	INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR	
		Stamp / Signature of the Invigilator

EXAMINATION (End Semester)										**SEMESTER (Autumn)**			
Roll Number										**Section**		**Name**	
Subject Number	C	S	6	0	0	4	5			**Subject Name**		*Artificial Intelligence*	
Department / Center of the Student										**Additional sheets**			

Important Instructions and Guidelines for Students

1. You must occupy your seat as per the Examination Schedule/Sitting Plan.
2. Do not keep mobile phones or any similar electronic gadgets with you even in the switched off mode.
3. Loose papers, class notes, books or any such materials must not be in your possession, even if they are irrelevant to the subject you are taking examination.
4. Data book, codes, graph papers, relevant standard tables/charts or any other materials are allowed only when instructed by the paper-setter.
5. Use of instrument box, pencil box and non-programmable calculator is allowed during the examination. However, exchange of these items or any other papers (including question papers) is not permitted.
6. Write on both sides of the answer script and do not tear off any page. **Use last page(s) of the answer script for rough work.** Report to the invigilator if the answer script has torn or distorted page(s).
7. It is your responsibility to ensure that you have signed the Attendance Sheet. Keep your Admit Card/Identity Card on the desk for checking by the invigilator.
8. You may leave the examination hall for wash room or for drinking water for a very short period. Record your absence from the Examination Hall in the register provided. Smoking and the consumption of any kind of beverages are strictly prohibited inside the Examination Hall.
9. Do not leave the Examination Hall without submitting your answer script to the invigilator. **In any case, you are not allowed to take away the answer script with you.** After the completion of the examination, do not leave the seat until the invigilators collect all the answer scripts.
10. During the examination, either inside or outside the Examination Hall, gathering information from any kind of sources or exchanging information with others or any such attempt will be treated as **'unfair means'**. Do not adopt unfair means and do not indulge in unseemly behavior.

Violation of any of the above instructions may lead to severe punishment.

Signature of the Student

To be filled in by the examiner

Question Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
Marks Obtained											

Marks obtained (in words)	**Signature of the Examiner**	**Signature of the Scrutineer**

CS60045 Artificial Intelligence, Autumn 2023–2024

End-Semester Test

23–November–2023 | 02:00pm–05:00pm | Maximum marks: 100

[*Write your answers in the question paper itself. Be brief and precise. Answer all questions.*]

Do not write anything on this page.

1. [Miscellaneous]

Write brief answers to the following parts. There is no need to write any explanation or justification.

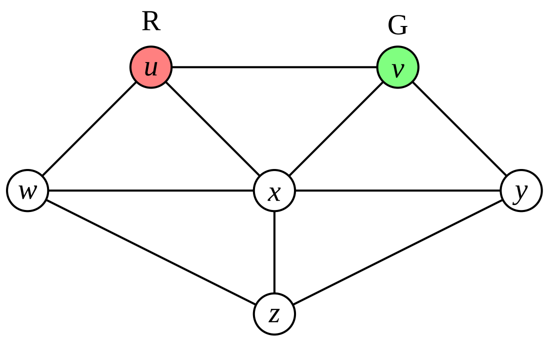
(a) Suppose that we use simulated annealing to minimize a function $f(s)$ of the state s. Let the current state be s, and a random neighbor s' of s is chosen. Let $\Delta f = f(s') - f(s)$. Fill in the following three blanks with the correct answers. Assume that the current temperature is Θ. **(4)**

If $\Delta f < 0$, the state change $s := s'$ is accepted with probability $\underline{\quad 1 \quad}$.

If $\Delta f > 0$, the state change $s := s'$ is accepted with probability $\underline{\quad e^{-\Delta f/\Theta} \quad}$.

If $\Delta f = 0$, the state change $s := s'$ is accepted with probability $\underline{\quad 1 \quad}$.

(b) Suppose that we want to 3-color the following graph with the colors R, G, and B. By u, v, w, x, y, z, we denote the variables standing for the colors that the nodes of the same names receive. We start by setting $u = \text{R}$, and $v = \text{G}$. Immediately after this, the updated domains $\mathscr{D}_y$ and $\mathscr{D}_z$ of the variables (nodes) y and z are to be determined in the following two cases. **(4)**



Case 1: Only forward checking is done after each variable assignment.

$\mathscr{D}_y = \underline{\quad \{\text{ R, B }\} \quad}$ $\qquad$ $\mathscr{D}_z = \underline{\quad \{\text{ R, G, B }\} \quad}$

Case 2: After forward checking, constraint propagation is done using the AC-3 algorithm for maintaining arc consistency.

