# 18-462/662 Assignment 4

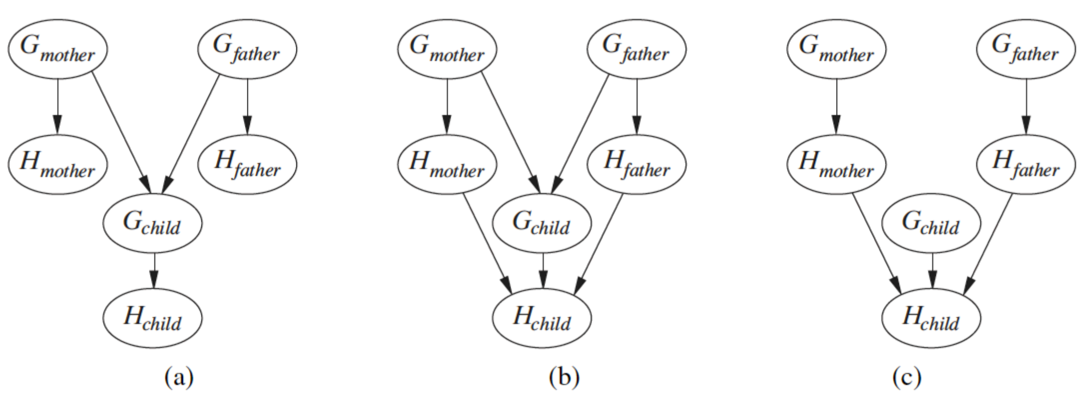
**Carnegie Mellon University**

Department of Electrical and Computer Engineering

{Due Date : April 25th , 6 PM EST}

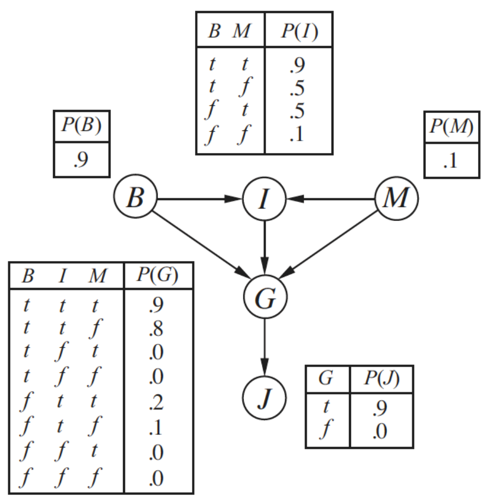
1. We have a bag of three biased coins a, b, and c with probabilities of coming up heads of 20%, 60%, and 80%, respectively. One coin is drawn randomly from the bag (with equal likelihood of drawing each of the three coins), and then the coin is flipped three times to generate the outcomes $X_1$, $X_2$, and $X_3$.
   a. Draw the Bayesian network corresponding to this setup and define the necessary CPTs. [Conditional Probability Table]
   b. Calculate which coin was most likely to have been drawn from the bag
   c. if the observed flips come out heads twice and tails once.

2. Let $H_x$ be a random variable denoting the handedness of an individual $x$, with possible values $l$ or $r$. A common hypothesis is that left- or right-handedness is inherited by a simple mechanism; that is, perhaps there is a gene $G_x$, also with values $l$ or $r$, and perhaps actual handedness turns out mostly the same (with some probability $s$) as the gene an individual possesses. Furthermore, perhaps the gene itself is equally likely to be inherited from either of an individual's parents, with a small nonzero probability $m$ of a random mutation flipping the handedness.
   a. Which of the following three networks claim that $P(G_{father}, G_{mother}, G_{child}) = P(G_{father}) P(G_{mother}) P(G_{child})$?



b. Which of the three networks make independence claims that are consistent with the hypothesis about the inheritance of handedness?
c. Which of the three networks is the best description of the hypothesis?
d. Write down the CPT for the $G_{child}$ node in network (a), in terms of $s$ and $m$.
e. Suppose that $P(G_{father} = l) = P(G_{mother} = l) = q$. In network (a), derive an expression for $P(G_{child} = l)$ in terms of $m$ and $q$ only, by conditioning on its parent nodes.
f. Under conditions of genetic equilibrium, we expect the distribution of genes to be the same across generations. Use this to calculate the value of $q$, and, given what you know about handedness in humans, explain why the hypothesis described at the beginning of this question must be wrong.

