# CSC 535 – Probabilistic Graphical Models

# Assignment Five

### Due: 11:59pm (*) Monday, October 22.

### (*) There is grace until 8am the next morning, as the instructor will not grade assignment before then. However, once the instructor starts grading assignments, no more assignments will be accepted.

### Weight about 7 points

### This assignment should be done individually

---

The purpose of this assignment is to become very familiar of undirected graphical models (Markov random fields). We will also consider causality which is featured in problem 1.

---

## Deliverables

Deliverables are specified below in more detail. For the high level perspective, you are to provide a program to output a few numbers and to create figures. You also need to create a PDF document that tells the story of the assignment including output, plots, and images that are displayed when the program runs. Even if the question does not explicitly remind you to put the resulting image into the PDF, if it is flagged with **($)**, you should do so. The instructor should not need to run the program to verify that you attempted the question. See

http://kobus.ca/teaching/grad-assignment-instructions.pdf

for more details about preparing write-ups. While it takes work, it is well worth getting better (and more efficient) at this. A substantive part of each assignment grade is reserved for exposition.

---

1. **Casual directed graphical models**. Consider people trying out a potential drug cure for some affliction in a medical experiment. Consider the binary variable D (whether a person takes the drug), the binary variable C (whether they get cured), and a binary variable H (whether they are health conscious). Whether they get cured is a (statistical) a function both of whether they took the drug, and whether they are health conscious. Further, the probability they took the drug can depend on whether they are health conscious. Provide a Bayes net for the statistical dependencies of the variables **($)**. Suppose we want to know whether the drug is in fact helpful or not, even if it requires running an expensive experiment. Propose an experiment, and provide a graphical model for this **($).**

   Now consider determining whether a doctor should prescribe the drug. Create numbers that shows that the answer can be different for the two graphs (**$**). Note that by referring to the second graph, you can directly determine the probability of being cured if they take the drug (because someone did the experiment). For the first graph, you need to deal with the probability that a given person is in one of the four cases.

   You should be able to show that the drug can be bad for both groups (health conscious or not), but using the first graph can suggest that the drug should be prescribed **($).** [Note: This is sometimes called Simpson's paradox].

   *Hint. Suppose health conscious folks take their meds, and non-health conscious folks do not. Now consider what would happen if health conscious folks generally get better, even if they take a drug that is bad for them (and makes them slightly less likely to get better).*

2. Recall the model from the previous assignment which was: Suppose that your undergraduate **grades** influence the **letters** that professors write for you (higher grade, better letter). Another thing that influences those letters is your **interactions** with those professors (better interaction, better letter). Now suppose that the **rank** of the graduate program you get into is influenced by your **GRE**, your grades, and your letter (better letter increases probability that rank will be higher). Finally, suppose that whether or not you get an academic **job** depends on the graduate program rank (better rank increases your chances), and the **state** where the job is located (political comment goes here).

   a) Remind the grader what your initial directed graphical model was **($)**, and convert it an undirected graph **($)**. Explain your steps **($)**. Does doing so loose any independence information **($)**?

   b) Provide potential functions for the undirected graph just made based on the directed graphical model **($)**. Make sure you map the potentials to expressions for the probabilities which are available from the directed graphical model.

3. The following matrix tells us who are friends in a cohort of 10 students who conveniently have a different first initial, and all initials are in the first ten letters of the alphabet (e.g., Alice, Bob, Carol, Doug, ...). The cells are "1" if the persons with the initials listed in the row label and the column label are friends. The matrix is symmetric.

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| A | | 1 | 1 | | | | | | | |
| B | 1 | | 1 | 1 | | | | | | |
| C | 1 | 1 | | 1 | | | | | | |
| D | | 1 | 1 | | 1 | | | | | |
| E | | | | 1 | | 1 | | | 1 | |
| F | | | | | 1 | | 1 | 1 | | |
| G | | | | | | 1 | | | | |
| H | | | | | | 1 | | | | |
| I | | | | | 1 | | | | | 1 |
| J | | | | | | | | | 1 | |

(A) Draw an undirected graph, one node per person, with links between friends **($)**. We will associate with each person a random variable representing whether their preferred computer is a mac or not, and friends might influence each others computer preference, but we assume that computer use by folks you do not know are irrelevant. We also have a Microsoft spy camera that detects whether each person is using a mac or not, but it is not very reliable. Represent the detector output on your graph as well **($)**.

