Lecture 4	Jan. 11-2006

Computational Inference

STAT 440 / 840, CM 461

Lecture: Ali Ghodsi Scribes: Alexandra Laflamme

1 Canonical Graphs

Recall that to derive the *Bayes Ball Algorithm* we have been studying the conditional independence properties of three canonical graphs. We have already analyzed the first two types:

1.1 Markov Chain

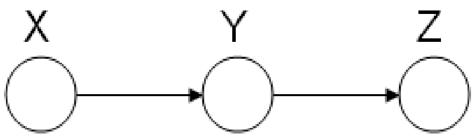

Figure 1: Markov Chain

Conditional Independence Property: $X \amalg Z|Y$

1.2 Hidden Cause

Conditional Independence Property: $X \amalg Z|Y$

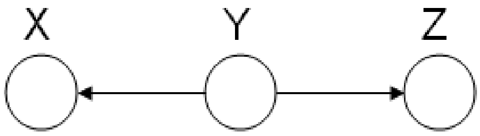

Figure 2: Hidden Cause Graph

1.3 Explaining-Away

Finally, we look at the third type of canonical graphs: *Explaining-Away Graphs.* These types of graph arise when a phenomena has multiple explanations. Here, the conditional independence statement is actually a statement of marginal independence: $X \perp\!\!\!\perp Z$.

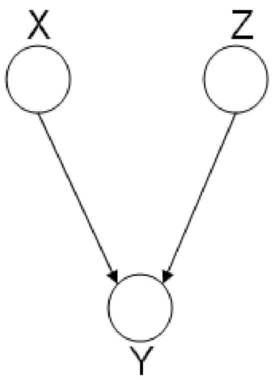

Figure 3: The missing edge between node X and node Z implies that there is a marginal independence between the two: $X \perp\!\!\!\perp Z$

In these types of scenarios, variables X and Z are independent. However, once the third variable Y is observed, X and Z become dependent (Fig.3).

To clarify these concepts, suppose Bob and Mary are supposed to meet for a noontime lunch. Consider the following events:

$$late = \begin{cases} 1, & \text{if Mary is late} \\ 0, & \text{otherwise} \end{cases}$$

$$aliens = \begin{cases} 1, & \text{if aliens kidnapped Mary} \\ 0, & \text{otherwise} \end{cases}$$

$$watch = \begin{cases} 1, & \text{if Bob's watch is incorrect} \\ 0, & \text{otherwise} \end{cases}$$

If Mary is late, then she could have been kidnapped by aliens. Alternatively, Bob may have forgotten to adjust his watch for daylight savings time, making him early. Clearly, both of these events are independent. Now, consider the following probabilities:

$$P(late = 1) \tag{1}$$

$$P(aliens = 1 \mid late = 1) \tag{2}$$

$$P(aliens = 1 \mid late = 1, watch = 0) \tag{3}$$

We expect $P(late = 1) < P(aliens = 1 \mid late = 1)$ since $P(aliens = 1 \mid late = 1)$ does not provide any information regarding Bob's watch. Similarly, we expect $P(aliens = 1 \mid late = 1) < P(aliens = 1 \mid late = 1, watch = 0)$. Since $Eq.2 \neq Eq.3$, *aliens* and *watch* are not independent given *late*. To summarize,

- If we do not observe *late*, then $aliens \perp\!\!\!\perp watch$ ($X \perp\!\!\!\perp Z$)
- If we do observe *late*, then $aliens \not\perp\!\!\!\perp watch \mid late$ ($X \not\perp\!\!\!\perp Z \mid Y$)

2 Bayes Ball Algorithm

Goal: We wish to determine whether a given conditional statement such as $X_A \perp\!\!\!\perp X_B \mid X_C$ is true given a directed graph.

The algorithm is as follows:

1. Shade nodes, X_C, that are conditioned on.
2. If the ball cannot reach X_B, then the nodes X_A and X_B must be conditionally independent.
3. If the ball can reach X_B, then the nodes X_A and X_B are not necessarily independent.

The biggest challenge in the *Bayes Ball Algorithm* is to determine what happens to a ball going from node X to node Z as it passes through node Y. The ball could continue its route to Z or it could be blocked. It is important to note that the balls are allowed to travel in any direction, independent of the direction of the edges in the graph.

We use the canonical graphs previously studied to determine the route of a ball traveling through a graph. Using these three graphs we establish base rules which can be extended upon for more general graphs.