$\mathscr{D}_y = \underline{\quad \{\text{R}\} \quad}$ $\qquad$ $\mathscr{D}_z = \underline{\quad \emptyset \quad}$

(c) What are the resolutions of the following pairs of clauses involving atomic propositions p, q, r, s? Fill in the blanks with the most simplified formulas. **(4)**

$\text{Res}_p(p \vee \neg q \vee r,\ \neg p \vee s) = \underline{\quad \neg q \vee r \vee s \quad}$

$\text{Res}_q(p \vee \neg q \vee r,\ q \vee r) = \underline{\quad p \vee r \quad}$

$\text{Res}_r(p \vee \neg q \vee r,\ q \vee \neg r \vee s) = \underline{\quad \text{T} \quad}$

$\text{Res}_{\neg s}(s,\ \neg s) = \underline{\quad \text{F} \quad}$

(d) Let S be a subset of the set $\mathbb{R}$ of real numbers. Encode the following two statements into predicate-logic formulas. You may use standard mathematical symbols.

Statement 1: S is an open set if and only if every $x \in S$ is contained in an open interval $(x - \varepsilon_x, x + \varepsilon_x) \subseteq S$ for some $\varepsilon_x > 0$. **(2)**

$$\forall S \left[\text{open}(S) \iff \forall x \left[(x \in S) \Rightarrow \exists \varepsilon \left[(\varepsilon > 0) \wedge \left((x - \varepsilon, x + \varepsilon) \subseteq S \right) \right] \right] \right]$$

Statement 2: S is a closed set if and only if its complement in $\mathbb{R}$ is an open set. **(1)**

$$\forall S \left[\text{closed}(S) \iff \text{open}(\mathbb{R} - S) \right]$$

Is the empty set $\emptyset$ an open set as per Statement 1? Write Yes/No only. Yes **(1)**

(e) Consider a planning problem as follows. There are three integer variables x, y, z. Initially, we have $x = 1$, $y = 2$, and $z = 3$. The goal is to achieve $x = 5$, $y = 7$, and $z = 9$. At each step, we can increment exactly one of x, y, z by 1 or 2. At all points of time, we must maintain $x < y < z$. We use a state-space planner P to solve the problem.

First, assume that P is a progressive (forward) planner. What are the possible first moves of the planner? **(2)**

$(1,2,3)$ may become $(1,2,4)$ or $(1,2,5)$

Next, assume that P is a regressive (backward) planner (which does not use variables). What are the possible first moves of the planner? **(2)**

$(5,7,9)$ may come from $(3,7,9)$ or $(4,7,9)$ or $(5,6,9)$ or $(5,7,8)$

2. [Resolution-refutation proof in predicate logic]

Let $A(x), B(x), C(x), D(x)$ be predicates with the values of x coming from a domain $\mathscr{D}_x$. The knowledge base (KB) consists of the following quantified statements.

$$[KB1] \qquad \forall x \left[A(x) \iff B(x)\right]$$

$$[KB2] \qquad \forall x \left[\big(B(x) \wedge C(x)\big) \Rightarrow D(x)\right]$$

$$[KB3] \qquad \exists x \left[C(x) \wedge \neg D(x)\right]$$

You want to prove the following statement (query) from the knowledge base.

$$[Q] \qquad \exists x \left[\neg A(x)\right]$$

You need to prove the query using the method of resolution-refutation. No other proof method is acceptable. Proceed exactly as directed in the following parts. Do not overdo any step in any of the parts.

(a) How do you convert the query for the proof? Do not simplify in this part. **(2)**

Solution Take the negation of Q.

$$[\neg Q] \qquad \neg\exists x \left[\neg A(x)\right]$$

(b) Eliminate all implications. Again, do not simplify. **(3)**

Solution $KB1$ and $KB2$ involve implications, and can be rewritten as follows.

$$[KB1] \qquad \forall x \left[\big(A(x) \Rightarrow B(x)\big) \wedge \big(B(x) \Rightarrow A(x)\big)\right] \equiv \forall x \left[\big(\neg A(x) \vee B(x)\big) \wedge \big(A(x) \vee \neg B(x)\big)\right]$$

$$[KB2] \qquad \forall x \left[\neg\big(B(x) \wedge C(x)\big) \vee D(x)\right]$$

(c) Move all the $\neg$ signs to the innermost scopes. **(2)**

Solution $KB2$ and $\neg Q$ are affected as follows.

$$[KB2] \qquad \forall x \left[\big(\neg B(x) \vee \neg C(x)\big) \vee D(x)\right]$$

$$[\neg Q] \qquad \forall x \left[A(x)\right]$$

(d) Perform Skolemization. **(2)**

Solution Existential quantifier appears only in $KB3$. Replacing x by a Skolem constant ξ gives the following.

$$[KB3] \qquad C(\xi) \wedge \neg D(\xi)$$

(e) Convert to the prenex normal form. Rename variables as required, and write the entire formula with all the quantifiers at the beginning. **(2)**

Solution In the current context, this step only involves renaming of variables.

$$\begin{array}{rl}
[KB1] & \forall x \left[\big(\neg A(x) \vee B(x)\big) \wedge \big(A(x) \vee \neg B(x)\big) \right] \\
[KB2] & \forall u \left[\big(\neg B(u) \vee \neg C(u)\big) \vee D(u) \right] \\
[KB3] & C(\xi) \wedge \neg D(\xi) \\
[\neg Q] & \forall z \left[A(z) \right]
\end{array}$$

The entire formula is now as follows.

$$\forall x \forall u \forall z \left[\Big(\big(\neg A(x) \vee B(x)\big) \wedge \big(A(x) \vee \neg B(x)\big) \Big) \wedge \Big(\big(\neg B(u) \vee \neg C(u)\big) \vee D(u) \Big) \wedge C(\xi) \wedge \neg D(\xi) \wedge \Big(A(z) \Big) \right]$$

(f) Convert the formula of Part (e) to the conjunctive normal form (CNF) with the quantifiers removed. Clearly write down all the clauses. Name the clauses as $C1, C2, C3, \ldots$. Make sure that each clause has its own variable(s). **(3)**

Solution Introducing new variables gives us the following clauses.

