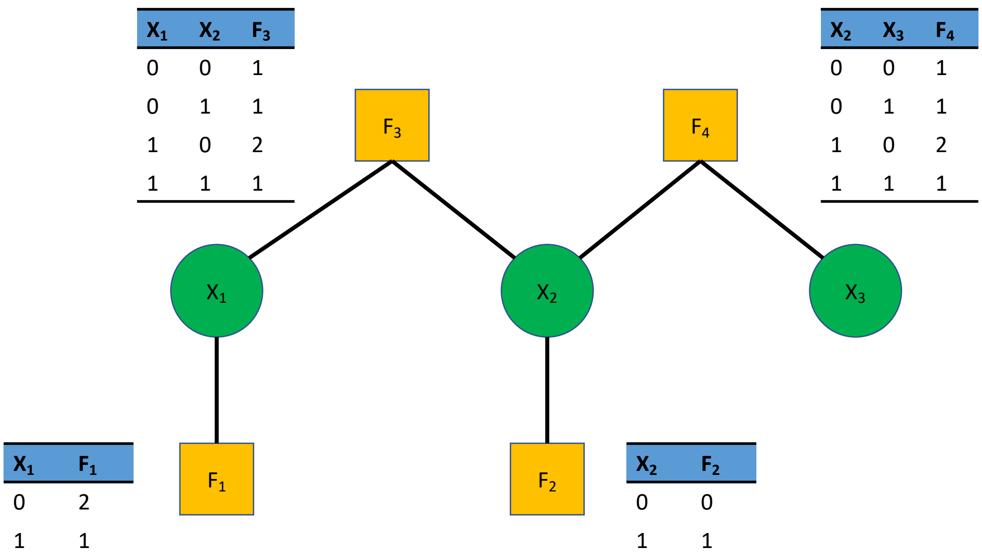
CS221 Problem Workout Solutions

Week 7

1) Problem 1

This problem will give you some practice on computing probabilities given a Markov network. Specifically, given the Markov network below, we will ask you questions about the probability distribution $p(X_1, X_2, X_3)$ over the binary random variables $X_1, X_2,$ and X_3.



X_1	X_2	F_3
0	0	1
0	1	1
1	0	2
1	1	1

X_2	X_3	F_4
0	0	1
0	1	1
1	0	2
1	1	1

X_1	F_1
0	2
1	1

X_2	F_2
0	0
1	1

(a) What is $p(X_1 = 0, X_2 = 0, X_3 = 0)$?

Solution $p(X_1 = 0, X_2 = 0, X_3 = 0) = \frac{F_1(X_1=0)F_2(X_2=0)F_3(X_1=0,X_2=0)F_4(X_2=0,X_3=0)}{Z} = \frac{2\times0\times1\times1}{Z} = 0$, where
$Z = \sum_{X_1\in\{0,1\}} \sum_{X_2\in\{0,1\}} \sum_{X_3\in\{0,1\}} F_1(X_1)F_2(X_2)F_3(X_1, X_2)F_4(X_2, X_3) = 9$

(b) What is $p(X_1 = 0, X_2 = 1, X_3 = 0)$?

Solution $p(X_1 = 0, X_2 = 1, X_3 = 0) = \frac{F_1(X_1=0)F_2(X_2=1)F_3(X_1=0,X_2=1)F_4(X_2=1,X_3=0)}{Z} = \frac{2\times1\times1\times2}{Z} = \frac{4}{9}$, where Z was computed in part a).

(c) What is $p(X_2 = 0)$?

Solution $p(X_2 = 0) = 0$

(d) What is $p(X_3 = 0)$?

Solution $p(X_3 = 0) = \frac{\sum_{X_1 \in \{0,1\}} \sum_{X_2 \in \{0,1\}} p(X_3=0, X_1, X_2)}{Z} = \frac{6}{9} = \frac{2}{3}$

2) **Problem 2** Farmer Kim wants to install a set of sprinklers to water all his crops in the most cost-effective manner and has hired you as a consultant. Specifically, he has a rectangular plot of land, which is broken into $W \times H$ cells. For each cell (i, j), let $C_{i,j} \in \{0, 1\}$ denote whether there are crops in that cell that need watering. In each cell (i, j), he can either install $(X_{i,j} = 1)$ or not install $(X_{i,j} = 0)$ a sprinkler. Each sprinkler has a range of R, which means that any cell within Manhattan distance of R gets watered. The maintenance cost of the sprinklers is the sum of the Manhattan distances from each sprinkler to his home located at $(1, 1)$. Recall that the Manhattan distance between (a_1, b_1) and (a_2, b_2) is $|a_1 - a_2| + |b_1 - b_2|$. Naturally, Farmer Kim wants the maintenance cost to be as small as possible given that all crops are watered. See figure below for an example.

