PROBLEM 1. (10pts)

Translate to Logic following the statement

"Every man likes all tasty apples "

1. Domain: \( X \neq \emptyset \)

2. Predicates:
   \( A(x) \) – x is an Apple,
   \( M(x) \) – x is a man,
   \( T(x) \) – x is tasty.
   \( L(x,y) \) – x likes y

3. Connectives: \( \land, \Rightarrow \)

4. Quantifiers:
   \( \forall_{M(x)} \) – "Every man",
   \( \forall (A(y) \land T(y)) \) – "All tasty apples" (restricted)

6. RESTRICTED FORMULA:

   \( \forall_{M(x)} \forall (A(y) \land T(y)) \ L(x,y) \)

7. LOGIC FORMULA:

   \( ( \forall_x (M(x) \Rightarrow \forall_y ((A(y) \land T(y)) \Rightarrow L(x,y))) \)
Write your Logic formula as the Intended Interpretation Formula

\[( \forall x (\text{Men}(x) \Rightarrow \forall y ((\text{Apple}(y) \land \text{Tasty}(y)) \Rightarrow \text{Likes}(x,y))) \]

**PROBLEM 2: RESOLUTION** (10pts)

**Q1.** Find all resolvents of the set \( CL = \{ C_1, C_2 \} \) of clauses for

- \( C_1 = \{ a, b, c, \neg d \} \)
- \( C_2 = \{ \neg a, \neg b, d \} \)

It means locate all clauses in \( CL \) that are Complementary Pairs and Resolve them

- \( C_1(a) \) and \( C_2(\neg a) \) resolves on \( \{ b, c, \neg d, \neg b, d \} \)
- \( C_1(b) \) and \( C_2(\neg b) \) resolves on \( \{ a, c, \neg d, \neg a, d \} \)
- \( C_1(\neg d) \) and \( C_2(d) \) resolves on \( \{ a, b, c, \neg a, \neg b \} \)

**Q2. Use the Resolution Completeness to prove that the set**

\( CL = \{ \{ a, b \}, \{ \neg a, c \}, \{ \neg b, c \} \} \) of clauses is SATISFIABLE.

There are only two possible derivations

**D1:** \( \{ a, b \}, \{ \neg a, c \}, \{ \neg b, c \} \)

\( \{ b, c \}, \{ \neg b, c \} \)

\( \{ c \} \)

**D1:** \( \{ a, b \}, \{ \neg a, c \}, \{ \neg b, c \} \)

\( \{ a, c \}, \{ \neg a, c \} \)

\( \{ c \} \)

We can never get a derivation of \( \{ \} \)- so by Completeness the set \( CL \) is satisfiable
PROBLEM 3: RULE BASED SYSTEMS  (10pts)

Here is a small set of RULES proposed for a simple rule-based system for financial advise.

R1  IF savings are not adequate  THEN invest in savings
R2  IF savings are adequate AND income is adequate THEN invest in stock
R3  IF there is no children THEN savings are adequate
R4  IF there is a partner AND partner has a job THEN income is adequate

Q1. (5pts) Conceptualize the rules R1-R4 in propositional convention that admits negation. Explain your solution.

We define.
• S = savings_adequate
• V = invest_savings
• I = income_adequate
• K = invest_stocks
• C = has_children
• P = has_partner
• J = partner_has_job

Here are the rules expressed in propositional logic conceptualization

• R1: ¬ S \(\rightarrow\) V
• R2: S \(\land\) I \(\rightarrow\) K
• R3: ¬ C \(\rightarrow\) S
• R4: P \(\land\) J \(\rightarrow\) I
Q2. (5pts)
1. Conceptualize the rules R1-R4 in predicate convention using predicates attribute(x, value of attribute), attribute(object, value of attribute).

We have the following ATTRIBUTES:

- **Savings**
  Values: adequate, not adequate
- **Income**
  Values: adequate, not adequate
- **InvestStocks**
  Values: yes, no
- **InvestSavings**
  Values: yes, no
- **Children**
  Values: yes, no
- **Partner**
  Values: yes, no
- **PartnerJob**
  Values: yes, no

**RULES:**

R1: Savings(x, not adequate) → InvestSavings(x, yes)

R2: Savings(x, adequate) ∧ Income(x, adequate) → InvestStocks(x, yes)

R3: Children(x, no) → Savings(x, adequate)

R4: Partner(x, yes) ∧ PartnerJob(x, yes) → Income(x, adequate)
2. Write a format of a database TABLE needed for the conceptualization

DATA TABLE - example of a record

<table>
<thead>
<tr>
<th>record</th>
<th>Savings</th>
<th>Income</th>
<th>Children</th>
<th>Partner</th>
<th>PartnerJob</th>
<th>InvestSavings</th>
<th>InvestStocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>not adequate</td>
<td>adequate</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

PROBLEM 4: Classification Rules 10pts

Q1 (5pts)
Given a dataset DB: C – class attribute

<table>
<thead>
<tr>
<th>Record</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>o1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>o2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>o3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>o4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>o5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

