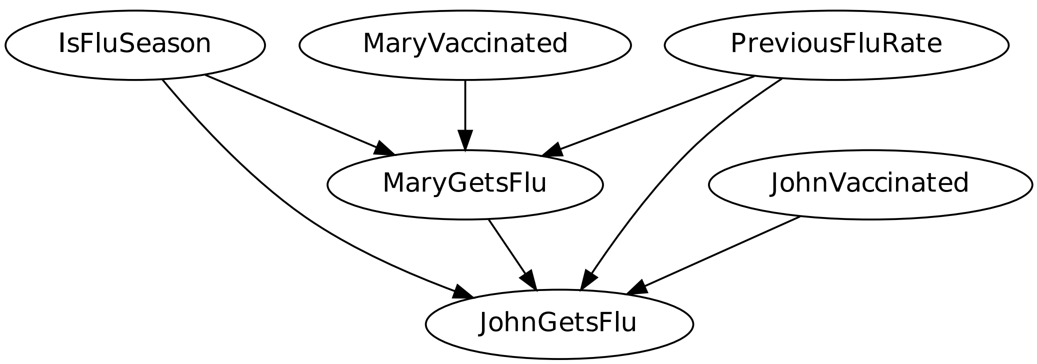
David Snyder Hopkins ID: dsnyde29
CS 476/676 — Fall 2013 **Homework 1b** Due: October 22rd 2013

Homework Notes: Notes

Problem 2.1.1

Figure 1: Drawing of $\mathcal{F}'$.



Problem 2.1.2

The procedure used to produce the drawing above is as follows. List all of the independencies in $\mathcal{F}$. Remove any independence which includes FluRate. Now, the remaining independencies guide our creation of G'.

Data: Bayesian network $\mathcal{F}'$ and a node to remove X
Result: Bayesian network $\mathcal{F}'$ such that X is eliminated
initialization;
$\mathcal{F}' = \{\}$;
for i *in* $Imap(\mathcal{F})$ **do**
  if X *not in* i **then**
    $Imap(\mathcal{F}') \leftarrow Imap(\mathcal{F}') \cup \{i\}$;
  else
  end
end
for *Nodes* n *and* $m \in Z$**do**
  if *not* $dsep(n; m \mid Z)$ **then**
    if $\exists g \in G$ *such that* $g = edge(n, m)$ *or* $g = edge(m, n)$ **then**
      $edges(G') \leftarrow edges(G') \cup \{g\}$;
    else
    end
    $edges(G') \leftarrow edges(G') \cup \{(n, m)\}$;
  else
end

Problem 2.2.1.1

Suppose that there is no active trail from $\hat{Z}$ to X and there exists some B_1 and B_2 such that $P_{B_1}(X \mid Y) \neq P_{B_2}(X \mid Y)$. Since there is no active trail between $\hat{Z}$ and X, by properties of d-separation, we can conclude that $dsep(X; \hat{Z} \mid Y)$. It must be the case that $P_{B_1}(X \mid Y) \neq P_{B_2}(X \mid Y)$ due to some node N on an active trail U to X given Y. There are two possibilities, either N is also on an active trail with $\hat{Z}$ or the CPD of N_{B_1} and N_{B_2} are different even though $dsep(N; \hat{Z}|Y)$. The first case is false since if N were on an active trail with $\hat{Z}$, then X would also be on U, but there is no active trail from $\hat{Z}$ to X. The second case is also false, since by definition of B_1 and B_2, the CPDs of N_{B_1} and $N_{B_2}$are identical. But this contradicts our assumption, so it must be the case that if there is no active trail from $\hat{Z}$ to X then for all B_1 and B_2 $P_{B_1}(X \mid Y) = P_{B_2}(X \mid Y)$.

Problem 2.2.1.2

Suppose not. Suppose that Z is a requisite node for P(X—Y) and Pb1(X—Y) = Pb2(X—Y) for all B1 and B2. Let B1 and B2 be some particular B1 and B2 such that G consists of the following topology Zhat -¿ Z -¿ X ¡- Y so that there is an active trail from Zhat to X given Y as required by the fact that Z is a requisite node for P(X—Y). Let all CPDs be nonzero for all values of the random variables. Further, let Zhat B1 SOME RELATIONSHIP with Zhat B2 STUFF HAPPENS so we see that pb1(X—Y) not equal pb2(X—Y). This contradiction allows us to negate our premise. We can then conclude that if Z is a requisite node for P(X—Y) then pb1(X—Y) = pb1(X—Y) for all B1 and B2.