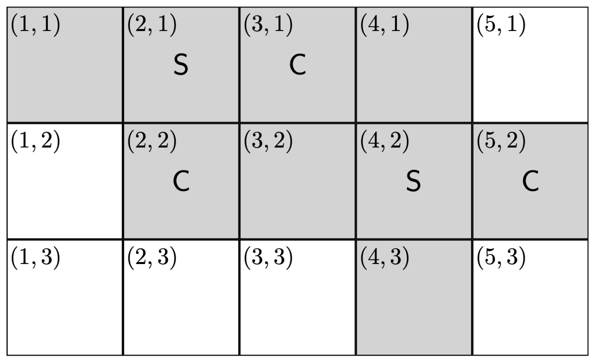


Figure 1: An example of a farm with $W = 5$ and $H = 3$. Each cell (i, j) is marked 'C' if there are crops there that need watering $(C_{i,j} = 1)$. An example of a sprinkler installation is given: a cell (i, j) is marked with 'S' if we are placing a sprinkler there $(X_{i,j} = 1)$. Here, the sprinkler range is $R = 1$, and the cells that are shaded are the ones covered by some sprinkler. In this case, the sprinkler installation is valid (all crops are watered), and the total maintenance cost is $1 + 4 = 5$.

Farmer Kim actually took CS221 years ago, and remembered a few things. He says: "I think this should be formulated as a factor graph. The variables should be $X_{i,j} \in \{0, 1\}$ for each cell (i, j). But here's where my memory gets foggy. What should the factors be?" Let $X = \{X_{i,j}\}$ denote a full assignment to all variables $X_{i,j}$. Your job is to define two types of factors:

- $f_{i,j}$: ensures any crops in (i, j) are watered,
- f_{cost}: encodes the maintenance cost,

so that a maximum weight assignment corresponds to a valid sprinkler installation with minimum maintenance cost.

$$f_{i,j}(X) = \left[C_{i,j} = 0 \text{ or } \left(\min_{i',j':X_{i',j'}=1} |i'-i| + |j'-j|\right) \leq R\right]. \quad (1)$$

$$f_{\text{cost}}(X) = \exp\left(-\sum_{i',j':X_{i',j'}=1} |i'-1| + |j'-1|\right). \quad (2)$$

(i) [5 points] Recall that Gibbs sampling sets $X_{i,j} = 1$ with some probability p. For convenience, use the notation $X \cup \{X_{i,j} : 1\}$ to denote a modification of X where $X_{i,j}$ has been assigned 1 (analogously for 0). Write an expression for p in terms of the factors (e.g., f_{cost}). Your expression should involve as *few* factors as possible.

$p =$

Solution The only factors that depend on $X_{i,j}$ are f_{cost} and the $f_{i',j'}$ of any cells (i', j') within range of R. Let's multiply only those factors together for a candidate choice $v \in \{0, 1\}$:

$$w_v = f_{\text{cost}}(X \cup \{X_{i,j} : v\}) \prod_{i',j':|i'-i|+|j'-j|\leq R} f_{i',j'}(X \cup \{X_{i,j} : v\}). \quad (3)$$

Then the probability p is just proportional to that:

$$p = \frac{w_1}{w_1 + w_0}. \quad (4)$$

(ii) [5 points] Recall Gibbs sampling is guaranteed to find the optimal assignment eventually if there is a non-zero probability of reaching any valid assignment X' from the initial assignment X. Prove that this is the case for any X, X'.