3. Consider the Bayes net shown in the figure below.




| $P(B)$ |
|---|
| .9 |

| $B$ | $M$ | $P(I)$ |
|---|---|---|
| $t$ | $t$ | .9 |
| $t$ | $f$ | .5 |
| $f$ | $t$ | .5 |
| $f$ | $f$ | .1 |

| $P(M)$ |
|---|
| .1 |

| $B$ | $I$ | $M$ | $P(G)$ |
|---|---|---|---|
| $t$ | $t$ | $t$ | .9 |
| $t$ | $t$ | $f$ | .8 |
| $t$ | $f$ | $t$ | .0 |
| $t$ | $f$ | $f$ | .0 |
| $f$ | $t$ | $t$ | .2 |
| $f$ | $t$ | $f$ | .1 |
| $f$ | $f$ | $t$ | .0 |
| $f$ | $f$ | $f$ | .0 |

| $G$ | $P(J)$ |
|---|---|
| $t$ | .9 |
| $f$ | .0 |

**Figure** A simple Bayes net with Boolean variables $B = BrokeElectionLaw$, $I = Indicted$, $M = PoliticallyMotivatedProsecutor$, $G = FoundGuilty$, $J = Jailed$.

a. Which of the following are asserted by the network structure?
   i. $\mathbf{P}(B, I, M) = \mathbf{P}(B)\mathbf{P}(I)\mathbf{P}(M)$.
   ii. $\mathbf{P}(J \mid G) = \mathbf{P}(J \mid G, I)$.
   iii. $\mathbf{P}(M \mid G, B, I) = \mathbf{P}(M \mid G, B, I, J)$.
b. Calculate the value of $P(b, i, \neg m, g, j)$.

c. Calculate the probability that someone goes to jail given that they broke the law, have been indicted, and face a politically motivated prosecutor.

d. A context-specific independence allows a variable to be independent of some of its parents given certain values of others. In addition to the usual conditional independences given by the graph structure, what context-specific independences exist in the Bayes net in Figure?

e. Suppose we want to add the variable *P =PresidentialPardon* to the network; draw the new network and briefly explain any links you add.

4. In this question, we examine what happens to the probabilities in the umbrella world in the limit of long time sequences.

   a. Suppose we observe an unending sequence of days on which the umbrella appears. Show that, as the days go by, the probability of rain on the current day increases monotonically toward a fixed point. Calculate this fixed point.

   b. Now consider *forecasting* further and further into the future, given just the first two umbrella observations. First, compute the probability $P(r_{2+k}|u_1, u_2)$ for $k=1. . . 20$ and plot the results. You should see that the probability converges towards a fixed point. Prove that the exact value of this fixed point is 0.5.

5. Often, we wish to monitor a continuous-state system whose behavior switches unpredictably among a set of k distinct "modes." For example, an aircraft trying to evade a missile can execute a series of distinct maneuvers that the missile may attempt to track. A Bayesian network representation of such a switching Kalman filter model is shown in Figure below.

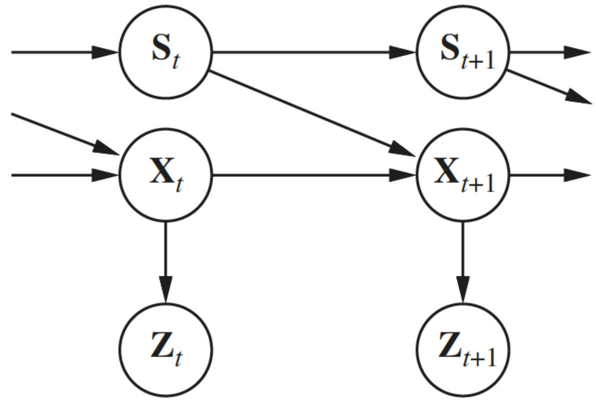



**Figure** A Bayesian network representation of a switching Kalman filter. The switching variable $S_t$ is a discrete state variable whose value determines the transition model for the continuous state variables $\mathbf{X}_t$. For any discrete state $i$, the transition model $\mathbf{P}(\mathbf{X}_{t+1}|\mathbf{X}_t, S_t = i)$ is a linear Gaussian model, just as in a regular Kalman filter. The transition model for the discrete state, $\mathbf{P}(S_{t+1}|S_t)$, can be thought of as a matrix, as in a hidden Markov model.

