UNIVERSITY OF MANITOBA
Midterm

Winter 2005

COMPUTER SCIENCE

Machine Learning

Date: **Friday, 10 March 2005**
Time: **15:30 - 16:20**
Room: **EITC E2-165, University of Manitoba**
(Time allowed: 50 Minutes)

NOTE:
Attempt all questions.
This is a *closed* book examination.
Use of non-programmable calculators is *permitted*.
Use of any other electronic equipment is strictly forbidden.
Show your work to receive full marks.

SURNAME:

FORENAME(S):

STUDENT ID:

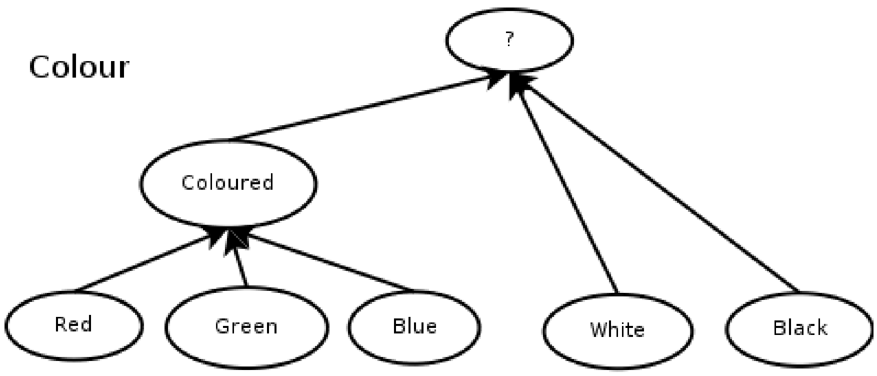
A 25	B 25	C 25	D 25	Total 100

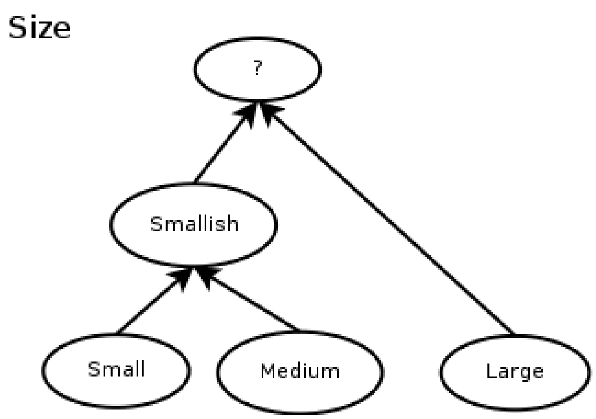
Surname: ____________________

Forename(s): ____________________

Section A: Candidate Elimination

1. Given is a clothing domain D. What is the maximum size of the G-set that can be created with a single example. Show the resulting S and G-sets and the instance that results in the largest possible G-set. Simply select one of the possible instances if there is more than one such instance.





[10 marks]

Surname: ____________________

Forename(s): ____________________

The maximum size of the G-set that can be created by a single negative example is 6

Instance: <Red,Small> -

Resulting S-set: <empty>

Resulting G-set:<Green,?>,<Blue,?>,<White,?>,<Black,?>,<?,Medium>,<?,Large>

---

2. Given is the clothing domain D shown above. What is the maximum size of the S-set that can be created with a single example. Show the resulting S and G-sets and the instance as well as its classification that results in the largest possible S-set. Simply select one of the possible instances if there is more than one such instance.

[5 marks]

The maximum size of the S-set that can be created by a single negative example is 1

Instance: <red,small> +

Resulting G-set:<?,?>

Surname: ______________________________

Forename(s): ______________________________

3. Show a trace of the execution of the Candidate Elimination algorithm on the clothing domain shown above given the following training sequence. For each instance, show the resulting S and G sets. I already did the first entry for you.

[10 marks]

```
<red,small>     +  S-Set = <red,small>
                   G-Set = <?,?>
<white,medium>  -  S-Set = <red,small>
                   G-Set = <coloured,?>,<?,small>
<black,medium>  -  S-Set = <red,small>
                   G-Set = <coloured,?>,<?,small>
<blue,medium>   +  S-Set = <coloured,smallish>
                   G-Set = <coloured,?>
<green,medium>  -  S-Set = Collapses to nil
                   G-Set = nil
<black,large>   -  S-Set = nil
                   G-Set = nil
```