2.1 Markov Chain

A ball traveling from X to Z or from Z to X will be blocked at node Y if this node is shaded. Alternatively, if Y is unshaded, the ball will pass through.

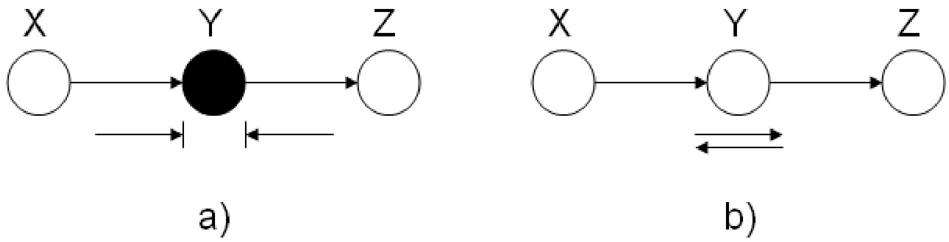

Figure 4: (a) When the middle node is shaded, the ball is blocked. (b) When the middle ball is not shaded, the ball passes through Y.

In Fig.4(a), X and Z are conditionally independent ($X \perp\!\!\!\perp Z \mid Y$) while in Fig.4(b) X and Z are not necessarily independent.

2.2 Hidden Cause

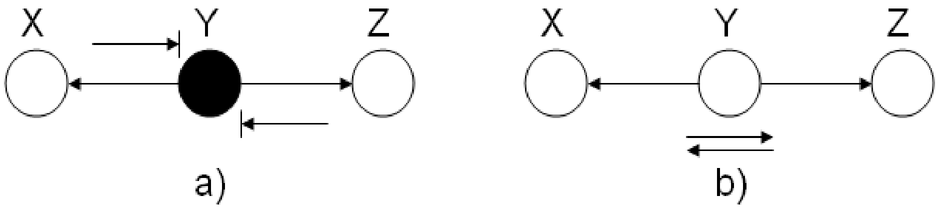

Figure 5: (a) When the middle node is shaded, the ball is blocked. (b) When the middle ball is not shaded, the ball passes through Y.

A ball traveling through Y will be blocked at Y if it is shaded. If Y is unshaded, then the ball passes through.

Fig.5(a) demonstrates that X and Z are conditionally independent when Y is shaded.

2.3 Explaining-Away

A ball traveling through Y is blocked when Y is unshaded. If Y is shaded, then the ball passes through. Hence, X and Z are conditionally independent when Y is unshaded.

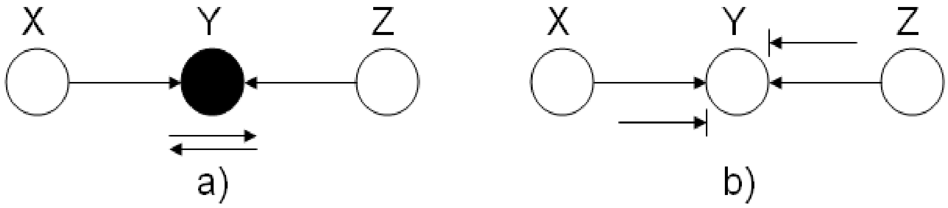

Figure 6: (a) When the middle node is shaded, the ball passes through Y. (b) When the middle ball is unshaded, the ball is blocked.

2.3.1 Example

In this first example, we wish to identify the behavior of a ball going from X to Y in two-node graphs.

The four graphs in Fig.7 show different scenarios. In (a), the ball is blocked at Y. In (b) the ball passes through Y. In both of these cases, we use the rules of the *Explaining Away Canonical Graph* (refer to Fig.6). Finally, for the last two graphs, we used the rules of the *Hidden Cause Canonical Graph* (Fig.5). In (c), the ball passes through Y while in (d), the ball is blocked at Y.

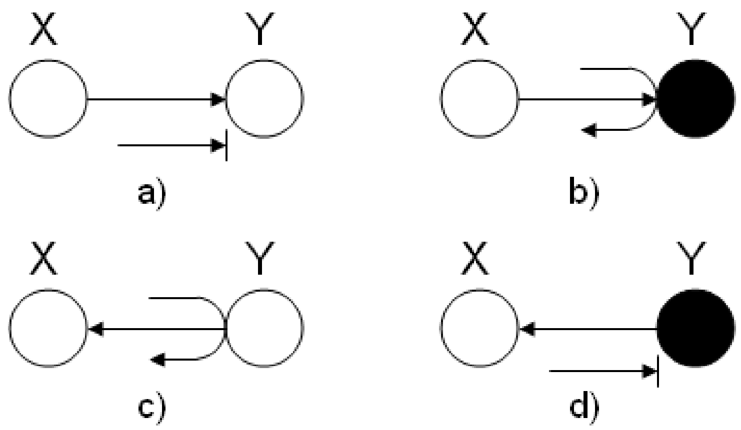

Figure 7: (a)The ball is blocked at Y. (b)The ball passes through Y . (c)The ball passes through Y. (d) The ball is blocked at Y.

