Tutorial 8 - Belief Propagation and Sampling

COMP9418 – Advanced Topics in Statistical Machine Learning

Lecturer: Gustavo Batista

Lecture: Belief Propagation and Sampling

Topic: Questions from lecture topics

Last revision: Monday 16th November, 2020 at 15:34

Question 1

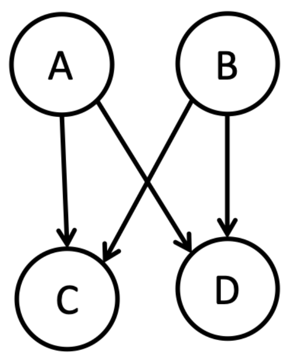
Consider the following Bayesian network:

$P(A)$

+a	.5
-a	.5

$P(B)$

+b	.25
-b	.75



$P(C|A,B)$

+a	+b	+c	.9
+a	+b	-c	.1
+a	-b	+c	.8
+a	-b	-c	.2
-a	+b	+c	.7
-a	+b	-c	.3
-a	-b	+c	.6
-a	-b	-c	.4

$P(D|A,B)$

+a	+b	+d	.6
+a	+b	-d	.4
+a	-b	+d	.4
+a	-b	-d	.6
-a	+b	+d	.3
-a	+b	-d	.7
-a	-b	+d	.3
-a	-b	-d	.7

Suppose we condition on evidence $e : D = true$. Suppose that we have run the Parallel Iterative Belief Propagation (IBP) algorithm on the network, where it converges and yields the following set of messages and family marginals:

A	$\pi_C(A)$	$\pi_D(A)$	$\lambda_C(A)$	$\lambda_D(A)$
+a		.5	.5	.6
-a		.5	.5	.4

B	$\pi_C(B)$	$\pi_D(B)$	$\lambda_C(B)$	$\lambda_D(B)$
+b	.3	.25	.5	
-b	.7	.75	.5	

A	B	C	$\beta(A,B,C)$
+a	+b	+c	.162
+a	+b	-c	.018
+a	-b	+c	
+a	-b	-c	.084
-a	+b	+c	.084
-a	+b	-c	.036
-a	-b	+c	.168
-a	-b	-c	.112

A	B	D	$\beta(A,B,D)$
+a	+b	+d	.2
+a	+b	-d	0
+a	-b	+d	
+a	-b	-d	0
-a	+b	+d	.1
-a	+b	-d	0
-a	-b	+d	.3
-a	-b	-d	0

a. Fill in the missing values for IBP messages.
b. Fill in the missing values for family marginals.
c. Compute marginals $\beta(A)$ and $\beta(B)$ using the IBP messages and those computed in (a).
d. Compute marginals $\beta(A)$ and $\beta(B)$ by summing out the appropriate variables from the family marginals $\beta(ABC)$ as well as $\beta(ABD)$.
e. Compute joint marginal $\beta(AB)$ by summing out the appropriate variables from family marginals $\beta(ABC)$ as well as $\beta(ABD)$.
f. Are the marginals computed in (d) consistent? What about those computed in (e)?

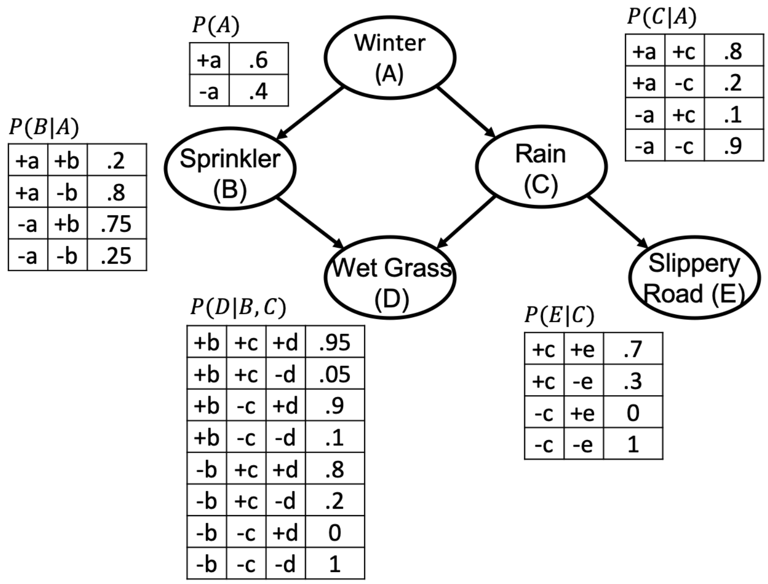
Consider the following IBP algorithm:

Data: N: Bayesian network
Data: $\mathbf{e}$: evidence
Result: Approximate marginals, $\beta(X\mathbf{U})$ of $P(X\mathbf{U}|e)$ for each family $X\mathbf{U}$ in N

begin
  $t \leftarrow 0$;
  initialize all messages;
  while *messages have not converged* **do**
    $t \leftarrow t+1$;
    for *each node X with parents $\boldsymbol{U}$* **do**
      for *each parent U_i* **do**
        $\lambda_X^t(U_i) = \eta \sum_{X\mathbf{U}\setminus\{U_i\}} \lambda_{\mathbf{e}}(X)\phi_X(X,\mathbf{U}) \prod_{k\neq i} \pi_X^{t-1}(U_k) \prod_j \lambda_{Y_j}^{t-1}(X)$;
      end
      for *each child Y_j* **do**
        $\pi_{Y_j}^t(X) = \eta \sum_{\mathbf{U}} \lambda_{\mathbf{e}}(X)\phi_X(X,\mathbf{U}) \prod_i \pi_X^{t-1}(U_i) \prod_{k\neq j} \lambda_{Y_k}^{t-1}(X)$;
      end
    end
  end
end
return $\beta(X\mathbf{U}) = \eta\lambda_{\mathbf{e}}(X)\phi_X(X,\mathbf{U}) \prod_i \pi_X^t(U_i) \prod_j \lambda_{Y_j}^t(X)$

Question 2

Consider the following Bayesian network:



$P(A)$

+a	.6
-a	.4

$P(B|A)$

+a	+b	.2
+a	-b	.8
-a	+b	.75
-a	-b	.25

$P(C|A)$

+a	+c	.8
+a	-c	.2
-a	+c	.1
-a	-c	.9

$P(D|B,C)$

+b	+c	+d	.95
+b	+c	-d	.05
+b	-c	+d	.9
+b	-c	-d	.1
-b	+c	+d	.8
-b	+c	-d	.2
-b	-c	+d	0
-b	-c	-d	1

$P(E|C)$

+c	+e	.7
+c	-e	.3
-c	+e	0
-c	-e	1

and the parameter $\theta_{\bar{a}}$ representing the probability of $A = false$ (i.e., it is not winter). For each of the following values of this parameter, 0.01, 0.4 and 0.99 do the following:

a. Compute the probability of $P(+d, +e)$: wet grass and slippery road. You can use the code from previous tutorials.
b. Estimate $P(+d, +e)$ using forward sampling with sample sizes ranging from $n = 100$ to $n = 15,000$.
c. Generate a plot with n on the x-axis and the exact value of $P(+d, +e)$ and the estimate for $P(+d, +e)$ on the y-axis.
d. Generate a plot with n on the x-axis and the exact variance of the estimate for $P(+d, +e)$ and the sample variance on the y-axis.