Surname: ______________________________

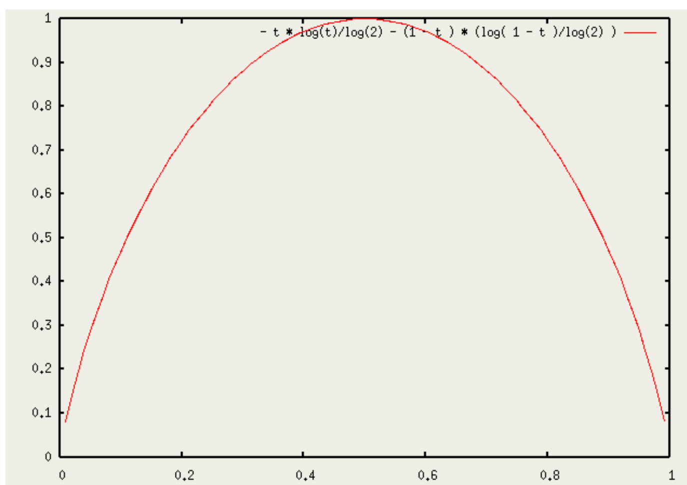
Forename(s): ______________________________

Section B: Decision Trees

The information gain Gain(S,A) of an attribute A for a sample set S is defined as

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v)$$

A graph of the entropy function is shown in the figure below. You can use this graph when answering the following questions.



Surname: ______________________________

Forename(s): ______________________________

4. Given below is a set of instances from a plant diagnosis domain with two attributes plant and type and whether the product was defective within the first month.

   Given the set of instances shown below, calculate the information gain for the attributes Plant and Type.

Instance	Plant	Type	Defect
x1	USA	A	Yes
x2	CAN	B	No
x3	SIN	A	No
x4	SIN	B	No
x5	USA	A	Yes
x6	USA	B	No
x7	SIN	A	No
x8	CAN	B	No
x9	USA	B	No
x10	SIN	A	No

[10 marks]

Information Gain(Plant) = 0.32

Information Gain(Type) = 0.23

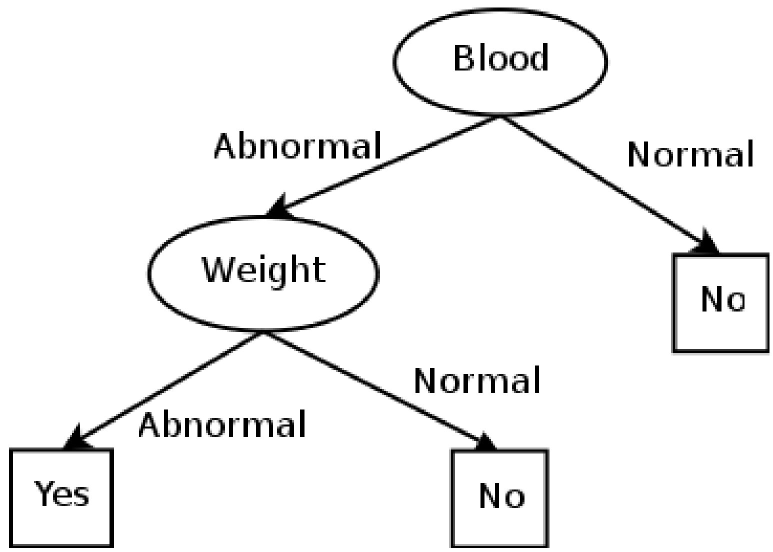
```
All: ent(2,8) = 0.72
Attribute Plant:
All: 0.72 -
Plant = USA: 4/10 * ent( 2, 2 ) = 0.4 -
Plant = SIN: 4/10 * ent( 0, 4 ) = 0.0 -
Plant = CAN: 2/10 * ent( 0, 2 ) = 0.0 = 0.32

All: 0.72 -
Type = A: 5/10 * ent( 2, 3) = 0.49
Type = B: 5/10 * ent( 0, 5) = 0.0
= 0.23
```

Surname: ____________________

Forename(s): ____________________

5. Given is a medical diagnosis domain with three attributes Blood, Weight, and Height. Each attribute has two possible values Normal and Abnormal.

   Given the instance set shown in the answerbox, fill in the missing classifications (Yes,No) so that the ID3 algorithm without post pruning will return the decision tree shown below.



   If it is impossible to fill in the missing classifications in such a way that the ID3 algorithm returns the decision tree shown above, then say so in your answer and explain why.

[15 marks]

Instance	Blood	Weight	Height	Diagnosis
x1	Normal	Abnormal	Abnormal	No
x2	Abnormal	Normal	Abnormal	No
x3	Abnormal	Abnormal	Normal	Yes
x4	Abnormal	Normal	Abnormal	No

It is impossible that the ID3 algorithm would have generated this tree, since the necessary assignments imply that the information gain of the attribute Height would have had a higher information gain.