For the following formulas use the proper definitions to prove whether they are or they are not discriminant or characteristic rules in the dataset DB

1) $a_1 = 1 \& a_2 = 1 \Rightarrow C = 1$
   
   \{o1\} is a subset of \{o1, o5\} so this is a DISCRIMINANT rule

2) $C = 1 \Rightarrow a_1 = 0 \& a_2 = 1 \wedge a_3 = 1$
   
   \{o: a1 = 0 \& a2 = 1 \& a3 = 1 \} is an empty set so this is not a CHARACTERISTIC rule
3) $a_1 = 1 \Rightarrow C = 1$
   \{o1\} is a subset of \{o1, o5\} so this is a DISCRIMINANT rule

4) $C = 1 \Rightarrow a_1 = 1$
   \{o1, o5\} intersection with \{o1\} is non-empty

5) $a_1 = 2 \& a_2 = 1 \& a_3 = 1 \Rightarrow C = 2$
   \{o5\} is not a subset of \{o2\}, so this is not a DISCRIMINANT rule

Q2 (5pts)

1. Prove that in any classification DB the inverse implication to the discriminant rule is a characteristic rule

By definition, for any database DB :

\[ \text{DESCRIPTION} \Rightarrow \text{CLASS} \]

is a discriminant rule iff

1. \{o: DESCRIPTION\} is not empty
2. \{o: DESCRIPTION\} is included in \{o: CLASS\}

We know that for any non-empty sets A, B, if A is included in B, then their intersection is non-empty.

Hence
\{o: DESCRIPTION\} intersection with \{o: CLASS\} is not empty and by Definition then inverse implication

\[ \text{CLASS} \Rightarrow \text{DESCRIPTION} \]

is a CHARACTERISTIC RULE
SHORT QUESTIONS (10pts)

Q1: (5pts) Define a Classifier

A classifier is a final product of a process that uses data set and a classification algorithm.

The classifier is build i.e. we terminate the process if it has been trained and tested and the predictive accuracy is on an acceptable level.

Q2: (5pts) Write down termination conditions for Decision Tree Model

1. All records (samples) for the given node belong to the same class.

2. There are no remaining attributes on which the samples (records in the data table) may be further partitioned – a LEAF is created with majority vote for training sample.

3. There is no records (samples) left – a LEAF is created with majority vote for training sample.
Majority voting involves converting node N into a leaf and labeling it with the most common class in D which is a set of training tuples and their associated class labels.

**Problem 5: Classification by Decision Tree Algorithm (20 pts)**

Given the following DATA

<table>
<thead>
<tr>
<th>TRAIN</th>
<th>Record</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0_1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0_2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0_3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0_4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0_5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0_6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0_7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0_8</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TEST</th>
<th>Record</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0_1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0_2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0_3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0_4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Use the TRAIN data to **build a CLASSIFIER** using the basic **Decision Tree Algorithm** with \(a_1\) as the ROOT. **Here are the STEPS you must follow**

**STEP 1: (5 pts)** Build the Decision Tree and write Rules in Predicate Form

**STEP 2: (10 pts)**
- Evaluate: (i) **rules accuracy**, (ii) **predictive accuracy**.
- (iii) Write down a **TEST** data that would give a 100% predictive accuracy for your set of rules.

**STEP 3: (5 pts)** Give your answer under a title: **MY CLASSIFIER IS**
Solutions

STEP 1: Build the Decision Tree and write Rules in Predicate Form

R1: $a_1(x,1) \land a_2(x,1) \rightarrow C(x,0)$
R2: $a_1(x,1) \land a_2(x,0) \rightarrow C(x,0)$
R3: $a_1(x,0) \rightarrow C(x,0)$
STEP 2:

Evaluate: (i) **rules accuracy**,
- o₁ is misclassified
- o₂ is classified by R3
- o₃ is classified by R3
- o₄ is classified by R3
- o₅ is misclassified
- o₆ is classified by R1
- o₇ is classified by R3
- o₈ is misclassified
5 Passed/out of 8 = 62.5% rule accuracy

(ii) **predictive accuracy**.
- o₁ is misclassified
- o₂ is classified by R2
- o₃ is misclassified
- o₄ is classified by R3
2 Passed/out of 4 = 50% predictive accuracy

(iii) Write down a **TEST** data that would give a 100% predictive accuracy for your set of rules.

<table>
<thead>
<tr>
<th>Record</th>
<th>a₁</th>
<th>a₂</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>o₁</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>o₂</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>o₃</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>o₄</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

STEP 3: Give your answer under a title: **MY CLASSIFIER IS**
(Answer 1): There is no classifier because predictive accuracy is only 50% and we do not accept the rules.

(Answer 2): I decide to accept predictive accuracy of 50%. **MY CLASSIFIER IS**

R1: a1(x,1) ∧ a2(x,1) → C(x,0), R2: a1(x,1) ∧ a2(x,0) → C(x,0), R3: a1(x,0) → C(x,0)