CSCI 4450/8456 Artificial Intelligence

## Fall 2018: Homework 2

**Instructor: Raj Dasgupta**

**Given on: September 20, 2018 (Thursday)**

**Due on: October 4, 2018 (Thursday), 11:59 PM**

**Programming part: Submissions on loki**

**Mathematical problems part: Submissions on Canvas**

**Maximum Points: 65**

**30 points (programming part)**

**35 points (math problems part)**

## Submission Instructions

**Programming Part**

You need to submit your source code that you have written for your assignment, and a README.txt file that gives the commands to compile and run your program for each question. Submissions for programming part are via loki only. DO NOT SUBMIT ON programming part on Canvas. **It is your responsibility to ensure that your program compiles and runs properly on loki using the commands in your README file.**

1. Login to loki. You will be logged into your home directory
2. Create a sub-directory inside your home directory called CSCI-4450-DG-F18-A2 by using the command mkdir CSCI-4450-DG-F18-A2 from inside your home directory.
3. You must create the directory in step 2 right under your home directory and **with the exact name given above**. Otherwise, your homework will not be collected and graded.
4. Place your source code in the directory that you created in step 2 above before the assignment deadline.
5. Note that assignments will be automatically collected using a pre-scheduled script at the due date/time from the above directory names.

**Math Problems Part**

Only electronically typed answers allowed in MS Word or pdf file format. Upload a single document containing your answers for the math problems on Canvas. Submissions for math problems part are via Canvas only. DO NOT SUBMIT ON math part on loki.

## Late Policy

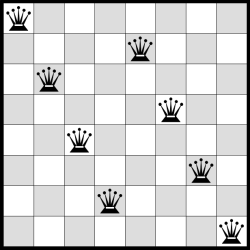
Programming part: Late submissions are NOT encouraged. Late policy for the programming part of the assignment will be according to the policy given in the first day handout in class, which is: For every 24 hour period following the submission deadline, 10% of the total points for the assignment (programming part + math part) will be deducted. No submissions for programming part will be accepted after 48 hours from the deadline. If you are submitting late for the programming part, send an email to Dr Dasgupta [pdasgupta@unomaha.edu](mailto:pdasgupta@unomaha.edu) between 24 hours of the regular deadline to receive late submission instructions.

Math Problems part: NO LATE SUBMISSIONS ALLOWED. If you aren’t able to submit on Canvas by the deadline, you will get zero points for the math problems part of the assignment.

## Programming Part

**Question 1 – Genetic Algorithm Crossover Operators (30 points)**

In the lectures slides on local and online search, we discussed about the 8 queens problem, where the objective is to place 8 queens on a chess board so that no two queens can attack each other. A solution corresponding to a goal state is shown in the picture below.



In this question, your objective is to compare the performance of different crossover operators. The crossover operators you need to program are the following: **single point crossover, two point crossover, cut and splice, uniform crossover.** More information about these crossover operators are given on the slides. Feel free to search on the Internet for references and implementation ideas for these operators

You should use the same representation of a chromosome, fitness function, etc. from the class slides/textbook. For example, the chromosome or state for the picture above is 86427531.

Your program should first prompt the user for the following input (just input values on a single line, separated by blank space):

* Initial population size
* Number of chromosomes to select at each iteration or generation of the GA
* Mutation rate
* Number of iterations or generations to run the GA for; -1 to run the GA until a goal state is found.

Other parameters can be hardcoded. Then, it should prompt the user to choose one of the crossover operators from a menu. Add an option to exit the program from this menu. After a crossover operator is selected, the program runs. At the end of the run, your program should print (each output on a separate line):

* Values of input and hardcoded parameters used (on a single line, comma-separated, format:- <parameter name: value>)
* Number of iterations or generations of the GA executed
* Final state reached
* Final state’s fitness function value.

After printing the output, the program should go back to the first menu prompt for another the crossover operator or exit.

## Mathematical Problems Part

**Question 1 - Decision Tree Learning (15 points, 3 points for each subpart)**

You have a decision tree algorithm and you are trying to figure out which attribute is the best to test on first. You are using the “information gain” metric.

* + You are given a set of 128 examples, with 64 positively labeled and 64 negatively labeled.
* There are three attributes: Homeowner (H), In Debt (ID), and Rich (R).
* For 64 examples, Home Owner is true. The Homeowner=true examples are 1/4 negative and 3/4 positive.
* For 96 examples, In Debt is true. Of the In Debt=true examples, 1/2 are positive and half are negative.
* For 32 examples, Rich is true. 3/4 of the Rich=true examples are positive and 1/4 are negative

You must show all mathematical calculations/steps to get full points for each subpart (a) – (d) below. Just writing the final answer in each subpart (correct or not) will get zero points.

1. What is the entropy of the initial set of examples?
2. What is the information gain of splitting on the Home Owner attribute as the root node?
3. What is the information gain of splitting on the In Debt attribute as the root node?
4. What is the information gain of splitting on the Rich attribute as the root node?
5. Which attribute do you split on?

**Question 2. Decision Tree Learning (adapted from Russell Norvig Exercise 18.6) (15 points)**

You are given the following training set with three attributes A1, A2 and A3 and one binary output y. Construct the decision tree for the hypothesis in the training set. You have to show the steps for the calculations of the information gain of each attribute and the complete decision tree to get full points.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Example** | A1 | A2 | A3 | Output y |
| **X1** | 1 | 0 | 0 | 0 |
| **X2** | 1 | 0 | 1 | 0 |
| **X3** | 0 | 1 | 0 | 0 |
| **X4** | 1 | 1 | 1 | 1 |
| **X5** | 1 | 1 | 0 | 1 |

**Question 3. Artificial Neural Network (Adapted from Russell-Norvig Problem 18.19) (5 points)**

Construct by hand a neural network for the Boolean XOR function with two inputs.