R228/406

DUBLIN INSTITUTE OF TECHNOLOGY
KEVIN STREET, DUBLIN 8

BSc (Hons) in Computer Science

Stage 4

SUPPLEMENTAL EXAMINATIONS 2010

ARTIFICIAL INTELLIGENCE 2

Dr. John Kelleher
Dr. D. Lillis
Dr. I. Arana

Duration: 2 Hours

Answer Question 1 (40 marks) **and**

any 2 Other Questions (30 marks each).

1. (a) Given the full joint distribution shown in Table 1, calculate the following:

Table 1: Full joint distribution for a dentist visit

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

(i) $P(toothache)$

(5 marks)

(ii) $\mathbf{P}(Cavity)$

(5 marks)

(iii) $\mathbf{P}(Toothache|cavity)$

(5 marks)

(iv) $\mathbf{P}(Cavity|toothache \vee catch)$

(5 marks)

(b) Describe the problems associated with measuring **classifier performance** using a single accuracy figure **and** describe a more appropriate alternative.

(20 marks)

2. (a) In you local power station, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the core of the power station. Consider the Boolean variables A (alarm sounds), F_A (alarm is faulty), and F_G (gauge is faulty); and multivalued nodes G (gauge reading) and T (actual core temperature).

   (i) Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.

   (5 marks)

   (ii) Suppose there are just two possible actual and measured temperatures, normal and high, and the probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with node G.

   (5 marks)

   (iii) Suppose the alarm works correctly unless it is faulty, in which case it never sounds. Give the conditional probability table associated with A.

   (5 marks)

   (b) Consider the following time keeping patterns of the lecturers in your college:

   - 25% of lecturers start 75% of their lectures on time and 25% late.
   - 50% of lecturers start 50% of their lectures on time and 50% late.
   - 25% of lecturers start 25% of their lectures on time and 75% late.

   (i) Given that both the 1^{st} and 2^{nd} Artificial Intelligence lectures of the year started on time, compute the posterior probability that your Artificial Intelligence lecturer follows that each of the three time-keeping patterns.

   (5 marks)

   (ii) Given that both the 1^{st} and 2^{nd} Artificial Intelligence lectures of the year started on time, what is the Bayesian Prediction that the 3^{rd} Artificial Intelligence lecture will start on time?

   (5 marks)

   (iii) Given that both the 1^{st} and 2^{nd} Artificial Intelligence lectures of the year started on time, what is the Maximum a Posterior (MAP) probability that the 3^{rd} Artificial Intelligence lecture will start on time?

   (5 marks)

3. (a) Define what is meant by **lazy learners** and **eager learners**, highlight the key differences between these approaches **and** give an example of each.

(10 marks)

(b) Formally define what is meant by the term **entropy** in the context of Decision Tree Learning.

(5 marks)

Table 2: The attributes of a newly discovered animal. A 1 indicates the animal possesses the feature listed in the column and 0 indicates that they do not. The column on the right contains a ? because the animal has not been classified yet.

Species	Births Live Young	Lays Eggs	Feeds Offspring Own Milk	Warm-Blooded	Cold-Blooded	Lives in Water and Land	Has Hair	Has Feathers	Class
Mystery	0	1	0	0	0	1	0	0	?

Table 3: Example feature vectors for animal classification. A 1 indicates the animal possesses the feature listed in the column and 0 indicates that they do not. The rightmost column lists the classification of each ainmal.

Species	Births Live Young	Lays Eggs	Feeds Offspring Own Milk	Warm-Blooded	Cold-Blooded	Lives in Water and Land	Has Hair	Has Feathers	Class
Cat	1	0	1	1	0	0	1	0	Mammal
Frog	0	1	0	0	1	1	0	0	Amphibian
Squirrel	1	0	1	1	0	0	1	0	Mammal
Duck	0	1	0	1	0	1	0	1	Bird

(c) You are working as an assistant-biologist to Charles Darwin on the Beagle voyage. You are at the Galápagos Islands and you have just discovered a new animal that has not yet been classified. Table 2 lists the attributes of the animal you have found. Mr. Darwin has asked you to classify the animal using a nearest-neighbour approach and he has supplied you with a case-base of already classified animals, see Table 3.

(i) A good measure of distance between two instances with categorical features is the number of features which have different values (the **overlap metric**, also known as the **hamming distance**). Using this measure of distance compute the distance between the mystery animal and each of the animals in the case base.

(5 marks)

(ii) If you used $1 - NN$ classification what class would be assigned to the mystery animal.

(5 marks)

(iii) If the you used $4 - NN$ classification what class would be assigned to the mystery animal.

(5 marks)

4. (a) Figure 1, on the next page, is a schematic of a 3 input perceptron. Input $a_0 = -1$, inputs a_1 and a_2 are binary. The perceptron uses a threshold activation function that outputs a 1 if the weighted sum of inputs is greater than 0 and a 0 otherwise. Define the truth-table of the function that this perceptron implements and identify the name of the function.

(5 marks)

(b) Describe the perceptron training rule?

(10 marks)

(c) Why does the perceptron training rule converge toward successful weight values?

(5 marks)

(d) The FOIL inductive logic programming algorithm is constructing a new rule with head $p(Y) \leftarrow$. Which of the following literals could be considered as candidate extensions $q(Y)$, $r(X)$, $s(X, Y)$, $\neg s(X, Y)$?

(5 marks)

(e) The FOIL inductive logic programming algorithm is considering adding a literal l to a rule r. The extension of the rule r before adding l is the following set of positive and negative examples $\{+, +, +, +, +, +, -, -, -, -\}$. The extension of the rule after the literal is added (i.e. the extension of $r + l$) is $\{+, +, +, +, +, +, -, -\}$. What is the information gain of adding the literal l to the rule r? (Note you do not need to calculate the logs in your answer (i.e. you can express your answer as an equation containing logs)).

(5 marks)

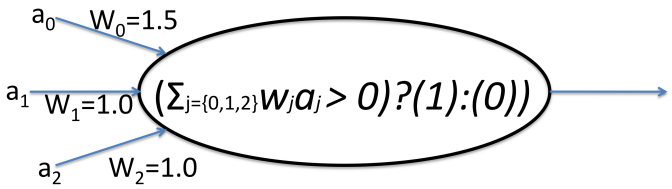

Figure 1: A 3 input perceptron. Input $a_0 = -1$, inputs a_1 and a_2 are binary. The perceptron uses a threshold activation function that outputs a 1 if the weighted sum of inputs is greater than 0 and a 0 otherwise.