$[C1]$	$\neg A(x) \vee B(x)$
$[C2]$	$A(y) \vee \neg B(y)$
$[C3]$	$\neg B(u) \vee \neg C(u) \vee D(u)$
$[C4]$	$C(\xi)$
$[C5]$	$\neg D(\xi)$
$[C6]$	$A(z)$

(g) Use a sequence of resolutions to arrive at the empty clause. Clearly mention the variable substitutions you make in each resolution step. **(4)**

Solution Resolve $C1$ with $C6$ under the variable substitution x/z to get:

$[C7]$ $B(z)$

Resolve $C3$ with $C7$ under the variable substitution u/z to get:

$[C8]$ $\neg C(z) \wedge D(z)$

Resolve $C4$ with $C8$ under the variable substitution z/ξ to get:

$[C9]$ $D(\xi)$

Finally, resolve $C5$ and $C9$ (no variable substitution is needed) to get the empty clause.

3. [SAT planning]

You have an array $A = (Y, B, R, G)$. You can swap any two distinct elements (not necessarily consecutive). Your goal is to convert A to the array (R, G, B, Y) by making a sequence of swaps. You want to achieve this using as few swaps as possible. You invoke a SAT planner to perform your task. You need to supply to the planner a SAT formula that encodes your problem. Solve the following parts to prepare that formula. Use 1-based array indexing.

(a) Write the proposition(s) you will use along with its/their intended meaning(s), for representing a state. Do not encode action(s) in this part. **(2)**

Solution For all $x \in \{R, G, B, Y\}$, for all $i \in \{1, 2, 3, 4\}$, and for all time t, we use the proposition

$$\text{at}(x, i, t)$$

to indicate that x is at index i at time t.

(b) Encode the initial state using the proposition(s) of Part (a). **(4)**

Solution We have the following list of propositions to encode the start state.

$$\begin{array}{ccccccccc}
\neg\text{at}(R,1,0) & \wedge & \neg\text{at}(R,2,0) & \wedge & \text{at}(R,3,0) & \wedge & \neg\text{at}(R,4,0) & \wedge \\
\neg\text{at}(G,1,0) & \wedge & \neg\text{at}(G,2,0) & \wedge & \neg\text{at}(G,3,0) & \wedge & \text{at}(G,4,0) & \wedge \\
\neg\text{at}(B,1,0) & \wedge & \text{at}(B,2,0) & \wedge & \neg\text{at}(B,3,0) & \wedge & \neg\text{at}(B,4,0) & \wedge \\
\text{at}(Y,1,0) & \wedge & \neg\text{at}(Y,2,0) & \wedge & \neg\text{at}(Y,3,0) & \wedge & \neg\text{at}(Y,4,0) &
\end{array}$$

(c) Encode the final (that is, goal) state using the proposition(s) of Part (a). **(2)**

Solution Suppose that we want to solve the problem in N steps. Then, the final state is given by:

$$\text{at}(R,1,N) \ \wedge\ \text{at}(G,2,N) \ \wedge\ \text{at}(B,3,N) \ \wedge\ \text{at}(Y,4,N)$$

(d) Encode the swap operation for the SAT planner. **(4)**

Solution Introduce a proposition $\text{swap}(x,y,i,j,t)$ to indicate the action that the i-th and the j-th elements of A, containing x and y respectively, are swapped at time t, where $i,j \in \{1,2,3,4\}$ with $i < j$, $x,y \in \{R,G,B,Y\}$ with $x \neq y$, and $t \in \{0,1,2,\ldots,N-1\}$. This is encoded as follows.

$$\begin{array}{rcll} \text{swap}(x,y,i,j,t) & \Rightarrow & \text{at}(x,i,t) \ \wedge \ \text{at}(y,j,t) & \text{[Preconditions]} \\ & \wedge & \neg\text{at}(x,i,t+1) \ \wedge \ \neg\text{at}(y,j,t+1) & \text{[Delete list]} \\ & \wedge & \text{at}(x,j,t+1) \ \wedge \ \text{at}(y,i,t+1) & \text{[Add list]} \end{array}$$

For $i > j$, the proposition $\text{swap}(x,y,i,j,t)$ is the same as $\text{swap}(y,x,j,i,t)$.

