# CMSC 671 (Introduction to AI) – Fall 2016

## Homework 4: CSPs, Games, and Probabilities (100 points)

Due: 11/7 at 11:59pm.

Turnin: Blackboard.

Please submit all parts together as a **single PDF file** named *lastname*_hw4.pdf, with parts clearly marked and delineated. **All** files must start with your last name and have your full name(s) in the file, at/near the top.

You are encouraged to work on this homework assignment in groups of up to **two** students. If you do so, you only need to turn in one solution, with everyone's name in the file (files should be named after the person who submits). Remember, if you work in a group, you must actually work on the problems *as a group*, not split up the work and combine the independent solutions.

Remember, you will only be able to turn in one final solution.

### PART I. LEARNING IN THE WILD (15 PTS.)

**Assignment:** Consider the problem faced by a small child learning to solve a shape sorter puzzle. *(Maximum: 300 words)*

- Explain how this problem fits into the general *learning model*.
- Describe the percepts (sensor inputs) and actions of the child, and the types of learning the child must do.
- Describe the subfunctions the child is trying to learn in terms of inputs, outputs, and available training data.

### PART II. INFORMATION GAIN (15 PTS.)

Suppose that:

- An attribute $a$ splits the set of examples $E$ into subsets $E_k$
- Each subset has $p_k$ positive examples and $n_k$ negative examples

**Assignment:** Show that the attribute has strictly positive information gain *unless* the ratio $\frac{p_k}{(p_k + n_k)}$ is the same for all $k$.

Show and explain each step of the solution in order to receive full credit. *(R & N, Exercise 18.5)*

### PART III. LEARNING BAYES NETS (20 PTS.)

Consider an arbitrary Bayesian network, a *complete* data set for that network, and the likelihood for the data set according to the network.

**Assignment:** Give a clear explanation (i.e., an informal proof, in words) of why the likelihood of the data *cannot decrease* if we add a new link to the network and recompute the maximum-likelihood parameter values.

## PART IV. DECISION TREE LEARNING (50 PTS.)

You find you must decide how to feed yourself *every single night*, which seems awfully inefficient! As a sensible CS student, you decide to build a decision tree to give a prediction whether you will stay home or go out, based on data from the last few weeks.

**Assignment:** Answer questions 1–3 using this data:

**Eat out or cook at home?**

| **Do I have homework? (*H*)** | **What's in the fridge? (*F*)** | **Friends going out? (*G*)** | **Yelp reviews (*Y*)** | **Class** |
|---|---|---|---|---|
| Yes | Food | Yes | ☆☆☆☆ | Stay home |
| No | Ingredients | Yes | ☆ | Stay home |
| No | Food | Yes | ☆½ | Eat out! |
| Yes | Ingredients | No | ☆½ | Stay home |
| Yes | Food | Yes | ☆☆½ | Stay home |
| Yes | Ingredients | Yes | ☆☆☆ | Eat out! |
| Yes | Nothing | No | ☆☆ | Stay home |
| Yes | Ingredients | No | ☆☆☆ | Eat out! |
| No | Ingredients | No | ½ | Stay home |
| Yes | Nothing | Yes | ☆☆☆½ | Eat out! |
| No | Nothing | Yes | ☆☆☆☆☆ | Eat out! |
| No | Nothing | No | ☆ | Eat out! |
| Yes | Food | No | ☆☆☆☆½ | Stay home |

1. **Information Gain (15 pts.)**

   a. At the root node for your decision tree in this domain, what is the information gain associated with a split on the attribute *F*?

   b. What would it be for a split at the root on the attribute *Y*? (Use a threshold of ✫✫½ for *Y* (i.e., assume a binary split: $Y \leq$ ✫✫½ , $Y >$ ✫✫½.)

2. **Gain Ratios (15 pts.)**

   a. Again at the root node, what is the *gain ratio* associated with the attribute *F*? What is the gain ratio for the *Y* attribute at the root (using the same threshold as in 1b)?

3. **Decision Tree (20 pts.)**

   a. (2 pts.) Suppose you build a decision tree that splits on the *H* attribute at the root node. How many child nodes are there are at the *first level* of the decision tree?

   b. (2 pts.) After *H*, which branches require a further split?

   c. (8 pts.) Draw the smallest (fewest nodes) decision tree you can you construct for this dataset. The tree should show which attribute you split on for each branch, and show the decisions (class predictions) at the leaves.

   d. (8 pts.) What method(s) did you use to find that tree? Show all calculations.