CSCI 561 - Introduction to Artificial Intelligence

Final Exam, August 2, 2016. **Total Time: 120 minutes (2 hours)**

This exam counts for 30% of the course grade.

Student ID										

USC-email									@usc.edu

First Name ________________________________

Last Name ________________________________

Instructions:

1. Date: **8/2/2016 from 4:00pm – 6:00 pm in SGM124**
2. Maximum credits/points/percentage for this midterm: 100
3. The percentages for each question are indicated in square brackets [ ] near the question.
4. **No books** (or any other material) are allowed.
5. **Write down your name, student ID and USC email address.**
6. **Your exam will be scanned and uploaded online.**
7. **Answers must be written in the provided boxed only.** Please make sure NOT to write the answer to one question in the box for another one.
8. **Do NOT write on the 2D barcode.**
9. **The back of the pages will NOT be graded.** You should use the back of the pages only for SCRATCH PAPER, not the actual answers.
10. No questions during the exam. **If something is unclear to you, write that in your exam.**
11. **Be brief: a few words are often enough if they are precise and use the correct vocabulary studied in class.**
12. When finished, raise completed exam sheets until approached by proctor.
13. **Adhere to the Academic Integrity code.**

1. True/False Questions (10 points)

1 point each

For each of the statements below, fill in the box 'T' if the statement is always and unconditionally true, or fill in the box 'F' if it is always false, sometimes false, or just does not make sense.

1	T	■
2	■	F
3	■	F
4	■	F
5	T	■
6	T	■
7	■	F
8	T	■
9	■	F
10	■	F

1) The exact evaluation function values affect minimax decision even if the ordering of these values is maintained.

2) Soundness of an inference algorithm means that the algorithm doesn't reach bogus conclusions.

3) The number of hidden layers needed in a neural network cannot be calculated from the input provided.

4) Bayes' Theorem takes you from a prior belief to a posterior belief after you've collected some additional data

5) If all cars would be self-driving, the world's roads would be a deterministic environment.

6) The entropy of a binary random variable decreases as data become less ordered.

7) Even a good knowledge base will not be able to answer any question.

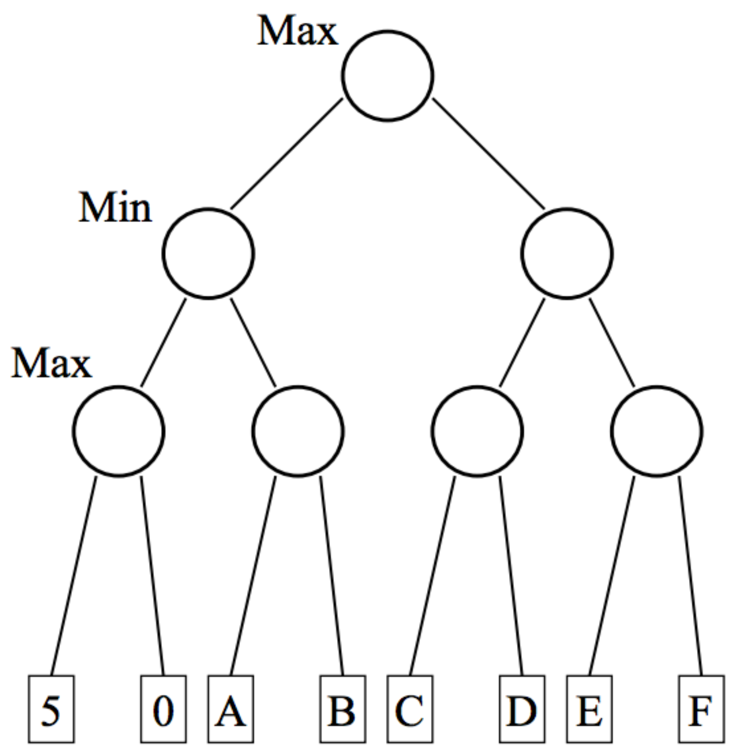
8) Perceptrons, arranged with no hidden layer, cannot express logical NOT.

9) SVM can classify data that is not linearly separable.

10) The back-propagation learning algorithm is based on the gradient descent method.

2. Game Trees (20 points)

Consider the game tree picture below where *A-F* represent some integer values. Assume the nodes are explored from left to right and standard alpha beta pruning is used.



(a) [4 pts] Give a value of *A* such that *B* is pruned.

Any thing > 5

(b) [4 pts] Give a value of *A* such that *B* is not pruned.

