## CS 540: Introduction to Artificial Intelligence

*Final Exam: 12:25-2:25pm, May 16, 2013*

*Room 228 Educational Sciences*

CLOSED BOOK

(two sheets of notes and a calculator allowed)

Write your answers on these pages and show your work. If you feel that a question is not fully specified, state any assumptions that you need to make in order to solve the problem. You may use the backs of these sheets for scratch work.

Write your name on this page and initial all other pages of this exam. Make sure your exam contains eight problems on ten pages.

**Name**  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Student ID** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Problem Score Max Score\_\_\_**

**1 \_\_\_\_\_\_ 15**

**2 \_\_\_\_\_\_ 16**

**3 \_\_\_\_\_\_ 15**

**4 \_\_\_\_\_\_ 15**

**5 \_\_\_\_\_\_ 10**

**6 \_\_\_\_\_\_ 10**

**7 \_\_\_\_\_\_ 10**

**8 \_\_\_\_\_\_ 9**

**TOTAL \_\_\_\_\_\_ 100**

**Problem 1 – Bayesian Networks (15 points)**

Consider the following Bayesian Network, where variables **A-D** are all Boolean-valued:

|  |  |
| --- | --- |
| **A** | **P(B =true | A)** |
| false | 0.9 |
| true | 0.2 |

**P(A=true)** = 0.7

# A

# B

# C

|  |  |
| --- | --- |
| A | P(C =true | A) |
| false | 0.6 |
| true | 0.3 |

|  |  |  |
| --- | --- | --- |
| B | C | P(D =true | B,C) |
| false | false | 0.4 |
| false | true | 0.5 |
| true | false | 0.8 |
| true | true | 0.1 |

# D

1. What is the probability that *A* and *B* are *true* but *C* and *D* are *false*? \_\_\_\_\_\_\_\_\_\_\_\_\_\_

[Be sure to show your work for Parts a-c.]

1. What is the probability that ***B*** is true, **A** is false, and ***D*** is true? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. What is the probability that ***B*** is true *given* that ***A*** is false and ***D*** is true? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Problem 2 – Probabilistic Reasoning and Learning (16 points)**

1. Consider the following training set, where all possible values of the three features appear in it. Assume you wish to apply the Naïve Bayes algorithm.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ex #** | **A** | **B** | **C** | **Output** |
| 1 | 2 | Small | Red | True |
| 2 | 1 | Small | Green | True |
| 3 | 3 | Large | Blue | False |
| 4 | 2 | Small | Blue | False |
| 5 | 4 | Large | Green | False |

Calculate the ratio below, showing your work below it and putting your (numeric) answer on the line to the right of the equal sign. Be sure to explain how you performed Laplace smoothing.

Prob(Output = True | A = 2, B = Large, C = Green)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_

Prob(Output = False | A = 2, B = Large, C = Green)

What is the most likely output for this example (A=2, B=Large, and C=Green)? \_\_\_\_\_\_\_\_\_

Briefly explain your answer below.

1. Assume we wish to use a *full joint probability table* for this task*.*

How many cells are needed if represented as a single multi-dimensional Java array? Briefly explain your answer.

Answer: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Draw the full joint probability table as a Bayes Net. Show the conditional probability table associated with each node, but you do not need to fill the cells with actual numbers (i.e., just leave them blank).

Here again is the dataset from Part *a*, copied for your convenience:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ex #** | **A** | **B** | **C** | **Output** |
| 1 | 2 | Small | Red | True |
| 2 | 1 | Small | Green | True |
| 3 | 3 | Large | Blue | False |
| 4 | 2 | Small | Blue | False |
| 5 | 4 | Large | Green | False |

**Problem 3 – Representing Knowledge with First-Order Logic (15 points)**

Convert each of the following English sentences into *First-Order Predicate Calculus* (FOPC), using reasonably named predicates, functions, and constants. If you feel a sentence is ambiguous, clarify which meaning you’re representing in logic. (Write your answers in the space *below* each English sentence.)

*Colin is a person who knows someone who owns a red bike.*

*For every species of tree in the world there is a bacterium that only lives in this type of tree.*

*Usually children are younger than their step parents, but not always.*

[You must use the notation of *Markov Logic Networks* here.]

**Problem 4 – Logical Reasoning (15 points)**

1. Is the following wff *valid*, *satisfiable*, or *unsatisfiable* (circle one)*?* Justify your answer formally in the space below.

[ (P → Q) ˅ (Q → P) ] → (P ↔ Q)

1. Put this wff in clausal form; show your work below and put your answer on the line.

¬ ∀ x [ p(x) → q(x) ] \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. We wish to use search to find an interpretation (i.e., a ‘world state’) that satisfies all the clauses below. Show how this set of clauses can be simplified before starting the search.   
   Justify your simplification(s).

(Q ∨ A ∨ ¬ R)

(P ∨ ¬ R)

(P ∨ ¬B ∨ ¬ Q)

(¬ P ∨ R)

(A ∨ B)

**Problem 5 – More Logical Reasoning (10 points)**

1. What is the *most-general unifier* (mgu), if any, of these two wff’s? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Q(?x, ?y, ?x) Q(2, f(1,?z, 2), ?z)*

1. Given the following background knowledge:
2. ∀ x { p(x) → [¬ q(x) ∨ r(x) ] }
3. ∀ y [¬ q(y) → s(y) ]
4. p(1)
5. p(2)
6. r(1)
7. ¬ r(2)

Show: ∃ z s(z) by filling out the table below, using as many lines as needed.

Number WFF Justification

7

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

8

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

9

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

10

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

11

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

12

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

13

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

14

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

15

**Problem 6 – Artificial Neural Networks (10 points)**

Consider a *perceptron* that has two real-valued inputs and an output unit that uses a *step function* as its output function. All the initial weights and the output unit’s threshold equal 0.05*.* Assume the teacher has said that the output should be 1 for the input *in1 = 0.2* and *in2 = -0.7.*

1. Show how the perceptron learning rule (also called the delta rule) would alter this neural network upon processing this training example. Let η (the learning rate) be 0.3.

## Perceptron BEFORE Training

## Perceptron AFTER Training

1. Did these changes improve the network? \_\_\_\_\_\_\_\_\_ Explain your answer below.
2. What kind of search does the perceptron learning rule perform? \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Explain.

**Problem 7 – Support Vector Machines (10 points)**

1. Draw and explain a picture that illustrates the idea and value of searching for a *large margin*. Be sure to compare (verbally and in your picture) to what the perceptron learning rule, without a “weight decay” term, might do.
2. Consider the dataset below. Assume we wish to use the idea of ‘kernels’ to convert this dataset into a new dataset, one where we will apply the perceptron learning rule (this time, with weight decay). Our chosen similarity function is below. Show the new dataset below the old one and be sure to clearly label the columns and rows.

The number of feature pairs (A+B, B+C, and C+A) that have the same values in the two examples (e.g., since ex1 and ex3 both have the same values for features *A* and *B*, that counts as one, whereas ex1 and ex4 do not have the same values for the feature pair A+B and nothing is added to the count).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ex #** | **A** | **B** | **C** | **Output** |
| 1 | True | False | False | True |
| 2 | False | True | True | True |
| 3 | True | False | True | False |
| 4 | True | True | False | False |

**Problem 8 – Important AI Concepts (9 points)**

Describe each of the following AI concepts and briefly explain its most significant aspect.   
(Write your answers in the space *below* the AI concept.)

*CLOSED List*

*description:*

*significance:*

*Overfitting*

*description:*

*significance:*

*Strong AI*

*description:*

*significance:*

Have a good summer vacation!