Surname: ______________________________

Forename(s): ______________________________

Section C: Neural Nets

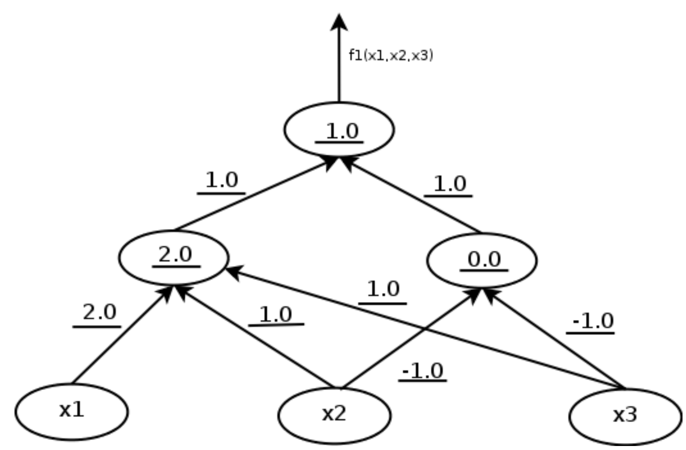
6. Given below is an artificial neural network (ANN) with three input nodes (X1,X2,X3), two hidden nodes, and one output node.

   The network uses simple threshold nodes (i.e., the node will output 1.0 if the sum of the weighted inputs is greater than or equal to the threshold, 0 otherwise).

   You are trying to learn the boolean target function f_1.

   Show a set of weights and thresholds for all nodes that implement the function f_1. If it is impossible to represent the function f_1 with the given neural network, then state this in your answer and explain why.

   [15 marks]



x_1	x_2	x_3	f_1
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

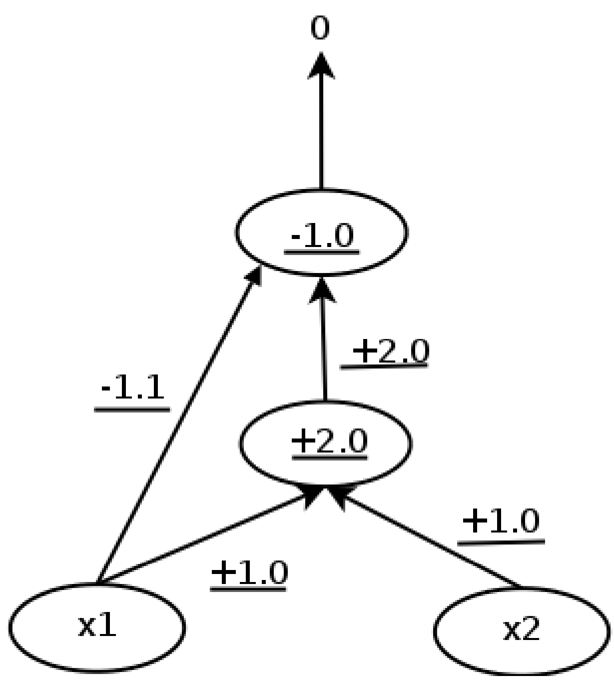
Surname: ______________________________

Forename(s): ______________________________

7. Given below is a small neural network with three nodes and two inputs $X1$ and $X2$. The network uses simple threshold neurons, that is a neuron will generate an output of 1, if the sum of the weighted inputs is greater than or equal to the threshold, 0 otherwise.

   Show one set of inputs $X1$ and $X2$ that will generate an output of 0 on the given network.

   If it is impossible to generate an output of 0 with the given network than say so in your answer and explain why.



[10 marks]

To generate an output of 0, the inputs have to be as follows:
$X1$=1

$X2$=0

Surname: ______________________________

Forename(s): ______________________________

Section D: Bayesian Learning

8. A robot has an infrared and an ultrasound sensor which help it to detect doors. The probabilites for the sensor readings and the existence of the door are shown below.

Infrared	Ultrasound	Door	Probability
N	N	N	0.2
N	N	Y	0.0
N	Y	N	0.1
N	Y	Y	0.2
Y	N	N	0.2
Y	N	Y	0.1
Y	Y	N	0.1
Y	Y	Y	0.1

Calculate the accuracy of the infrared sensor, in other words P(Infrared | Door) + P(˜Infrared | ˜Door) and the accuracy of the ultrasound sensor P(Ultrasound | Door) + P(˜Ultrasound | ˜Door) and compare them.

If it is impossible to compute and compare the accuracy of the two sensors, then say so in your answer and explain why.

[10 marks]

The accuracy of the infrared sensor is: <u>50%</u>
0.2 + 0.1 + 0.1 + 0.1 = 0.5
The accuracy of the ultrasound sensor is: <u>70%</u>
0.2 + 0.2 + 0.2 + 0.1 = 0.7

Surname: ____________________

Forename(s): ____________________

9. Given that the infrared sensor detects a door, and that the ultrasound sensor does not detect a door, calculate the probability that a door exists, that is calculate the posterior probability P(Door | Infrared = Yes and Ultrasound = No).

[10 marks]

> P(Door | Infrared = Yes and Ultrasound = No) = <u>33%</u>
> 0.1/(0.1+0.2)

10. Calculate the prior probability of the existence of a door.

[5 marks]

> P(Door) = 0.2 + 0.1 + 0.1 = 0.4

Surname: ____________________

Forename(s): ____________________

Additional work pages

Surname: ____________________

Forename(s): ____________________

Additional work pages