(e) How do you ensure that after each step, an element is at a given index at a given time? Note, in particular, that if a swap takes place between indices i and j, then the elements at the other two indices do not change. **(2)**

Solution x is at location i at time $t+1$ if and only if one of the following things happens.

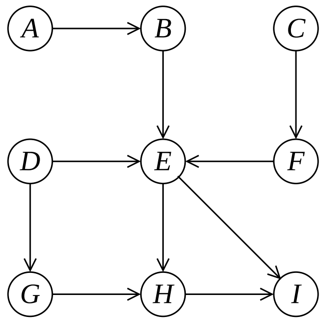
(1) A swap operation $\text{swap}(x,y,j,i,t)$ brings x to index i.
(2) A swap operation $\text{swap}(x,y,i,j,t)$ does not move x from position i.

This is encoded as:

$$\text{at}(x,i,t+1) \iff \left[\bigvee_{y\neq x, j\neq i} \text{swap}(x,y,j,i,t)\right] \bigvee \left[\text{at}(x,i,t) \bigwedge \neg\left(\bigvee_{y\neq x, j\neq i} \text{swap}(x,y,i,j,t)\right)\right]$$

4. [Probabilistic reasoning]

(a) Consider the Bayesian network shown to the right. Answer whether the following statements are True / False, clearly showing the justification by applying the D-separation method. **(6)**



(i) $A \perp\!\!\!\perp C$, that is, "A is conditionally independent of C given nothing"

(ii) $A \perp\!\!\!\perp I \mid E$, that is, "$A$ is conditionally independent of I given E"

(iii) $F \perp\!\!\!\perp A \mid H$, that is, "$F$ is conditionally independent of A given H"

(iv) $D \perp\!\!\!\perp I \mid \{E, G\}$, that is, "$D$ is conditionally independent of I given E, G"

Solution (i) True

No active paths present, because the path connecting A and C gets blocked due to the inactive triple $B \rightarrow E \leftarrow F$.

(ii) False

Among three possible paths, one is active:

- $A \rightarrow B \rightarrow E \rightarrow I$ is blocked due to inactive triple $B \rightarrow E$(observed) $\leftarrow I$;
- $A \rightarrow B \rightarrow E \rightarrow H \rightarrow I$ is blocked due to inactive triple $B \rightarrow E$(observed) $\leftarrow H$;
- However, $A \rightarrow B \rightarrow E$ (observed) $\leftarrow D \rightarrow G \rightarrow H \rightarrow I$ is an active path.

(iii) False

An active path is as follows: $F \rightarrow E$ (descendant H observed) $\leftarrow B \leftarrow A$.

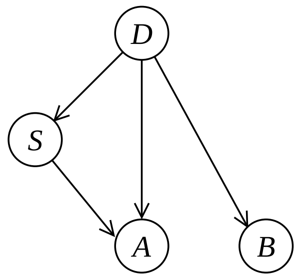
(iv) True

Among three possible paths, all are inactive:

- $D \rightarrow E$(observed) $\rightarrow I$ is a blocked or inactive triple;
- $D \rightarrow E \rightarrow H \rightarrow I$ is blocked due to inactive triple $D \rightarrow E$(observed) $\leftarrow H$;
- $D \rightarrow G$(observed) $\rightarrow H \rightarrow I$ is blocked due to inactive triple $D \rightarrow G$(observed) $\rightarrow H$;

(b) The following Bayesian network (right) has the following conditional probability distributions (left). Here, the variables D, S, A and B intuitively indicate having a Disease, having a Symptom, positive result on Test A, and positive result on Test B, respectively.

$$\begin{aligned} \Pr\left(A \mid D,S\right) &= 0.9, & \Pr\left(A \mid \neg D,S\right) &= 0.6,\\ \Pr\left(A \mid D,\neg S\right) &= 0.8, & \Pr\left(A \mid \neg D,\neg S\right) &= 0.1,\\ \Pr\left(S \mid D\right) &= 0.7, & \Pr\left(S \mid \neg D\right) &= 0.8,\\ \Pr\left(B \mid D\right) &= 0.7, & \Pr\left(B \mid \neg D\right) &= 0.5,\\ \Pr\left(D\right) &= 0.1. \end{aligned}$$



Estimate the following probability values. Show your calculations in detail. **(8)**

(i) What is the probability of having Disease D and getting a positive result on Test A?

(ii) What is the probability of *not* having Disease D and getting a positive result on Test A?

(iii) What is the probability of having Disease D given a positive result on Test A?

(iv) What is the probability of having Disease D given a positive result on Test B?