Any thing < 5

(c) [4 pts] **True of False:** There are SOME values of A and B such that the subtree containing C and D is pruned. Explain Why.

False. Because the leftmost subtree of any subtree cannot be pruned as that is the first value that can be known for that subtree and all the decisions of the pruning can be done after that only.

1 point for False and 3 points for explanation

(d) [4 pts] Assuming that $B = 5$ and $A = 5$, give a value of C and D such that the subtree containing E and F is pruned.

Anything such that max(B,C)<5

(e) [4 pts] If you are allowed to assign A-F arbitrarily, what is the MAXIMUM number of leaves that can be pruned? Provide the name of the leaves as well.

3 (B,E and F)

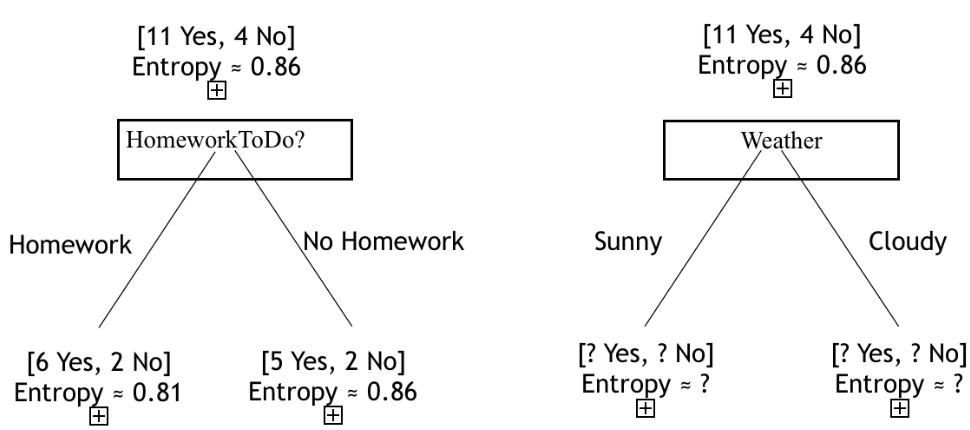
2 point for 3
2 point for the rest

3. Decision Trees (20 points)

In trying to decide whether or not to go jogging, some of the factors I consider are the weather and whether or not I have homework to do. There are other factors as well of course. Some example decisions are summarized in the table below.

Weather	HomeworkToDo	Jogging
Sunny	No Homework	Yes
Sunny	Homework	No
Sunny	No Homework	Yes
Sunny	Homework	No
Sunny	Homework	Yes
Sunny	No Homework	No
Sunny	Homework	Yes
Sunny	No Homework	No
Cloudy	Homework	Yes
Cloudy	Homework	Yes
Cloudy	Homework	Yes
Cloudy	No Homework	Yes
Cloudy	Homework	Yes
Cloudy	No Homework	Yes
Cloudy	No Homework	Yes

With 11 positive examples of jogging and 4 negatives, the entropy or information of this decision in bits is 0.84. Consider the following decision trees, reflecting splitting on the attributes of **Weather** or **HomeworkToDo**.



A. [8 pts] The HomeworkToDo tree has been filled out. Complete the values for the Weather tree, including entropy. (Show your work)

weather: Sunny -> 4 Yes, 4 No, Entropy = 1, entropy = -0.5*log(0.5) – 0.5*log(0.5) = 1

weather: Cloudy -> 7 Yes, 0 No, Entropy = 0, entropy = -1*log(1) – 0*log(0) = 0

4 point each

B. [8 pts] Calculate the information gain IG from splitting on HomeworkToDo and Weather. Please show formulas used and steps clearly. You do not need to compute a final numerical result. Plug in values to the formulas so that the result could be calculated with a calculator.

IG(HomeworkToDo) =

IG(HomeworkToDo) = I(Jogging) – remainder(HomeworkToDo) = 0.86 – 8/15 * I(6/8, 2/8) – 7/15 * I(5/7, 2/7) = 0.86 – 8/15 * 0.81 – 7/15 * 0.86 = 0.0267

4 point each

IG(Weather) =

IG(Weather) =(Jogging) – remainder(Weather) = 0.86 – 8/15 * I(4/8, 4/8) – 7/15 * I(7/7, 0/7) = 0.86 – 8/15 * 1 – 7/15 * 0 = 0.3267

C. [4 pts] Which of the two attributes is a better choice in constructing a decision tree? Why?

Weather is the better choice. Because the information gain (IG) from it is more.