Solution Note that Gibbs sampling chooses an assignment with probability proportional to its weight, so therefore it cannot choose an assignment with weight 0. We must show that we can reach any assignment without going through any zero-weight assignment. The key insight is adding sprinklers to an assignment X with non-zero weight cannot make its weight zero (although it can decrease its weight). Take any two assignments X, X' with non-zero weight. We can construct a path through intermediate assignments with non-zero weight as follows: add sprinklers one by one until all of them are added, and then remove sprinklers one by one until we obtain X'. Note that this is only one such positive probability path, which is probably not the best one, but it suffices to prove the claim.

3) **Problem 3** Some artists have finished new paintings and are trying to display them. Luckily, some galleries are looking for new paintings to display. It is your job to match artists and galleries taking into account the preferences of the artists, the preferences of the galleries, and the capacity of each gallery.

Here is the formal art gallery matching problem setup:

(a) There are m artists $A_1, ..., A_m$ who each have a single painting they would like displayed.

(b) There are n galleries $G_1, ..., G_n$ that have space to display paintings.

(c) Each artist A_i specifies arbitrary non-negative preferences $PA_1^{(i)}, ..., PA_n^{(i)} \geq 0$ for each of the n galleries. A large preference value of $PA_j^{(i)}$ means that artist A_i really wants their painting to be displayed in gallery G_j, and a preference value of 0 for $PA_j^{(i)}$ means that artist A_i does not want their painting to be displayed in gallery G_j.

(d) Each gallery G_i specifies arbitrary non-negative preferences $PG_1^{(i)}, ..., PG_m^{(i)} \geq 0$ for each of the m artists. A large preference value of $PG_j^{(i)}$ means that gallery G_i really wants to display artist A_j's painting, and a preference value of 0 for $PG_j^{(i)}$ means that gallery G_i does not want to display artist A_j's painting

The art gallery matching process has the following requirements:

(a) Each gallery G_i can have a maximum of 1 painting displayed.

(b) Each artist must be matched to exactly one gallery for which they have specified a *positive* preference (assume each artist has at least one such preference) and for which the chosen gallery specifies a *positive* preference for the artist (assume each gallery has at least one such preference).

a. (*16 points*) We can model the art gallery matching process as a CSP. Our CSP should find the assignment with the *maximum weight* as determined by the product of the preference weights of the artists and galleries all together. There are two possible formulations of this CSP - one with m variables, one for each artist $A_1, ..., A_m$, and one with n variables, one for each gallery $G_1, ..., G_n$.

Finish the specification of this CSP for each of the formulations by stating the domains of each variable and the factors needed. You may define any notation/helper functions to help you concisely express your answers below.

Formulation 1: Artists as Variables (For this formulation, you should use only unary and binary factors.)

- **Variables (Already given):** We have m variables for the artists $A_1, ..., A_m$
- **Domains (how large is each and what are the values?):**

Solution The domain for each artist is of cardinality n with values $G_1, ..., G_n$

- **Factors (Use only unary and binary factors. State the arity of each and write them as functions from variables to scalars):**

  Solution There are three sets of factors.

  The first set encodes the maximum number of paintings that can be displayed at each art gallery. Noting that for all pairs of artists, their assigned galleries must be unique, we can account for this with a set of binary factors for each pair of artists. For A_i and A_j where $i, j \in \{1, ..., m\}$ and $i \neq j$, we have a factor

  $$f_{i,j}(A_i, A_j) = \mathbb{1}[A_i \neq A_j]$$

  The second set encodes individual artist preferences for where they would like their paintings displayed. This can be written as unary factors $g_1, ..., g_m$ where

  $$g_i(A_i) = PA^{(i)}_{A_i}$$

  which is the preference of the ith artist to have their painting displayed by gallery A_i (read: value of A_i)

  The third set encodes each gallery's preferences for which artist's paintings they would like to display. This can be written as unary factors $h_1, ..., h_n$ where

  $$h_i(A_i) = PG_i^{(A_i)}$$

  which is the preference of the gallery A_i (read: value of A_i) to have the ith artist's painting displayed.

Formulation 2: Galleries as Variables

- **Variables (Already given):** We have n variables for the galleries $G_1, ..., G_n$
- **Domains (how large is each and what are the values?):**

  Solution The domain for each gallery is of cardinality m+1 with values $0(unassigned), A_1, ..., A$

- **Factors (state the arity of each and write them as functions from variables to scalars):**

Solution There are sets of factors.