Solution (i)

$$\begin{aligned} \Pr\left(D,A\right) &= \Pr\left(D,S,A\right) + \Pr\left(D,\neg S,A\right)\\ &= \Pr\left(A \mid D,S\right)\cdot\Pr\left(S \mid D\right)\cdot\Pr\left(D\right) + \Pr\left(A \mid D,\neg S\right)\cdot\Pr\left(\neg S \mid D\right)\cdot\Pr\left(D\right)\\ &= \Pr\left(D\right)\cdot\left[\Pr\left(A \mid D,S\right)\cdot\Pr\left(S \mid D\right) + \Pr\left(A \mid D,\neg S\right)\cdot\left(1-\Pr\left(S \mid D\right)\right)\right]\\ &= 0.1\times\left[0.9\times0.7+0.8\times\left(1-0.7\right)\right] \quad = \quad 0.087 \end{aligned}$$

(ii)

$$\begin{aligned} \Pr\left(\neg D,A\right) &= \Pr\left(\neg D,S,A\right) + \Pr\left(\neg D,\neg S,A\right)\\ &= \Pr\left(A \mid \neg D,S\right)\cdot\Pr\left(S \mid \neg D\right)\cdot\Pr\left(\neg D\right) + \Pr\left(A \mid \neg D,\neg S\right)\cdot\Pr\left(\neg S \mid \neg D\right)\cdot\Pr\left(\neg D\right)\\ &= \left(1-\Pr\left(D\right)\right)\cdot\left[\Pr\left(A \mid \neg D,S\right)\cdot\Pr\left(S \mid \neg D\right) + \Pr\left(A \mid \neg D,\neg S\right)\cdot\left(1-\Pr\left(S \mid \neg D\right)\right)\right]\\ &= \left(1-0.1\right)\times\left[0.6\times0.8+0.1\times\left(1-0.8\right)\right] \quad = \quad 0.45 \end{aligned}$$

(iii)

$$\begin{aligned} \Pr\left(D \mid A\right) &= \frac{\Pr\left(D,A\right)}{\Pr\left(A\right)} = \frac{\Pr\left(D,A\right)}{\Pr\left(D,A\right)+\Pr\left(\neg D,A\right)}\\ &= \frac{0.087}{0.087+0.45} \approx 0.162 \end{aligned}$$

(iv)

$$\begin{aligned} \Pr\left(D \mid B\right) &= \frac{\Pr\left(D,B\right)}{\Pr\left(B\right)} = \frac{\Pr\left(B \mid D\right)\cdot\Pr\left(D\right)}{\Pr\left(B \mid D\right)\cdot\Pr\left(D\right)+\Pr\left(B \mid \neg D\right)\cdot\Pr\left(\neg D\right)}\\ &= \frac{0.7\times0.1}{0.7\times0.1+0.5\times(1-0.1)} \approx 0.135 \end{aligned}$$

5. **[Decision tree learning]**

We want to predict the outcome of the next cricket match between the two top-ranked teams: India and Australia. We are given the following dataset from recent matches. The outcome of a match is T if India wins, or F if Australia wins. Assume that a match cannot end in a draw.

Time	Match Type	Pitch Type	Outcome
Evening	T20	Fast	T
Day & Night	ODI	Dusty	T
Day	Test	Bouncy	T
Day & Night	Test	Neutral	F
Day & Night	T20	Dusty	F
Day & Night	ODI	Fast	T
Day & Night	ODI	Bouncy	T
Day & Night	ODI	Bouncy	T
Evening	T20	Fast	T
Day & Night	ODI	Dusty	F
Day	Test	Bouncy	F
Day	T20	Neutral	F
Day & Night	T20	Dusty	F
Day & Night	T20	Fast	T
Day & Night	ODI	Bouncy	T
Day & Night	ODI	Dusty	T

You are required to create a Decision Tree from this data, and use it to decide whether India is likely to win the next match against Australia. In particular, answer the following parts.

(a) Calculate and show the entropy of the overall Outcome. **(2)**

Solution Out of 16 possible outcomes, 10 outcomes are T and 6 outcomes are F.

Therefore, the entropy of the outcome is given as,

$$Entropy\,[\text{Outcome}] = E(10,6) = -\frac{10}{16}\log_2\left(\frac{10}{16}\right) - \frac{6}{16}\log_2\left(\frac{6}{16}\right) = 0.9544$$

(b) Which attribute should you choose at the root of your decision tree? Show the detailed calculations including information gains obtained for each attribute you choose. **(6)**

Solution If we consider the attribute "Time", we get:

$$\begin{aligned}
\text{For Time = `Evening'}: \quad E(2,0) &= 0 \\
\text{For Time = `Day \& Night'}: \quad E(7,4) &= -\frac{7}{11}\log_2\left(\frac{7}{11}\right) - \frac{4}{11}\log_2\left(\frac{4}{11}\right) = 0.946 \\
\text{For Time = `Day'}: \quad E(1,2) &= -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right) = 0.918
\end{aligned}$$

Therefore, $\quad Entropy\,[\text{Time}] = \frac{2}{16}E(2,0) + \frac{11}{16}E(7,4) + \frac{3}{16}E(1,2) = 0.8225$

Hence, $\quad IG\,[\text{Time}] = Entropy\,[\text{Outcome}] - Entropy\,[\text{Time}] = 0.9544 - 0.8225 = 0.132$

