Prof. Wolfgang Nejdl
M.Sc. Ambreen Zaina, M.Sc. Jeff Reimer
Department of Knowledge-based Systems (KBS)

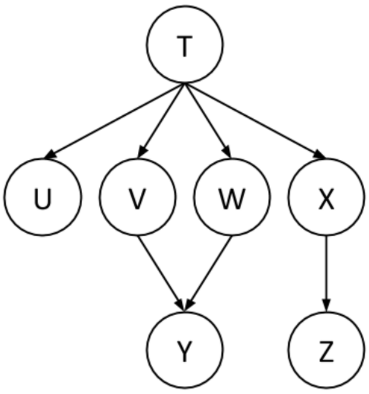
November 18, 2019

Artificial Intelligence II

Exercise 3

Q1. D-Separation

Indicate whether each of the following conditional independence relationships is guaranteed to be true in the Bayes Net below. If the independence relationship does not hold, identify all active (d-connected) paths in the graph.



1. $U \perp\!\!\!\perp X$

   Solution:
   Not guaranteed, path U–T–X is active

2. $U \perp\!\!\!\perp X|T$

   Solution:
   Guaranteed

3. $V \perp\!\!\!\perp W|Y$

Solution:
Not guaranteed, paths V–T–W and V–Y–W are both active

4. $V \perp\!\!\!\perp W|T$

Solution:
Guaranteed

5. $T \perp\!\!\!\perp Y|V$

Solution:
Not guaranteed, path T–W–Y is active

6. $Y \perp\!\!\!\perp Z|W$

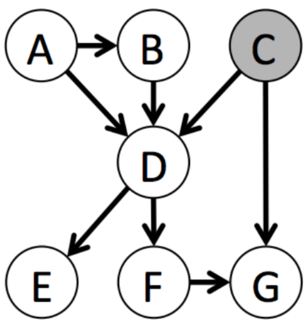
Solution:
Not guaranteed, path Y–V–T–X–Z is active

7. $Y \perp\!\!\!\perp Z|T$

Solution:
Guaranteed, with T being observed, there are no active paths from Y to Z

Q2. Variable Elimination

(a) For the Bayes' net below, we are given the query $P(A; E| + c)$. All variables have binary domains. Assume we run variable elimination to compute the answer to this query with the following variable elimination ordering: B, D, G, F.



Complete the following description of the factors generated in this process:

After inserting evidence, we have the following factors to start out with:

$$P(A), P(B|A), P(+c), P(D|A, B, +c), P(E|D), P(F|D), P(G| + c, F)$$

When eliminating B we generate a new factor f_1 as follows:

$$f_1(A, +c, D) = \Sigma_b P(b|A)P(D|A, b, +c)$$

This leaves us with the factors:

$$P(A), P(+c), P(E|D), P(F|D), P(G| + c, F), f_1(A, +c, D)$$

When eliminating D we generate a new factor f_2 as follows:

Solution:

$f_2(A, +c, E, F) = \Sigma_d P(E|d)P(F|d)f_1(A, +c, d)$

This leaves us with the factors:

Solution:

$P(A), P(+c), P(G| + c, F), f_2(A, +c, E, F)$

When eliminating G we generate a new factor f_3 as follows:

Solution:

$f_3(+c, F) = \Sigma_g P(g| + c, F)$

This leaves us with the factors:

Solution:

$P(A), P(+c), f_2(A, +c, E, F), f_3(+c, F)$

When eliminating F we generate a new factor f_4 as follows:

Solution:
$f_4(A, +c, E) = \Sigma_f f_2(A, +c, E, f) f_3(+c, f)$

This leaves us with the factors:

Solution:
$P(A), P(+c), f_4(A, +c, E)$

(b) Write a formula to compute $P(A, E| + c)$ from the remaining factors.

Solution:
$P(A, E| + c) = \frac{P(A)P(+c)f_4(A,+c,E)}{\Sigma_{a,e} P(a)P(+c)f_4(a,+c,e)}$ or alternatively: $P(A, E| + c)\ \alpha\ P(A)P(+c)f_4(A, +c, E)$ and include statement that says renormalization is needed to obtain $P(A, E| + c)$.

(c) Among f_1, f_2, f_3, f_4 , which is the largest factor generated, and how large is it? Assume all variables have binary domains and measure the size of each factor by the number of rows in the table that would represent the factor.

Solution:
$f_2(A, +c, E, F)$ is the largest factor generated. It has 3 non-instantiated variables, hence $2^3 = 8$ entries.

(d) Find a variable elimination ordering for the same query, i.e., for $P(A, E| + c)$, for which the maximum size factor generated along the way is smallest. Hint: the maximum size factor generated in your solution should have only 2 variables, for a size of $2^2 = 4$ table. Fill in the variable elimination ordering and the factors generated into the table below.

Variable Eliminated	Factor Generated
B	$f_1(A, +c, D)$
G	$f_2(+c, F)$
F	$f_3(+c, D)$
D	$f_4(A, +c, E)$

For example, in the naive ordering we used earlier, the first row in this table would have had the following two entries: B, $f_1(A, +c, D)$.

Solution:

Note: multiple orderings are possible. An ordering is good if it eliminates all non-query variables (B, D, F, G) and its largest factor has only two variables.