The first set encodes each gallery's preferences for which artist's paintings they would like to display if they are assigned. This can be written as unary factors $f_1, ..., f_n$ where

$$f_i(G_i) = PG^{(i)}_{G_i} * \mathbb{1}[G_i \neq 0] + \mathbb{1}[G_i = 0]$$

which is the preference of the ith gallery to have the artist G_i's painting displayed.

The second set encodes the each artist preferences for where they would like their paintings displayed if a gallery is assigned to them. This can be written as unary factors $g_1, ..., g_m$ where

$$g_i(G_i) = PA^{G_i}_{(i)} * \mathbb{1}[G_i \neq 0] + \mathbb{1}[G_i = 0]$$

which is the preference of artist G_i to have their painting displayed by gallery i (read: value of G_i)

The third set encodes that the each gallery must have a unique painting if it is assigned, as each artist only has one painting. Noting that for all pairs of galleries, their artists must be unique, we can account for this with a set of binary factors for each pair of galleries. For G_i and G_j where $i, j \in \{1, ..., m\}$ and $i \neq j$, we have a factor

$$h_{i,j}(G_i, G_j) = \mathbb{1}[\mathbb{1}[G_i \neq G_j] + \mathbb{1}[G_i = G_j = 0] \neq 0]$$

The final set encodes that each artist must be matched to a gallery. We can encode this as an n-ary factor:

$$j(G_1, ..., G_n) = \mathbb{1}[\{A_1, ..., A_m\} \subseteq \{G_1, ...G_n\}]$$

Note there are several other formulations of the above factors - the above are examples.

b. (*6 points*) Imagine a small setting with 3 artists A_1, A_2, A_3 and 3 galleries G_1, G_2, G_3.
The artist preferences are below:

	G_1	G_2	G_3
$PA^{(1)}$	0	3	0
$PA^{(2)}$	1	2	4
$PA^{(3)}$	3	2	1

The gallery preferences are below:

	A_1	A_2	A_3
$PG^{(1)}$	2	4	0
$PG^{(2)}$	1	4	5
$PG^{(3)}$	3	2	1

Assume that we are modeling the problem using the **first formulation** of the CSP and are using the artists as variables.

Apply the CSP you designed to this small setting and enforce arc-consistency amongst its variables. In particular, write out each variable and its **final** domain after arc consistency has been enforced. For example, if you have a variable X_i with a domain $\{a, b, c\}$, after enforcing arc-consistency, you should write $X_i : \{a, b, c\}$

Solution $A_1 : \{G_2\}$ $A_2 : \{G_1\}$ $A_3 : \{G_3\}$

Also accepted because of lack of clarification that gallery preferences for the artist assigned must be non-zero (although accounting for preferences as a factor implicitly handles this and makes the first case the only legal one.) $A_1 : \{G_2\}$ $A_2 : \{G_1, G_3\}$ $A_3 : \{G_1, G_3\}$

c. (*4 points*) We will now return to the generalized version of the art gallery matching problem using the first formulation.

Circle all of the options below that would be applicable techniques for our art gallery matching CSP if we want a solution that is guaranteed to be the maximum weight solution.

- Least Constrained Value
- Most Constrained Variable
- Iterated Conditional Modes
- Backtracking Search

Solution **Most constrained variable** and **backtracking search** should be circled.
Least constrained value: Not useful - we use LCV when all factors are constraints, which does not hold since we have a factor that encodes preferences.
Most constrained value: Not useful - does not exist.
Most constrained variable: Useful - we use MCV when some factors are constraints, which holds true in this formulation CSP since we have a binary constraint to ensure that the galleries where an artist's painting is displayed must be unique.
Iterated conditional Modes: Not useful - ICM doesn't guarantee finding an optimal solution.
Backtracking search: Useful - Backtracking search does guarantee finding an optimal

solution.

True/False: If we use beam search with different beam sizes k to solve our Art galleries CSP, our solution's assignment weight will always increase as we increase the beam size k. Justification:

Solution False. Consider going from k = 1 (greedy) to k = 2. The greedy solution might be the globally optimal assignment, but when k = 2 we may find more partially optimal solutions as we expand more paths that cause us to drop the greedy solution from our beam. We are only guaranteed a global optimum with an unbounded beam size. A common mistake was to state that the weight will never decrease or the weight will stay the same. This is actually not true even though intuitively it seems like it should be!