If we consider the attribute "Match Type", we get:

$$\begin{aligned}
\text{For Match Type = `T20'}: \quad E(3,3) &= 1 \\
\text{For Match Type = `ODI'}: \quad E(6,1) &= -\frac{6}{7}\log_2\left(\frac{6}{7}\right) - \frac{1}{7}\log_2\left(\frac{1}{7}\right) = 0.592 \\
\text{For Match Type = `Test'}: \quad E(1,2) &= -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right) = 0.918
\end{aligned}$$

Therefore, $\quad Entropy\,[\text{Match Type}] = \frac{6}{16}E(3,3) + \frac{7}{16}E(6,1) + \frac{3}{16}E(1,2) = 0.8061$

Hence, $\quad IG\,[\text{Match Type}] = Entropy\,[\text{Outcome}] - Entropy\,[\text{Match Type}] = 0.9544 - 0.8061 = 0.148$

If we consider the attribute "Pitch Type", we get:

$$\begin{aligned}
\text{For Pitch Type = `Fast'}: \quad E(4,0) &= 0 \\
\text{For Pitch Type = `Dusty'}: \quad E(2,3) &= -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right) = 0.971 \\
\text{For Pitch Type = `Bouncy'}: \quad E(4,1) &= -\frac{4}{5}\log_2\left(\frac{4}{5}\right) - \frac{1}{5}\log_2\left(\frac{1}{5}\right) = 0.722 \\
\text{For Pitch Type = `Neutral'}: \quad E(0,2) &= 0
\end{aligned}$$

Therefore, $\quad Entropy\,[\text{Pitch Type}] = \frac{4}{16}E(4,0) + \frac{5}{16}E(2,3) + \frac{5}{16}E(4,1) + \frac{2}{16}E(0,2) = 0.529$

Hence, $\quad IG\,[\text{Pitch Type}] = Entropy\,[\text{Outcome}] - Entropy\,[\text{Pitch Type}] = 0.9544 - 0.529 = 0.425$

We choose Pitch Type as the root attribute of the decision tree because it yields the maximum information gain.

(c) Determine the choice of next (second) level attributes leveraging the entropy-based information gain measures, and build (show) the approximate decision tree up to two levels, by selecting proper attributes at every level. **(6)**

Solution With Pitch Type = 'Dusty' ($E(2,3) = 0.971$),

- if we consider Match Type = 'ODI', we get $E(2,1) = 0.918$ and
- if we consider Match Type = 'T20', we get $E(0,2) = 0$.

Hence, $IG[Match\ Type] = 0.971 - \left(\frac{3}{5} \times 0.918 + \frac{2}{5} \times 0\right) = 0.4202$

On the other hand, since all Time = 'Day & Night' for Pitch Type = 'Dusty', so $IG[Time] = 0$

So, after Pitch Type = 'Dusty', we select Match Type attribute (giving higher IG).

With Pitch Type = 'Bouncy' ($E(4,1) = 0.722$),

- if we consider Match Type = 'ODI', we get $E(3,0) = 0$ and
- if we consider Match Type = 'Test', we get $E(1,1) = 1$.

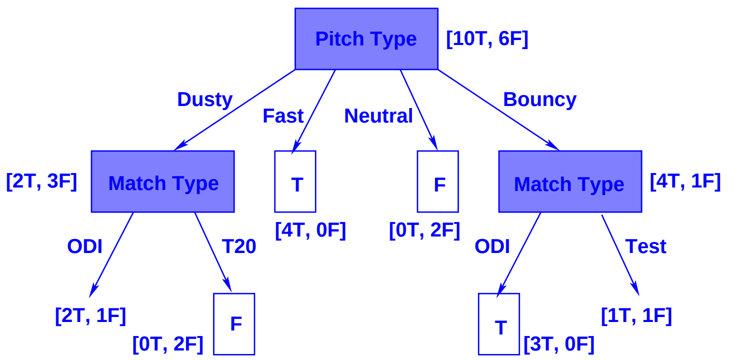
Hence, $IG[Match\ Type] = 0.722 - \left(\frac{3}{5} \times 0 + \frac{2}{5} \times 1\right) = 0.322$

With Pitch Type = 'Bouncy' ($E(4,1) = 0.722$),

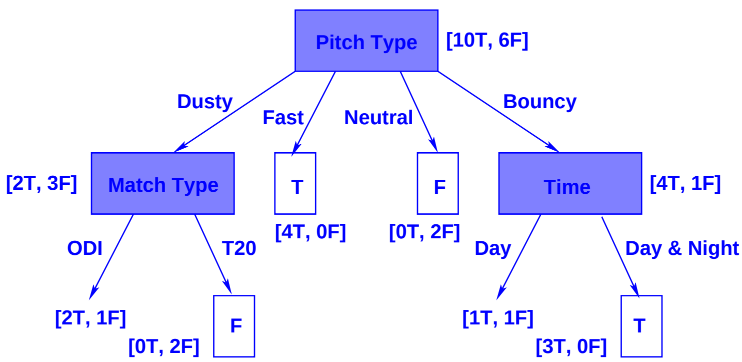
- if we consider Time = 'Day & Night', we get $E(3,0) = 0$ and
- if we consider Time = 'Day', we get $E(1,1) = 1$.

