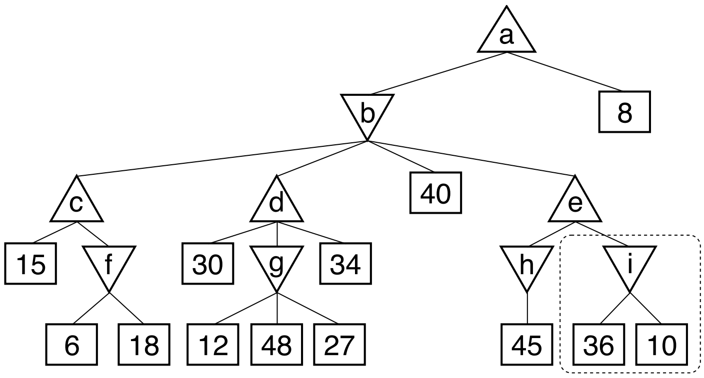
COMP7015 Artificial Intelligence **Semester 1, 2024–25**

Additional Exercise

Instructor: Dr. Kejing Yin Oct. 10, 2024

Exercise 1: Adversarial Search

Given the following game tree, where the upward-facing and downward-facing triangles represent max and min nodes, respectively. The rectangles represent terminal states with their utilities written inside the rectangles.



(a) Compute the minimax values $a, b, ..., i$.
$a = 15, b = 15, c = 15, d = 34, e = 45, f = 6, g = 12, h = 45, i = 10.$

(b) Suppose we use alpha-beta pruning, can we prune the subtree i as shown in the dotted rectangle? Briefly explain the reasons.
Yes, we can. From the tree, we know that when evaluating e, $\beta = 15$. Since $h = 45 > \beta$, the value of i will not affect the value of b. Thus i can be pruned.

(c) Suppose we have a game that is much more complicated than the tree shown above and it becomes infeasible to construct the complete tree (like Go). Can you suggest an algorithm that can find a good policy?
We can use Monte Carlo Tree Search (MCTS).

Exercise 2: Logics

(a) Given P is true and Q is false, determine the truth value for the following propositions.

1. $\neg Q \wedge P$ true
2. $Q \vee \neg\neg P$ true
3. $(P \wedge Q) \vee Q \Rightarrow \neg P \vee Q$ true
4. $\neg P \vee (P \wedge Q) \Leftrightarrow P \vee \neg Q$ false
5. $\neg(\neg Q \Rightarrow \neg P) \wedge (P \Leftrightarrow Q)$ false

(b) Are the following correct? Justify your answer.

1. $(A \vee B) \wedge \neg(A \Rightarrow B)$ is satisfiable.

   The statement is correct.

For variables A and B, there are four possible models: $m_1 = \{A = 1, B = 1\}$, $m_2 = \{A = 1, B = 0\}$, $m_3 = \{A = 0, B = 1\}$, and $m_4 = \{A = 0, B = 0\}$, where we use 1 to represent true and 0 to represent false. We can examine the four models one by one.

i) $A = 1$ and $B = 1$, $(A \vee B) \wedge \neg(A \Rightarrow B)$ is false. Thus, m_1 does not satisfy this formula.
ii) $A = 1$ and $B = 0$, $(A \vee B) \wedge \neg(A \Rightarrow B)$ is true. Thus, m_2 satisfies this formula.
iii) $A = 0$ and $B = 1$, $(A \vee B) \wedge \neg(A \Rightarrow B)$ is false. Thus, m_3 does not satisfy this formula.
iv) $A = 0$ and $B = 0$, $(A \vee B) \wedge \neg(A \Rightarrow B)$ is false. Thus, m_4 does not satisfy this formula.

The set of all models of the given formula is $M\{(A \vee B) \wedge \neg(A \Rightarrow B)\} \neq \emptyset$. Therefore, the formula is satisfiable.

2. $(A \vee B) \vDash (A \Leftrightarrow B)$.

   The statement is not correct.

   For variables A and B, there are four possible models: $m_1 = \{A = 1, B = 1\}$, $m_2 = \{A = 1, B = 0\}$, $m_3 = \{A = 0, B = 1\}$, and $m_4 = \{A = 0, B = 0\}$, where we use 1 to represent true and 0 to represent false. Similar to the steps above, we can obtain that:

   i) $M(A \vee B) = \{m_1, m_2, m_3\}$
   ii) $M(A \Leftrightarrow B) = \{m_1, m_4\}$

   Obviously, $M(A \vee B) \nsubseteq M(A \Leftrightarrow B)$. Therefore, $(A \vee B)$ does not entail $(A \Leftrightarrow B)$.

(c) Write a first-order logic formula for each of the following English sentence.