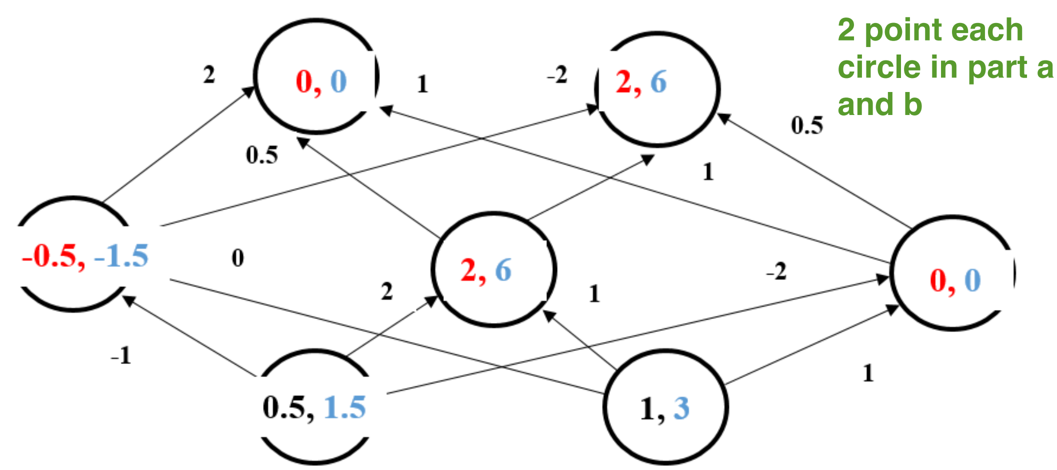
2 point each

4. Neural Networks (25 points)

The following is a network of **linear** neurons, that is, neurons whose output is identical to their net input. The numbers in the circles indicate the output of a neuron, and the numbers at connections indicate the value of the corresponding weight. (Check the directions of the arrows on the connections while answering the questions).

a) [10 pts] Compute the output of the hidden-layer and output-layer neurons for the given input and enter those values into the corresponding circles.

b) [10 pts] What is the output of the network for the input (1.5,3) i.e. the left input neuron having the value 1.5 and the right one having the value 3? Fill in values for all the circles.



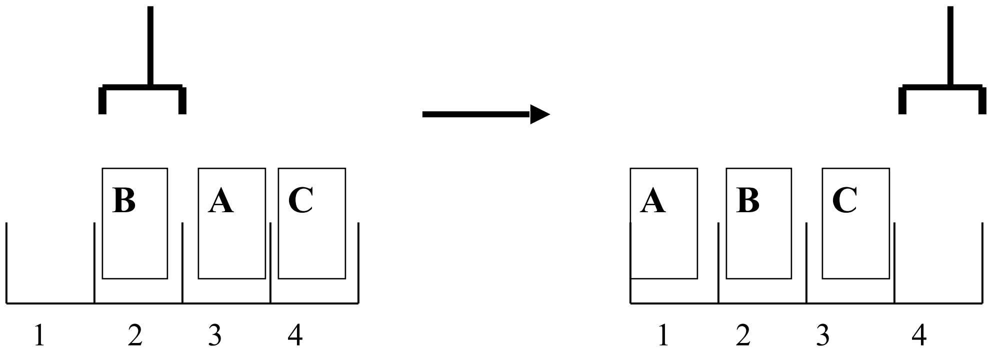
c) [5 pts] To answer part (b), do you have to do all the network computations once again? Explain why you do or do not have to do this.

Since they are linearly proportional we do not need to recalculate all of them.

2 point , 3 point
explanation

5. Planning (15 points)

Consider writing a planner that will solve the following block-moving problem:
6.1 [10pts] Consider the given domain, and formalize the following sentences:



There are three operators available:

- **pickup (obj1,pos1)**
  Pick up the object **obj1**, which is currently at position **pos1**. The gripper must also be at **pos1**.
- **putdown (obj2,pos2)**
  Put down the object **?obj2**, which is currently being held, at position **pos2**. Gripper must be at **pos2** as well.
- **move-gripper (pos1,pos2)**
  Move the gripper from position **pos1** to position **pos2**.