Hence, $IG[Time] = 0.722 - \left(\frac{3}{5} \times 0 + \frac{2}{5} \times 1\right) = 0.322$

So, after Pitch Type = 'Bouncy', we can select either Match Type or Time attribute (both giving same IG).



OR

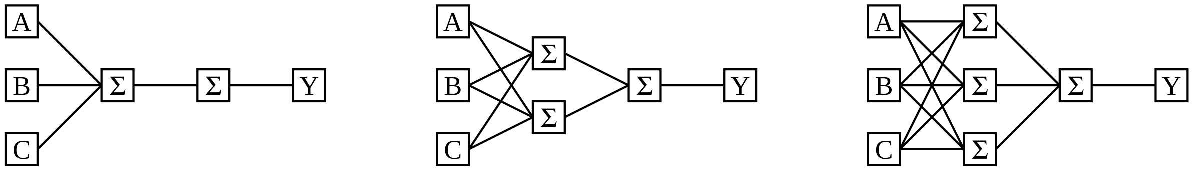


(d) We are given that the next match is an 'ODI' to be played as a 'Day & Night' match, and the pitch is 'Dusty'. Use the decision tree that you built above to decide whether India is likely to win the match. If the decision is not certain, then indicate the probability of win. **(2)**

Solution The decision here is not certain, and the probability of India winning is $\frac{2}{3}$.

6. [Neural networks]

(a) The following figure shows three multi-layer perceptrons PT_1, PT_2, and PT_3.



PT_1 $\qquad$ PT_2 $\qquad$ PT_3

Booleans (A, B, and C) take values 0 and 1, and each perceptron outputs values 0 and 1. The perceptron unit is denoted using [Σ] which outputs $h(X) = \text{sign}\left(\sum_i W_i.X_i\right)$ as usual. You may assume that each perceptron also has a mandatory input feature (threshold or bias) that always takes the value 1 having weight W_t (not explicitly shown in the figure above). Connection weights (W_i's) are allowed to take on any values.

Design the weight parameters so that you can realize the following Boolean functions using each of these three perceptrons (architectures are shown above) separately. In the case that some of these functions is not realizable by some of the above perceptrons, mention the same. **(9)**

(i) $A \wedge B$, (ii) $B \oplus C$, (iii) $A \rightarrow B$

Solution (i) The following weights can realize ($A \wedge B$):

PT_1 : $W_t^1 = -1.5, W_t^2 = -0.5, W_a = 1, W_b = 1, W_c = 0, W_s^1 = 1$

PT_2 : $W_t^{11} = W_t^{12} = -1.5, W_t^2 = -0.5, W_s^1 = W_s^2 = 1$

$W_a^1 = W_a^2 = 1, W_b^1 = W_b^2 = 1, W_c^1 = W_c^2 = 0$

PT_3 : $W_t^{11} = W_t^{12} = W_t^{13} = -1.5, W_t^2 = -0.5, W_s^1 = W_s^2 = W_s^3 = 1,$

$W_a^1 = W_a^2 = W_a^3 = 1, W_b^1 = W_b^2 = W_b^3 = 1, W_c^1 = W_c^2 = W_c^3 = 0$

(ii) The following weights can realize ($B \oplus C$):

Here, $B \oplus C = (B \wedge \neg C) \vee (\neg B \wedge C)$

PT_1 : Cannot realize this function!

PT_2 : $W_t^{11} = W_t^{12} = W_t^2 = -0.5, W_s^1 = W_s^2 = 1$

$W_a^1 = W_a^2 = 0, W_b^1 = 1, W_b^2 = -1, W_c^1 = -1, W_c^2 = 1$

PT_3 : $W_t^{11} = 0, W_t^{12} = W_t^{13} = W_t^2 = -0.5, W_s^1 = 0, W_s^2 = W_s^3 = 1,$

$W_a^1 = W_a^2 = W_a^3 = 0, W_b^1 = 0, W_b^2 = 1, W_b^3 = -1, W_c^1 = 0, W_c^2 = -1, W_c^3 = 1$

(iii) The following weights can realize ($A \rightarrow B$):

Here, $A \rightarrow B = \neg A \vee B$

PT_1 : $W_t^1 = 0.5, W_t^2 = -0.5, W_a = -1, W_b = 1, W_c = 0, W_s^1 = 1$

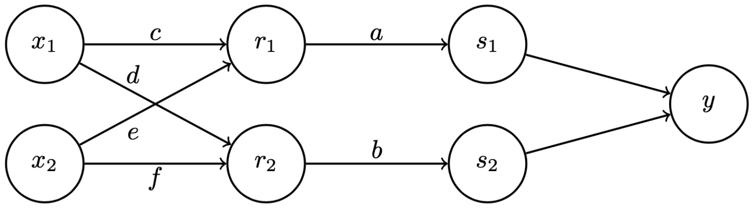
PT_2 : $W_t^{11} = W_t^{12} = 0.5, W_t^2 = -0.5, W_s^1 = W_s^2 = 1$