Problem 2.2.3.1.a

Let $MB_G(X)$ denote the Markov blanet of a node X in an undirected graph G, whose set of nodes is denoted $\mathcal{X}$. We will show that for any X such that $W = \mathcal{X} - \{X\} - MB_G(X)$ then X and W are separated given $MB_G(X)$.
Suppose that X and W are not separated given $MB_G(X)$. That is, ¬sep(X;Y | $MB_G(X)$). Then there is some path U from a node $N \in W$to X, given $MB_G(X)$. Since this path exists, there must be some $M \in U$ such that M is a neighbor of X. However, by definition of the Markov Blanket, $M \in MB_G(X)$. Thus, it cannot be that $M \in U$, and this extends to all possible M. But then, U cannot exist and so by Contradiction, the negation of our premise must be true, that is, X and W are separated given $MB_G(X)$.

Problem 2.2.3.1.b

Let $MB_G(X)$ denote the Markov blanet of a node X in an undirected graph G, whose set of nodes is denoted $\mathcal{X}$. We will show that for any X such that $W = \mathcal{X} - \{X\} - MB_G(X)$ then X and W are separated given $MB_G(X)$ and $MB_G(X)$ is the minimal set with this property.
Suppose that $MB_G(X)$ is not minimal. Then there exists some neighbor of X $N \in \mathcal{X}$ such that $MB'_G(X) = MB_G - \{N\}$ separates X and W, that is, $sep(X; W \mid MB'_G(X))$. Let $M \in W$ be some node such that there is a path U shared by M and N. Then it must be that $X \in U$, since N is a neighbor of X. Because such a path exists, $\neq sep(X; M \mid MB'_G(X))$, which contradicts our supposition. Therefore, $MB_G(X)$ must be the minimal set with this property.

Problem 2.2.3.2

Let B be a Bayesian network with a graph G. By Definition 4.16 Corollary 4.2, we can create an undirected moral graph from G denoted $M[G]$, such that it is an I-map for P_B. Therefore, we can show the Markov blanket separates a node $X \in B$ from the remaining nodes and that this set is minimal by first converting B into $M[G]$ and repeating the proofs 2.2.3.1.a and 2.2.3.1.b on the resultant Markov network.

Problem 2.4.1.a

HMM Yes. Influence can flow from $S_{3,1}$ to Y_1. We can examine each overlapping triplet of nodes and apply the rules for an active trail. Since there is an active trail, it is indeed possible for $P(Y_1 = severe|S_{3,1} = yes) \leq P(Y_1 = severe|S_{3,1} = no)$.

MEMM No. Again, we examine the potential active trails. We see that due to the v-structure at Y_3 and the fact that Y_3 is not observed, influence cannot flow to Y_1. As a result, $P(Y_1 = severe|S_{3,1} = yes) = P(Y_1 = severe|S_{3,1} = no)$

CRF Yes. Influence can flow $S_{3,1}$ to Y_1 over the undirected edges and so $P(Y_1 = severe|S_{3,1} = yes) \leq P(Y_1 = severe|S_{3,1} = no)$

is possible.

Problem 2.4.1.b

HMM No. This is a causal trail from Y_2 to $S_{3,1}$ which is made d-separated by observing Y_3.

MEMM Yes. In the v-structure at $S_{3,1} \leftarrow Y_3 \rightarrow Y_2$ influence can flow from $S_{3,1}$ to Y_2 since Y_3 is now observed.

CRF No. The active trail is broken by observing Y_3.

Problem 2.4.1.c

In all of the graphs, $S_{3,1}$ becomes d-separated or separated from Y_1 given Y_2. This is evident from the properties of Causal trails in directed graphs, and separation in undirected graphs.

Problem 2.4.2.a

HMM Because of the causal direction of the arrows from Y to S, we can naturally compute $P(S \mid Y)$. To compute $P(Y \mid S)$ we need to be able to obtain $P(S)$ and $P(Y)$ that give us the probability of the symptoms and the disease respectively. If we are experts in Biology then we should be able to construct reasonable models for $P(Y)$ and $P(S)$, making the HMM efficient for situations in which we don't have much data. If we don't have good priors, then we may be better off using one of the other models.

MEMM In this model we can more naturally compute $P(Y|S)$. Since we have causality from the features S to the disease state Y it seems easier to handle large numbers of features. Also, in this model the features are allowed to be dependent given the disease label. This could allow for more flexibile features, such as overlapping features.

CRF It seems that this model has the same advantages as MEMM regarding flexible features. In addition, as observed in the previous questions, it allows for long range dependencies of features through time.

Problem 2.4.2.b

As discussed above, it seems that MEMM and CRFs are better at handling correlated symptom variables (and CRFs are superior to MEMMs at correlated symptoms over time).

Problem 2.4.3.b

Z	x^0, y^0	x^0, y^1	x^1, y^0	x^1, y^1
z^0	.1	.8	.8	.1
z^1	.9	.2	.2	.9

Problem 2.4.3.c

Z is in dependent of Y in the context of $X = x^1$. Z is also independent of X in the context of $Y = y^1$.