1. CSPs & Tournament Design (*30 points*)

In the FIFA World Cup, there are n national soccer teams, $1, \ldots, n$. Each team has a strength $0 \le s_i \le 100$ representing the percentage of matches team i has won in the past year (i.e., how good the team is).

These teams need to be partitioned into $m = \frac{n}{4}$ groups of size 4 (you may assume n is a multiple of 4). Let $X_i \in \{1, \ldots, m\}$ be the group that team i is assigned to.

The strength of an assignment $X = (X_1, \ldots, X_n)$ is the product of the group strengths, where the strength of each group is the sum of the team strengths; formally:

$$\text{strength}(X) = \prod_{j=1}^{m} \sum_{i:X_i=j} s_i.$$

For example, if we have $n = 8$ teams that are split into two groups with strengths $\{10, 20, 30, 40\}$ and $\{50, 60, 70, 80\}$, then assignment strength would be $100 \cdot 260 = 2600$.

The goal is to find the assignment with the maximum strength.

a. (*8 points*)

We can model the assignment of teams to groups with a CSP with n variables, where the weight of an assignment is its strength.

Write down all the factors of the CSP in formal notation. Be precise about how many factors there are.

b. (*4 points*)

For $m = 2$ groups, suppose that it is possible to assign the teams so that both groups have strength C.

Prove that the *maximum weight assignment* has weight C^2.

c. (*6 points*)

Answer the following True/False questions. Clearly circle your choice of True or False. You will receive 1 point for choosing the correct option and 2 more points for a correct one sentence justification of your choice.

(i) **True** / **False**: The least constrained value (LCV) heuristic would be useful for solving our CSP.

(ii) **True** / **False**: The most constrained variable (MCV) heuristic would be useful for solving our CSP.

(iii) **True** / **False**: Backtracking search could be a useful optimization for our CSP.
Note: this question was poorly worded. It should have read "Backtracking search would be useful for solving our CSP", as backtracking search is not an optimization. We were very lenient with answers for this one.

d. (*4 points*)

Suppose team 1 corresponds to the nation hosting the World Cup. Introduce a set of binary constraints enforcing that team 1 is at least as strong as any other team in its group. Note that each factor should have two arguments and m is not necessarily 2, as in (b).

e. (*8 points*)

Suppose we have 8 teams and 2 groups, and the the host advantage constraint from (d) applies in addition to all the constraints mentioned at the beginning. Only team USA, the host, has been assigned to a group:

Team	Strength	Group Assignment
USA (host)	60	1
Iran	50	
France	100	
Brazil	90	
Japan	60	
Morocco	80	
UK	70	
Senegal	50	

What does the AC-3 algorithm output? In particular, write out each variable and its domain after arc-consistency has been enforced. For example, if you have a variable X_{Iran} with a domain $\{1, 2\}$, you should write

$$X_{\text{Iran}} : \{1, 2\}$$

.

2. Two Coins (*53 points*)

Consider a gambling machine with two coins X and Y inside.

At each time step t, the machine secretly selects a coin $C_t \in \{X, Y\}$ with probability $\lambda \in (0, 1)$ of being the same selection as time $t - 1$. The machine then tosses the selected coin. You can observe whether the result $O_t \in \{H, T\}$ is Head (H) or Tail (T), but you can't observe which coin was selected.

At time $t = 1$, the machine starts by selecting coin X with probability $\lambda_0 \in (0, 1)$ and coin Y with probability $1 - \lambda_0$. The two coins show head with probability $p_X, p_Y \in (0, 1)$. You decide to model this gambling machine as an HMM.

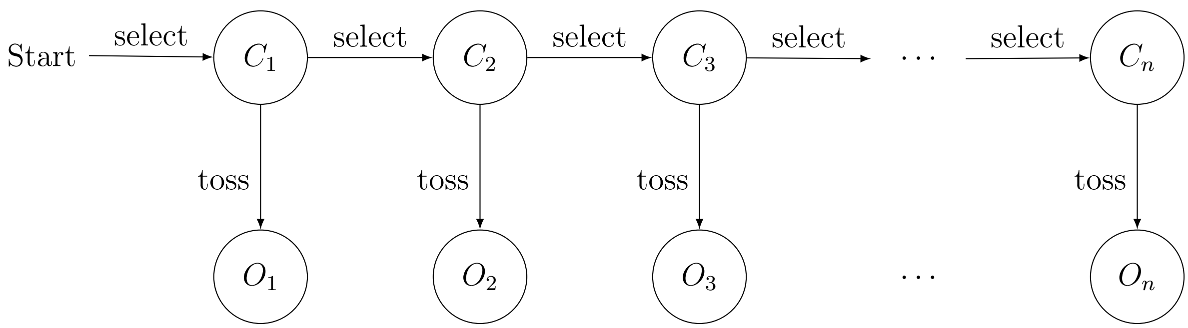

Figure 1: An HMM for the gambling machine.

