Examen sur les Réseaux Bayésiens

mercredi 4 mai 2016

Durée de l'épreuve: 2h. Notes de cours interdites, calculatrice autorisée.
Barême approximatif donné à titre purement indicatif et susceptible de changements.
Les deux exercices peuvent être traités indépendamment l'un de l'autre.

Exercice 1 (12 points)

On considère ici l'ensemble $\mathcal{U} = (X_{1:n}, Y_{1:n}, Z_{1:n}, B_{1:2}, E_{1:2})$ de $3n+4$ variables aléatoires réelles dont la loi jointe est définie par:

$$\mathbb{P}(\mathcal{U}) = \mathbb{P}(X_{1:n})\mathbb{P}(Y_{1:n})\mathbb{P}(Z_{1:n})\mathbb{P}(B_1|X_1,Y_1)\mathbb{P}(B_2|Y_1,Z_1)\mathbb{P}(E_1|X_n,Y_n)\mathbb{P}(E_2|Y_n,Z_n)$$

avec

$$\mathbb{P}(X_{1:n}) = \mathbb{P}(X_1)\prod_{i=2}^{n}\mathbb{P}(X_i|X_{i-1}) \quad \mathbb{P}(Y_{1:n}) = \mathbb{P}(Y_1)\prod_{i=2}^{n}\mathbb{P}(Y_i|Y_{i-1}) \quad \mathbb{P}(Z_{1:n}) = \mathbb{P}(Z_1)\prod_{i=2}^{n}\mathbb{P}(Z_i|Z_{i-1}).$$

1) [**2 points**] Rappeler la définition d'un réseau bayésien et montrer que $\mathbb{P}(\mathcal{U})$ en définit bien un. Représenter graphiquement ce réseau bayésien.

2) [**2 points**] Calculer $\mathbb{P}(X_{1:n}, Y_{1:n}, Z_{1:n}|B_1)$. Quelles relations d'indépendance conditionnellement à B_1 avons nous entre les variables de $(X_{1:n}, Y_{1:n}, Z_{1:n})$?

3) [**3 points**] Dans le cas où $n = 3$, construire un *junction tree* correspondant au réseau bayésien ayant une *tree-width* inférieure à 4. On pourra pour cela utiliser l'ordre d'élimination: $B_1 - B_2 - E_1 - E_2 - X_1 - Z_1 - Y_1 - X_2 - Z_2 - X_3 - Z_3 - Y_{2:3}$.

4) [**1 point**] Enoncer la définition d'un *junction tree* et vérifier que celui que vous avez proposé vérifie bien cette définition.

5) [**2 point**] On suppose que l'ensembles des variables aléatoires de $\mathcal{U}$ sont binaires. On considère l'évidence $\text{ev} = \{B_1 = B_2 = E_1 = E_2 = 1\}$. Calculer en fonction de n la complexité du calcul de $\mathbb{P}(\mathcal{U}, \text{ev})$: a) par la force brute (énumération de l'ensemble des configurations possibles); b) à l'aide de votre *junction tree*.

6) [**2 points**] Choisir une racine à votre *junction tree* et donner (sans preuve) les formules permettant de calculer les 4 premiers messages de la recursion *inward*.

Exercice 2 (8 points)

On considère dans cet exercice $X_{1:n}$ un n-échantillon d'une loi de Poisson de paramètre λ de sorte que $\mathbb{P}(X_i = k) = e^{-\lambda}\lambda^k/k!$. Pour tout i, on note $S_i = X_1 + \ldots + X_i$ la somme partielle. L'objectif de cet exercice est d'étudier la loi $\mathbb{P}(X_{1:n}|S_n = K)$ pour un $K \in \mathbb{N}$ donné.

1) **[2 points]** Calculer cette loi dans pour $K = 0$. Pour $K = 1$. Donner le nombre de configurations de $X_{1:n}$ possibles sous la contrainte $S_n = K$ en fonction de K et de n.

2) **[2 points]** On introduit pour $i = 1, \ldots, n$ les quantités *forward* $F_i(k) = \mathbb{P}(S_i = k)$ et *backward* $B_i(k) = \mathbb{P}(S_n = K|S_i = k)$ (avec la convention $B_n \equiv 1$). Montrer que $\mathbb{P}(S_i = k, S_n = K) = F_i(k)B_i(k)$ et que $\mathbb{P}(S_{i-1} = j, S_i = k, S_n = K) = F_{i-1}(j)\mathbb{P}(X_i = k - j)B_i(k)$.

3) **[2 points]** En déduire des formules récursives permettant le calcul des quantités *forward* et *backward*. Calculer en fonction de n et K la complexité nécessaire au calcul de l'ensemble des quantités *forward/backward*.

4) **[2 points]** Proposer un algorithme permettant de simuler $X_{1:n}$ sous la contrainte $S_n = K$. Illustrer cet algorithme dans le cas particulier où $K = 2$.