$W_a^1 = W_a^2 = -1, W_b^1 = W_b^2 = 1, W_c^1 = W_c^2 = 0$

PT_3 : $W_t^{11} = W_t^{12} = W_t^{13} = 0.5, W_t^2 = -0.5, W_s^1 = W_s^2 = W_s^3 = 1,$

$W_a^1 = W_a^2 = W_a^3 = -1, W_b^1 = W_b^2 = W_b^3 = 1, W_c^1 = W_c^2 = W_c^3 = 0$

(b) Below is a neural network with weights a, b, c, d, e, f. The inputs are x_1 and x_2. The first hidden layer computes $r_1 = \max(c \cdot x_1 + e \cdot x_2, 0)$ and $r_2 = \max(d \cdot x_1 + f \cdot x_2, 0)$. The second hidden layer computes $s_1 = \dfrac{1}{1+\exp(-a \cdot r_1)}$ and $s_2 = \dfrac{1}{1+\exp(-b \cdot r_2)}$. The output layer computes $y = s_1 + s_2$. Note that the weights a, b, c, d, e, f are indicated along the edges of the neural network here.



Suppose that the network has inputs $x_1 = 1$ and $x_2 = -1$. The weight values are $a = 1$, $b = 1$, $c = 4$, $d = 1$, $e = 2$, and $f = 2$. Forward propagation then computes $r_1 = 2$, $r_2 = 0$, $s_1 = 0.9$, $s_2 = 0.5$, $y = 1.4$. (Note: some values are rounded.)

Using the values computed from forward propagation, use back propagation to numerically calculate the following partial derivatives. Show your detailed calculations. **(9)**

(i) $\dfrac{\partial y}{\partial a}$, (ii) $\dfrac{\partial y}{\partial b}$, (iii) $\dfrac{\partial y}{\partial c}$, (iv) $\dfrac{\partial y}{\partial d}$, (v) $\dfrac{\partial y}{\partial e}$, (vi) $\dfrac{\partial y}{\partial f}$

Hint: For $g(z) = \dfrac{1}{1+\exp(-z)}$, the derivative is $\dfrac{\partial g}{\partial z} = g(z) \cdot \big(1 - g(z)\big)$

Solution (i) $\dfrac{\partial y}{\partial a} = \dfrac{\partial y}{\partial s_1} \cdot \dfrac{\partial s_1}{\partial a} = 1 \cdot \dfrac{\partial g(a \cdot r_1)}{\partial a} = r_1 \cdot g(a \cdot r_1)\big(1 - g(a \cdot r_1)\big) = r_1 \cdot s_1\big(1 - s_1\big) = 2 \times 0.9 \times (1 - 0.9) = 0.18$

(ii) $\dfrac{\partial y}{\partial b} = \dfrac{\partial y}{\partial s_2} \cdot \dfrac{\partial s_2}{\partial b} = 1 \cdot \dfrac{\partial g(b \cdot r_2)}{\partial b} = r_2 \cdot g(b \cdot r_2)\big(1 - g(b \cdot r_2)\big) = r_2 \cdot s_2\big(1 - s_2\big) = 0 \times 0.5 \times (1 - 0.5) = 0$

(iii) $\dfrac{\partial y}{\partial c} = \dfrac{\partial y}{\partial s_1} \cdot \dfrac{\partial s_1}{\partial r_1} \cdot \dfrac{\partial r_1}{\partial c} = 1 \cdot \Big[a \cdot g(a \cdot r_1)\big(1 - g(a \cdot r_1)\big)\Big] \cdot x_1 = \Big[a \cdot s_1\big(1 - s_1\big)\Big] \cdot x_1 =$
$\Big[1 \times 0.9 \times (1 - 0.9)\Big] \times 1 = 0.09$

(iv) $\dfrac{\partial y}{\partial d} = \dfrac{\partial y}{\partial s_2} \cdot \dfrac{\partial s_2}{\partial r_2} \cdot \dfrac{\partial r_2}{\partial d} = \dfrac{\partial y}{\partial s_2} \cdot \dfrac{\partial s_2}{\partial r_2} \cdot 0 = 0$

(v) $\dfrac{\partial y}{\partial e} = \dfrac{\partial y}{\partial s_1} \cdot \dfrac{\partial s_1}{\partial r_1} \cdot \dfrac{\partial r_1}{\partial e} = 1 \cdot \Big[a \cdot g(a \cdot r_1)\big(1 - g(a \cdot r_1)\big)\Big] \cdot x_2 = \Big[a \cdot s_1\big(1 - s_1\big)\Big] \cdot x_2 =$
$\Big[1 \times 0.9 \times (1 - 0.9)\Big] \times (-1) = -0.09$

(vi) $\dfrac{\partial y}{\partial f} = \dfrac{\partial y}{\partial s_2} \cdot \dfrac{\partial s_2}{\partial r_2} \cdot \dfrac{\partial r_2}{\partial f} = \dfrac{\partial y}{\partial s_2} \cdot \dfrac{\partial s_2}{\partial r_2} \cdot 0 = 0$