2.3.2 Example

Suppose your home is equipped with an alarm system. There are two possible causes for the alarm to ring:

- Your house is being burglarized
- There is an earthquake

Hence, we define the following events:

$$burglary = \begin{cases} 1, & \text{if your house is being burglarized} \\ 0, & \text{if your house is not being burglarized} \end{cases}$$

$$earthquake = \begin{cases} 1, & \text{if there is an earthquake} \\ 0, & \text{if there is no earthquake} \end{cases}$$

$$alarm = \begin{cases} 1, & \text{if your alarm is ringing} \\ 0, & \text{if your alarm is off} \end{cases}$$

$$report = \begin{cases} 1, & \text{if a police report has been written} \\ 0, & \text{if no police report has been written} \end{cases}$$

The *burglary* and *earthquake* events are independent if the alarm does not ring. However, if the alarm does ring, then the *burglary* and the *earthquake* events are not necessarily independent. Also, if the alarm rings then it is possible for a police report to be issued.

We can use the *Bayes Ball Algorithm* to deduce conditional independence properties from the graph. Firstly, consider figure 8(a) and assume we are trying to determine whether there is conditional independence between the *burglary* and *earthquake* events. In figure 8(a), a ball starting at the *burglary* event is blocked at the *alarm* node.

Nonetheless, this does not prove that the *burglary* and *earthquake* events are independent. Indeed, figure 8(b) disproves this as we have found an alternate path from *burglary* to *earthquake* passing through *report*. It follows that $burglary \not\perp\!\!\!\perp earthquake \mid report$

2.3.3 Example

Referring to figure 9, we wish to determine whether the following conditional probabilities are true:

$$X_1 \perp\!\!\!\perp X_3 \mid X_2 \tag{4}$$

$$X_1 \perp\!\!\!\perp X_5 \mid \{X_3, X_4\} \tag{5}$$

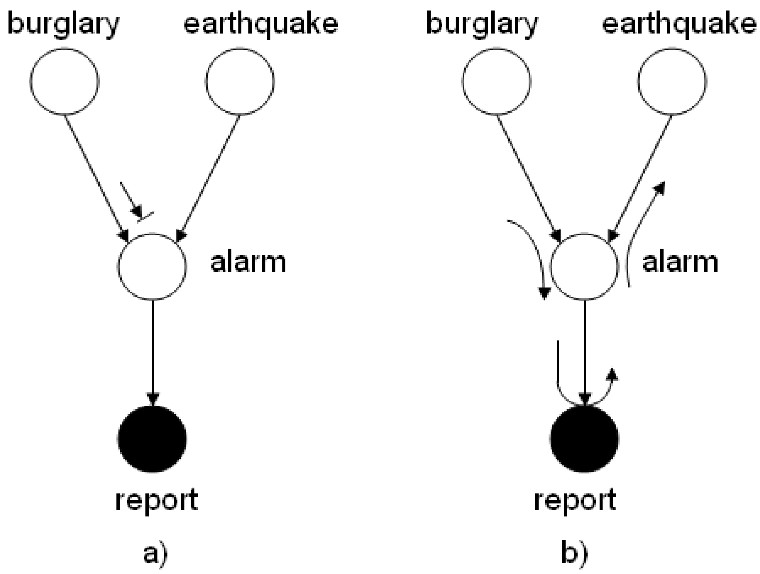

Figure 8: If we only consider the events *burglary*, *earthquake*, and *alarm*, we find that a ball traveling from *burglary* to *earthquake* would be blocked at the *alarm* node. However, if we also consider the *report* node, we can find a path between *burglary* and *earthquake*.