a. (*8 points*) **Setup**

(i) [4 points] Fill in the table for the transition probability $P(C_{t+1}|C_t)$:

C_t \ C_{t+1}	X	Y
X		
Y		

(ii) [4 points] Fill in the table for the emission probability $P(O_t|C_t)$:

C_t \ O_t	H	T
X		
Y		

b. (*25 points*) **Inference**

You wish to use the forward-backward algorithm to guess which coin was selected at each step. Note that in this part, we assume that all the HMM parameters (i.e. $\lambda_0, \lambda, p_X, p_Y$) are known to you.

Let's play with the toy scenario where the machine stops after two steps. You observe that the results of the two tosses are $\{H, H\}$.

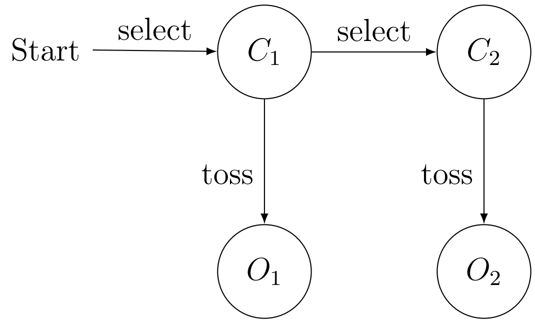

Figure 2: A toy scenario: The machine stops after two steps

(i) [4 points] Please draw the lattice representation of this two-step HMM (No need to label the edge weights).

(ii) [4 points] What is the weight of the edge going from $C_1 = X$ to $C_2 = Y$ in the lattice representation? Your answer should be expressed in terms of one or more parameters from $\{\lambda_0, \lambda, p_X, p_Y\}$.

(iii) [12 points] You start to run the forward-backward algorithm by computing the values of the forward messages $F_i(c)$ and backward messages $B_i(c)$ (where, $i \in \{1, 2\}$ indexes the time step, and $c \in \{X, Y\}$ denotes the coin selection).

① [4 points] Please write down the values of $F_1(X), F_1(Y), B_2(X), B_2(Y)$.

② [8 points] Please write down the values of $F_2(X), F_2(Y), B_1(X), B_1(Y)$ in terms of $F_1(c_1), B_2(c_2)$ and the weights of the edges going from $C_1 = c_1$ to $C_2 = c_2$ (denoted as $w(c_1, c_2)$), where $c_1, c_2 \in \{X, Y\}$.

(iv) [5 points] The gambling machine now asks you to guess which coin was picked at time $t = 1$. Given the current observation, under which condition should you predict coin X as the answer? Formulate this condition with an inequality using $F_i(\cdot)$s and $B_i(\cdot)$s ($i = 1, 2$). Briefly justify your answer.

c. (*20 points*) **Learning**

Now assume that you know how the machine works (i.e. λ_0, λ are known), but you have no idea about the coins (i.e. p_X, p_Y are unknown). You want to use the EM algorithm to estimate p_X and p_Y.

(i) [4 points] Which one of the following algorithms can NOT be used for posterior inference in the E-step? No justification required.

(a) Gibbs sampling

(b) Forward-backward algorithm

(c) Maximum likelihood

(d) Particle filtering

(ii) [12 points] Let's consider the scenario where the machine stops after three steps. You observe that the results of the three tosses are $o = \{H, H, T\}$.

① [4 points] The E-step computes the posterior probability $q_i(c), i \in \{1, 2, 3\}, c \in \{X, Y\}$ of the coin selections. Please rewrite $q_2(X)$ as a conditional probability (Hint: No computation needed. Your answer should be a 1-line equation that includes the observation o and the parameters p_X, p_Y). No justification required.

② [8 points] In one EM iteration, suppose you already get the values of $q_i(c)$ from the E-step (shown in the Table below). What is the value of p_X and p_Y after the M-step? Show your work.

c	$q_1(c)$	$q_2(c)$	$q_3(c)$
X	0.1	0.5	0.3
Y	0.9	0.5	0.7

(iii) [4 points] The gambling machine asks you to guess which coin was picked at time $t = 1$. How will you utilize the result of the EM algorithm to make your prediction? Describe your answer with 1-2 sentences (Hint: No extra computation needed).