The gripper can only be holding one thing at a time, and there can be at most one block at a position at a time.

a) [5 pts] Write the operator definitions for **pickup**, **putdown**, and **move-gripper**.

pickup(b, p)
Pre: bp(b, p), ge, gp(p)
Effects: -ge, -bp(b, p), +gh(b), +pe(p)

putdown(b, p)
Pre: gh(b), gp(p), pe(p)
Effects: -pe(p) –gh(b), +bp(b, p), +ge

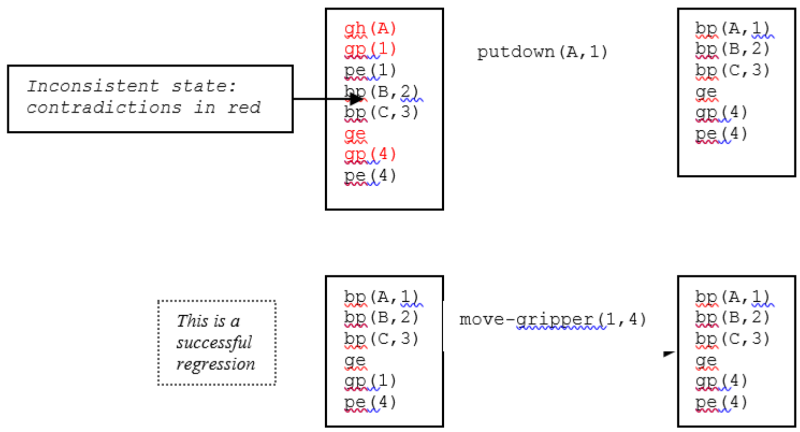
move-gripper(p1, p2)
Pre: gp(p1)
Effects: -gp(p1), +gp(p2)

-1 for missing each

b) [5 pts] Demonstrate a regression search for a solution: show the goal and two regressions, one that fails, and one that succeeds.

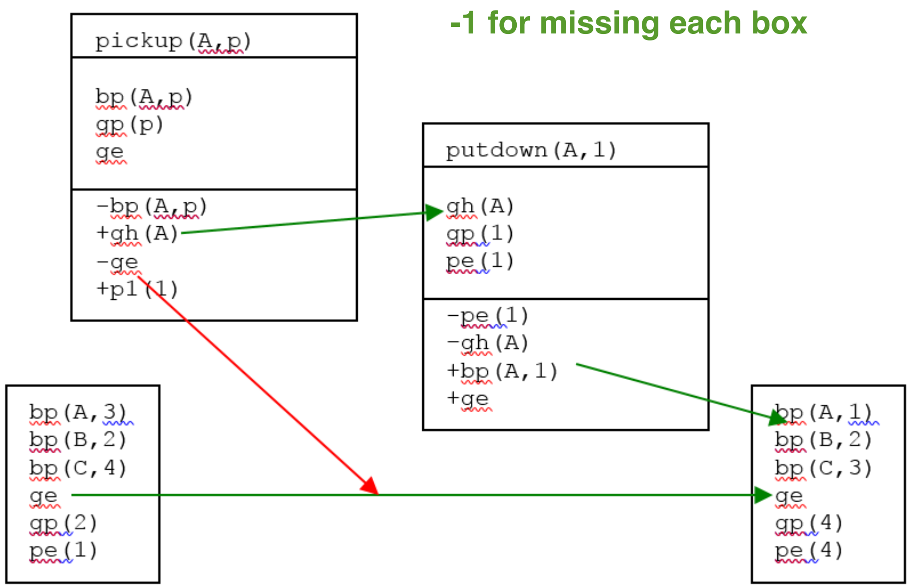
Here's an example of a failed regression and a successful regression. The last action in the plan must be move-gripper(x, 4) so for failure we will make the last action which is part of a solution plan, but not the last action.

-1 for missing each



c) [5 pts] Demonstrate a plan-space search for a solution: show the initial plan and three refinements: one that links to the initial state, one that involves adding an operator, and one that involves resolving a threat.

-1 for missing each box



6. Short Answers (10 points)

1. [2 pts] List 2 reasons why machine learning is needed?

Unknown environments
Adaptability
Lazy
Autonomous

+1 for a valid reason

2. [2 pts] What is Ockham's razor? How is it used in decision tree learning?

Bias for simplest hypothesis +1

Finds smallest tree first +1

3. [6 pts] Choose 2 research topics from the poster sessions and describe each topic, including the limitations/pitfalls and future directions. Compare and contrast how uncertainty was handled.

+1 for each topic

+1 for each description including the limitations/pitfalls and future directions

+1 for discussing uncertainty

+1 for compare and contrasting uncertainty