a. Suppose that the discrete state $S_t$ has $k$ possible values and that the prior continuous state estimate $\mathbf{P}(\mathbf{X}_0)$ is a multivariate Gaussian distribution. Show that the prediction $\mathbf{P}(\mathbf{X}_1)$ is a mixture of Gaussians—that is, a weighted sum of Gaussians such that the weights sum to 1.

b. Show that if the current continuous state estimate $\mathbf{P}(\mathbf{X}_t|\mathbf{e}_{1:t})$ is a mixture of $m$ Gaussians, then in the general case the updated state estimate $\mathbf{P}(\mathbf{X}_{t+1}|\mathbf{e}_{1:t+1})$ will be a mixture of $km$ Gaussians.

c. What aspect of the temporal process do the weights in the Gaussian mixture represent?

The results in (a) and (b) show that the representation of the posterior grows without limit even for switching Kalman filters, which are among the simplest hybrid dynamic models.

6. A professor wants to know if students are getting enough sleep. Each day, the professor observes whether the students sleep in class, and whether they have red eyes. The professor has the following domain theory:

a. The prior probability of getting enough sleep, with no observations, is 0.7.
b. The probability of getting enough sleep on night t is 0.8 given that the student got enough sleep the previous night, and 0.3 if not.
c. The probability of having red eyes is 0.2 if the student got enough sleep, and 0.7 if not.
d. The probability of sleeping in class is 0.1 if the student got enough sleep, and 0.3 if not.

Formulate this information as a dynamic Bayesian network that the professor could use to filter or predict from a sequence of observations. Then reformulate it as a hidden Markov model that has only a single observation variable. Give the complete probability tables for the model.

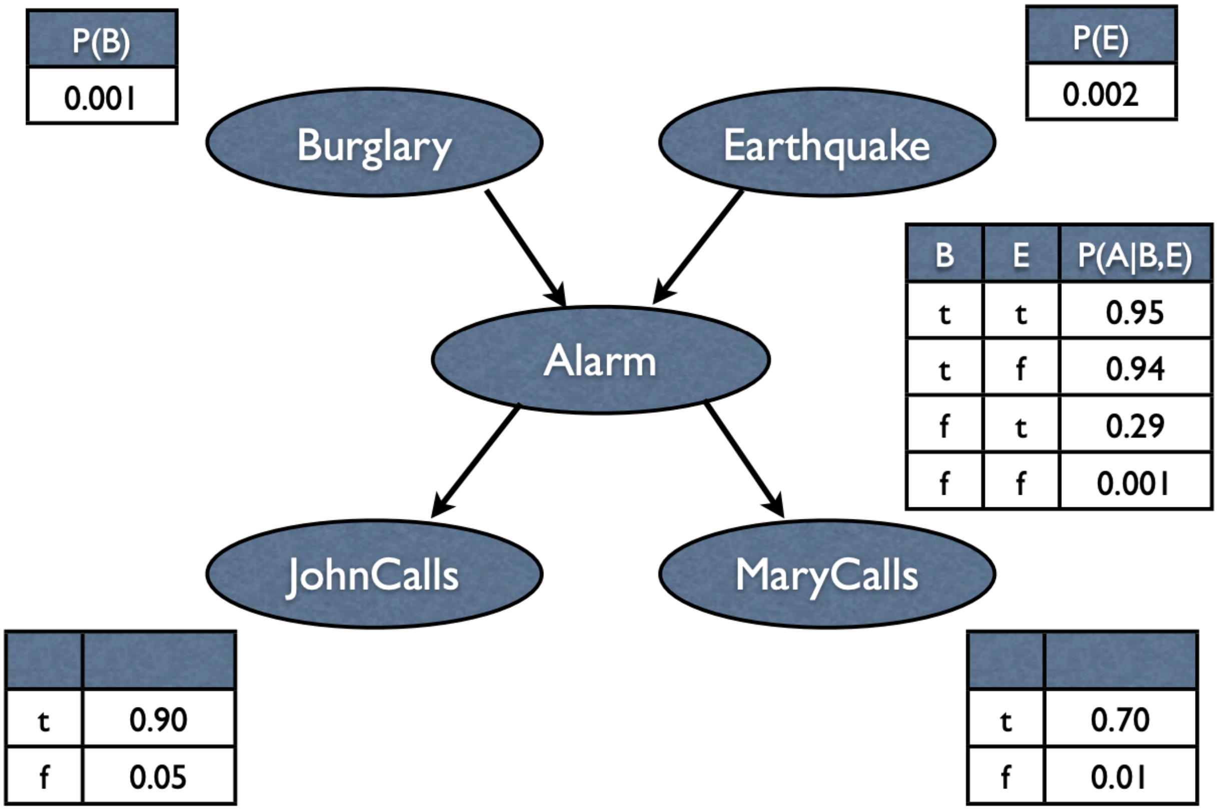
7. Model a moving object following a simple track as given in the following function:

