

Durée : 1h30

Documents autorisés - Calculatrice autorisée

Votre voisin n'est pas un document

Albert and Bertrand are back**(10 points)**

Albert l'étudiant et son collègue Bertrand sont dans la même pièce et discutent. Comme ils ne savent pas faire deux choses en même temps, ils sont obligés de s'arrêter de parler pour écouter l'autre. Leur vocabulaire est limité à 3 mots *Miam*, *Dodo*, *Wii*.

- Albert parle essentiellement de *Dodo* (à 85%) puis à égalité de *Miam* et *Wii*
- Bertrand parle essentiellement de *Wii* (à 85%) puis de *Miam* (10%) et *Dodo* (5%)
- Albert est souvent à l'écoute de Bertrand : lorsqu'Albert prononce un mot, il y a 2 chances sur 3 qu'il laisse ensuite la parole à Bertrand
- Ce n'est pas le cas de Bertrand qui a tendance à garder la parole à 80%
- La coutume veut qu'Albert commence toujours les discussions.

Un micro placé dans la pièce enregistre cette conversation : *Miam Wii Dodo*

1. (2 points) Dessiner le modèle de Markov caché qui correspond à ce problème, déroulé sur 4 pas de temps. Indiquer quelles sont les tables de probabilité conditionnelle associées à chaque variable du réseau bayésien ainsi obtenu.
2. (6 points) Appliquer l'algorithme de Message Passing pour répondre à la question suivante : Quel est le mot suivant le plus probable ?
3. (2 points) Proposer un arbre de jonction associé au modèle de la question 1 (en détaillant les étapes).

Bonaparte Disaster Victim Identification System**(10 points)**

Bonaparte DVI est un logiciel d'identification de victimes à l'aide d'analyses ADN. Le principe est d'identifier une victime à l'aide de ADN, dans une liste de personnes disparues, à partir de l'ADN de plusieurs membres de la famille de la personne disparue.

Prenons un exemple simple, avec la couleur des yeux, en se restreignant aux parents immédiats. Cette information génétique se trouve "décomposée" en deux exemplaires (les allèles) dans nos cellules. L'un de ces exemplaires vient de notre père (pris aléatoirement parmi les deux siens), l'autre de notre mère (idem). Ces deux informations peuvent être identiques (yeux marrons/ yeux marrons), ou différentes (yeux bleus/ yeux marrons).

Lorsque les deux allèles codent la même information, celle-ci sera "réalisée". Ainsi, si un individu possède deux allèles "yeux marrons", il aura les yeux marrons ; deux allèles "yeux bleus", et il possèdera des yeux bleus. Par contre, quand les deux allèles codent des informations contradictoires, il se peut que l'un soit automatiquement perdant, c'est le cas ici des "yeux bleus". On parle alors de caractère physique récessif "yeux bleus" ou dominant "yeux marrons" : un individu avec deux allèles "bleux / marrons" aura donc les yeux marrons.

1. (5 points) **Structure**
 - (a) (3 points) Proposer une structure de réseau bayésien permettant de modéliser la couleur des yeux d'une personne en fonction de la couleur des yeux de ses parents, via les allèles de cette personne et de ses parents.
 - (b) (2 points) En supposant que le modèle soit déjà fixé, quelles "questions" faut-il poser au réseau bayésien et que faut-il faire avec les résultats obtenus pour identifier une victime d'après les couleurs des yeux des parents de personnes disparues ?
2. (5 points) **Paramètres**
 - (a) (3 points) Détailler les différents paramètres indépendants induits par la structure précédente. Parmi ces différents paramètres, quels sont ceux déjà fixés par l'énoncé ? Quels sont ceux qui sont inconnus mais identiques (i.e. que l'on retrouve à plusieurs endroits dans le réseau) ?
 - (b) (2 points) Proposer plusieurs méthodes permettant de déterminer les paramètres restants.