcal{P}(B)$
2. $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$
3. $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$
4. $|\mathcal{P}(A \times B)| = 2^{|A| \cdot |B|}$

Submission Guidelines:

- Show all work and reasoning clearly for computational problems.
- For proofs, state what you're proving, provide clear logical steps, and conclude with QED or $\square$.
- For false statements, provide specific counterexamples.
- Collaborate with classmates, but write all solutions independently.
- Submit as PDF with clearly labeled problems and legible work.

Grading Rubric:

- Computational accuracy: 50% (Getting the right answer)
- Mathematical reasoning and proof quality: 30% (Showing clear logical thinking)
- Presentation and clarity: 20% (Making your solutions easy to follow)