$$f(t) = 0.1 * (t^2 - t)$$

Here 't' is time in seconds, $f(t)$ is position (from origin) in m.
Write a clearly commented Python Code using Kalman Filtering that tracks the above object for $t = 0$ to 100 s. Plot your results with x-axis representing time(seconds) and y-axis representing position(meters). Make use of following information in your code:

- Velocity magnitude is 2 m/s
- Standard deviation of acceleration is 0.25 m/s$^2$
- Position will be estimated every 0.1 s
- Standard deviation of measurement is 1.2 m

8. Implement the following Bayesian network in Python.




Make use of your Python implementation to determine the following probabilities. Please comment your Python code clearly, explaining different parts of your code.

- What's the probability that an alarm has sounded, there was neither an earthquake nor a burglary, and both John and Mary called?
- What's the probability that the alarm doesn't sound given that there was an earthquake, but no burglary?
- What's the probability that both burglary and earthquake occurred, alarm sounded and only John called?

9. Assume the weather operates as a discrete Markov Chain in this problem. There are 2 different weather states: “Rainy” and “Sunny”. There is a graduate student “A” in Carnegie Mellon University and he/she visits the campus everyday using different modes of transportation: “Walk”, “Bike”, “Bus”. Due to some reason, suppose we cannot observe the weather directly, but we can observe A’s behavior. Therefore,

**states** = ('Rainy', 'Sunny')
**observations** = ('Walk', 'Bike', 'Bus')

**start_probability** = {'Rainy': 0.3, 'Sunny': 0.7}

**transition_probability** = {
'Rainy’: {'Rainy': 0.8, 'Sunny': 0.2},
'Sunny’: {'Rainy': 0.3, 'Sunny': 0.7},
}

**emission_probability** = {
'Rainy’: {'Walk': 0.1, 'Bike': 0.2, 'Bus': 0.7},
'Sunny’: {'Walk': 0.4, 'Bike': 0.5, 'Bus': 0.1},
}

It means, for day 0, there is 0.3 possibility to rain, and 0.7 possibility to be sunny.
If day t is a rainy day, then day t+1 has 0.7 possibility to continue raining while it has 0.3 possibility to be sunny.
If it is a rainy day, *A* will have different possibility shown in emission probability to use different modes of transportation. Please note that the term “emission probability” above is the same thing as the “sensor model” covered in lectures when describing Hidden Markov Models (HMM). Suppose we have the following data set:

```
data = ['Bike', 'Bus', 'Bike', 'Bus', 'Walk', 'Bike', 'Bus', 'Walk', 'Bike', 'Bike',
        'Bus', 'Bus', 'Bus', 'Bike', 'Walk', 'Bus', 'Walk', 'Walk', 'Bus', 'Walk',
        'Bike', 'Bike', 'Walk', 'Bike', 'Bike', 'Bus', 'Bike', 'Bus', 'Bike', 'Bus']
```

Write a python script to solve the problems below:

a) **Filtering:**

Given the observations in data (which has the data for 30 days, t0 - t29), what is the probability for day 29 (t29) to be a rainy day?

b) **Smoothing:**

- Given the observations in data (which has the data for 30 days, t0 - t29), what is the probability for day 10 (t10) to be a rainy day?
- Given the observations in data (which has the data for 30 days, t0 - t29), what is the probability for day 20 (t20) to be a rainy day?

c) **Prediction:**

- Given the observations in data (which has the data for 30 days, t0 - t29), what is the probability for day 30 (t30) to be a rainy day?
- Given the observations in data (which has the data for 30 days, t0 - t29), what is the probability for day 40 (t40) to be a rainy day?

d) **Most likely sequence:**

Given the observations in data (which has the data for 30 days, t0 - t29), what is the Most likely sequence for weather?

Please make sure that you comment your Python code clearly for each part.