1. There exists a student that can play Tennis.

   $\exists x \quad \text{IsStudent}(x) \wedge \text{CanPlayTennis}(x)$

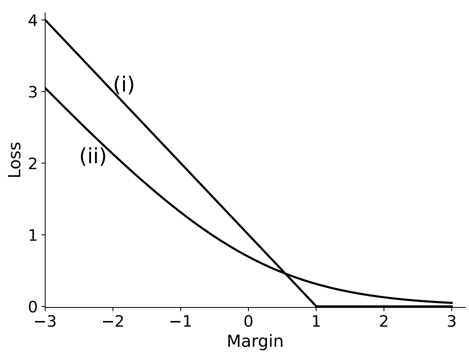
2. All students need to take a course that covers AI.

   $\forall x \quad \text{IsStudent}(x) \Rightarrow (\exists y \ \text{IsCourse}(y) \wedge \text{TakeCourse}(x, y) \wedge \text{CoversAI}(y))$

3. There are some students who take all courses that cover AI.

   $\exists x \ \text{IsStudent}(x) \wedge (\forall y \quad \text{IsCourse}(y) \wedge \text{CoversAI}(y) \Rightarrow \text{TakeCourse}(x, y))$

Exercise 3: Machine Learning Basics



Visualization of two loss functions

The dataset

x_1	x_2	x_3	x_4	Labels
6.3	2.5	4.9	1.5	+1
5.4	3.4	1.7	0.2	-1
4.6	3.6	1.0	0.2	+1
6.5	2.8	4.6	1.5	+1
6.0	2.7	5.1	1.6	-1
5.6	3.0	2.6	1.5	+1
5.4	3.7	1.5	0.2	+1
5.0	3.6	1.4	0.2	-1

(a) The above figure on the left visualizes two loss functions for binary classification. Give the names and formulas of (i) and (ii), respectively. Also, briefly discuss the merits of them.

The loss function (i) is hinge loss and (ii) is the logistic loss. They are given by:

(i) Hinge loss: $\text{Loss}_{\text{hinge}}(\mathbf{x}, y, \mathbf{w}) = \max\left\{1 - \left(\mathbf{w}^\top \mathbf{x}\right) y, 0\right\}$;

(ii) Logistic loss: $\text{Loss}_{\text{logistic}}(\mathbf{x}, y, \mathbf{w}) = \log\left(1 + \exp\left(-\left(\mathbf{w}^\top \mathbf{x}\right) y\right)\right)$.

The merits of them (any one of the following, or other reasonable answer will be correct):

- They both have non-zero gradient when the model make wrong predictions, so the gradient descent algorithm can be used.
- The hinge loss will try to increase the margin even when a sample is correctly classified. It tries to make the model classify all points correctly and confidently.
- The logistic loss always have non-zero loss and tries to increase margin even when it already exceeds 1.

(b) Given the dataset as shown above on the right. Consider a linear model with the weight vector and bias given by:

$$\mathbf{w} = [0.4, \ -0.7, \ 2.0, \ 0.8]^\top \qquad b = -6.$$

Compute the scores, margins, and the three loss function values for each data point and fill in the missing cells in the table below.

Features				Label	Score	Margin	One-zero loss	Hinge loss	Logistic loss
x_1	x_2	x_3	x_4						
6.3	2.5	4.9	1.5	+1	5.77	5.77	0	0	0
5.4	3.4	1.7	0.2	-1	-2.66	2.66	0	0	0.07
4.6	3.6	1.0	0.2	+1	-4.52	-4.52	1	5.52	4.53
6.5	2.8	4.6	1.5	+1	5.04	5.04	0	0	0
6.0	2.7	5.1	1.6	-1	5.99	-5.99	1	6.99	5.99
5.6	3.0	2.6	1.5	+1	0.54	0.54	0	0.46	0.46
5.4	3.7	1.5	0.2	+1	-3.27	-3.27	1	4.27	3.31
5.0	3.6	1.4	0.2	-1	-3.56	3.56	0	0	0.03

(c) Can we use zero-one loss function under the loss minimization framework with gradient descent algorithm? Briefly explain the reasons.

No, since the gradient of zero-one loss function is zero almost everywhere. Therefore, the parameters cannot be updated using gradient descent with zero-one loss function.

(d) Suppose we use $\hat{y} = \text{sign}(\mathbf{w}^\top \mathbf{x} + b)$ as predicted labels. Plot the confusion matrix, compute the accuracy, precision, recall, and F1 scores.

(1) The confusion matrix is as follows:

	Predicted Positive	Predicted Negative
Ground-truth Positive	TP=3	FN=2
Ground-truth Negative	FP=1	TN=2

(2) The accuracy is given by:

$$\text{Accuracy} = \frac{TP + TN}{N} = \frac{5}{8} = 0.625$$

(3) The precision is given by:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{3}{4} = 0.75$$

(4) The recall is given by:

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{3}{5} = 0.6$$

(5) The F1 score is given by:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.75 \times 0.6}{0.75 + 0.6} = 0.67$$