(B) Identify the maximal cliques in the graph **($)**.

(C) Provide an expression using potential functions for the probability distribution for the assignment of all the values over the network (e.g., a distribution over any particular assignment of who prefers macs and what they were detected using) **($)**. The arguments of the potential functions matter, but you do not need to produce expressions for them.

(D) Is mac preference by A independent of mac preference of J? Provide an argument based on rules learned in class **($)**.

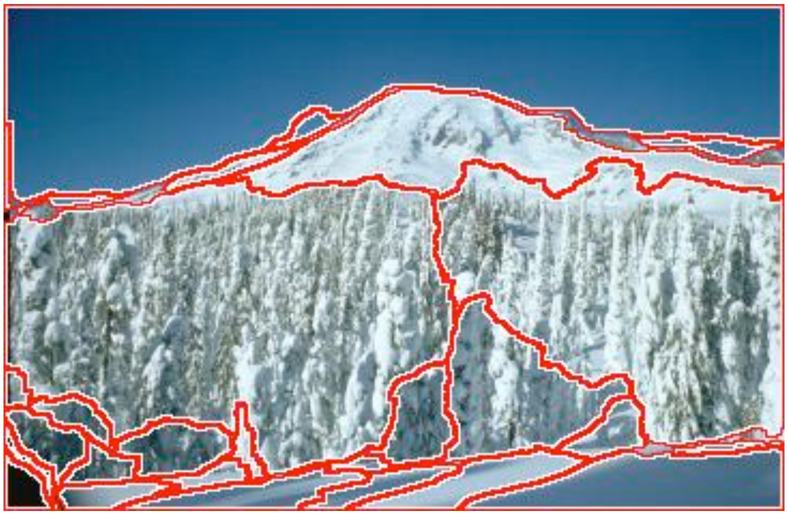
(E) Is mac preference by A independent of mac preference by D given mac preference by C? Provide an argument based on rules learned in class **($)**.

(F) Is mac preference by A independent of mac preference by E given mac preference by D? Provide an argument based on rules learned in class **($)**.

(G) Construct an energy function for this model that has the following properties: 1) Everyone has a fixed energy (bias) preference for macs. 2) If one of your friends has a mac, then you are more likely to have one; 3) Three mutual friends having a mac is less likely because mac users have to be different, and so one of them will switch to a PC running Linux, rather than be in a group of three mac users; 4) The detector output for which computer you were spotted using is statistically correlated with your actual preferred computer.

*Note that this does not specify the energy function exactly. You are creating an energy function of your choice with those properties.*

4. Consider images with captions like the one shown below. Assume that you have very simple captions and a high quality parser than can extract noun-phrases and prepositions and their bindings. In other words, in the example, you know there are three nouns (sky, peak, trees) and that there are two spatial relations between (sky, peak) and (trees, peak). (It may seem obvious that you would know that sort of thing, but in natural language it can be very hard to get the correct parse). Assume that we have functions which are proportional to p(noun, region-features). Finally, we have a segmentation of the image into regions which we will assume do not span noun areas, but that may need to be merged to get a perfect semantic segmentation of the image.



The sky is above the peak. The trees are below the peak.

a) Considering only nouns, provide a scheme for creating a Markov random field (MRF) for modeling the probability of region labels using the fact that adjacent regions are more likely to be associated with the same noun, and that nouns are linked to features (e.g., color and texture) via the function mentioned above, illustrated by a PGM (**$**). Also provide an abstract formulation of the probability distribution (**$**). Finally, explain in your own words how random assignments of labels to regions can be compared for fitness with respect to the model (**$**).

b) Extend your scheme for your MRF to include prepositions. They are one level higher in abstraction, since they are about relations between nouns. Again provide a description **($)**, an illustration **($),** and an abstract formulation of the probability distribution (**$)**

*One way to think about this is to assume that you build an MRF for each segmented image on the fly based on the data (and then you might use that MRF to label the image (or produced a distribution over labelings)). Then, in the example, working with the proposition "above", you could imagine a link between all pairs of nodes where the segments associated with the nouns could sensibly be related by "above" based on segment geometry. Those factors could reward assignments where the nouns appeared in the correct slots in the text.*

## What to Hand In

If you wrote code for this assignment, you should hand in a program hw5.<suffix> (e.g., hw5.m if you are working in Matlab) and the PDF file hw5.pdf with the story of your efforts into D2L.