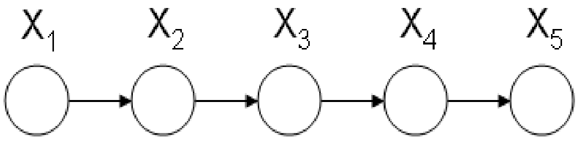

Figure 9: Simple Markov Chain graph

To determine if the conditional probability Eq.4 is true, we shade node X_2. This blocks balls traveling from X_1 to X_3 and proves that Eq.4 is valid.

Similarly, after shading nodes X_3 and X_4 and applying the *Bayes Balls Algorithm* rules, we find that Eq.5 holds.

2.3.4 Example

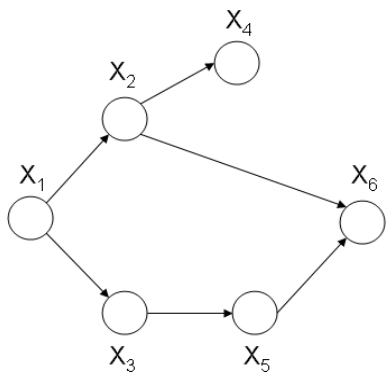

Figure 10: Directed graph

Consider figure 10. Using the *Bayes Ball Algorithm* we wish to determine if each of the following statements are valid:

$$X_4 \perp\!\!\!\perp \{X_1, X_3\} \mid X_2 \tag{6}$$

$$X_1 \perp\!\!\!\perp X_6 \mid \{X_2, X_3\} \tag{7}$$

$$X_2 \perp\!\!\!\perp X_3 \mid \{X_1, X_6\} \tag{8}$$

To disprove Eq.6, we must find a path from X_4 to X_1 and X_3 when X_2 is shaded (Refer to Fig.11(a)). Since there is no route from X_4 to X_1 and X_3 we conclude that Eq.6 is true.

Similarly, we can show that there does not exist a path between X_1 and X_6 when X_2 and X_3 are shaded (Refer to Fig.11(b)). Hence, Eq.7 is true.

Finally, Fig.11(c) shows that there is a route from X_2 to X_3 when X_1 and X_6 are shaded. This proves that the statement 7 is false.

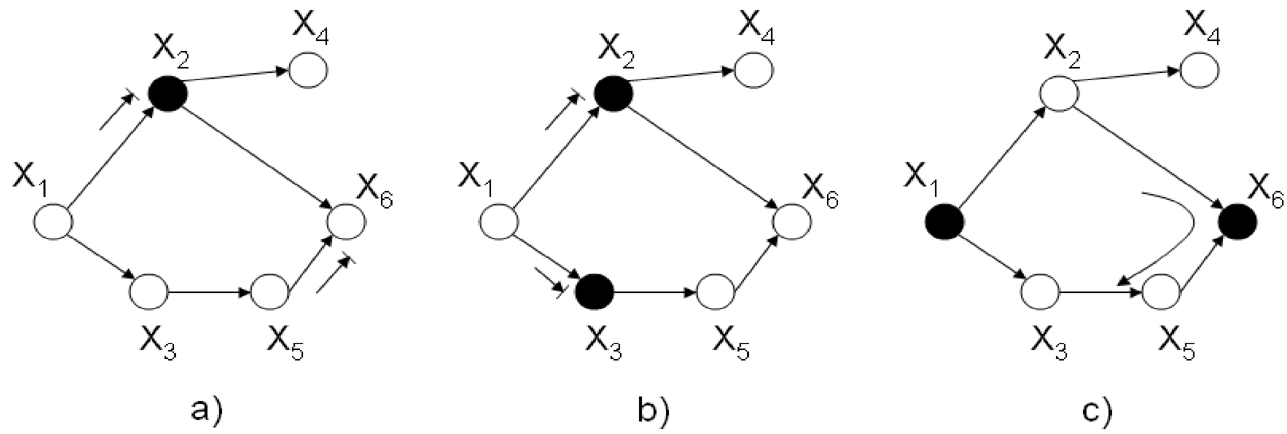

Figure 11: a) A ball cannot pass through X_2 or X_6. b) A ball cannot pass through X_2 or X_3. c) A ball can pass from X_2 to X_3.

Define $p(x_v) = \prod_{i=1}^{n} p(x_i \mid x_{\pi_i})$ to be the factorization as a multiplication of some local probability of a directed graph. Then we have the following theorem:

Theorem: Let $D_1 = \{p(x_v) = \prod_{i=1}^{n} p(x_i \mid x_{\pi_i})\}$ and $D_2 = \{p(x_v) :$ satisfy all conditional independence statements associated with graph $\}$. Then $D_1 = D_2$.