

Durée : 1h30

Documents autorisés - Calculatrice autorisée

Votre voisin n'est pas un document

Learning Analytics

Dans cet exercice¹, nous allons prendre le point de vue d'un enseignant qui veut déterminer si un étudiant a compris une matière, en se basant sur la note de l'examen. La figure 1 propose un réseau bayésien pour cette tâche. Comme vous pouvez le voir, le fait que l'étudiant obtienne une note élevée (E , *high Exam score*) est influencé à la fois par le fait qu'il a l'habitude de passer des examens (T , *good Test taker*), et qu'il a compris la matière (U , *Understand material*). Ces deux variables sont à leur tour influencées par l'intelligence de l'étudiant (I , *Intelligent*). La compréhension de la matière est aussi influencée par le fait que l'étudiant travaille beaucoup (H , *Hardworking*).

Toutes les variables sont booléennes, les valeurs d'une variable V étant notées $+v$ et $-v$.

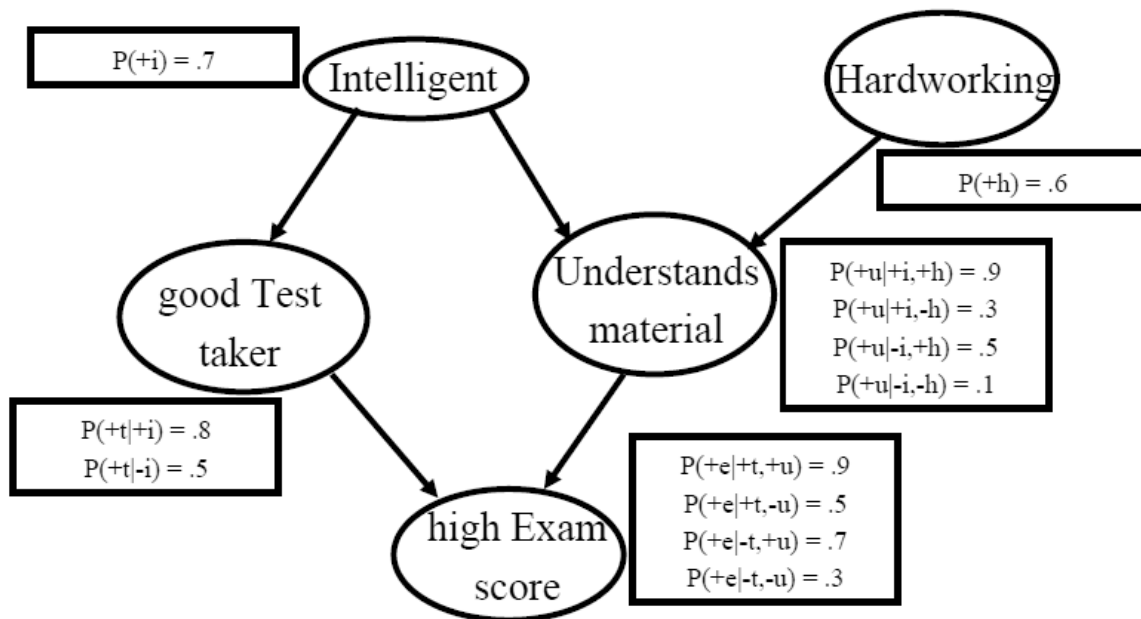


FIGURE 1 – Réseau bayésien. Seules les valeurs "indépendantes" des tables de probabilité sont données, les autres s'obtenant par complément à 1.

1. (5 points) Modèle de dépendances probabilistes

En ne considérant que l'information fournie par le graphe du réseau bayésien, répondez (par VRAI ou FAUX) aux affirmations suivantes :

- T et U sont indépendants.
- T et U sont indépendants conditionnellement à I , E et H .
- T et U sont indépendants conditionnellement à I et H .
- E et H sont indépendants conditionnellement à U .
- E et H sont indépendants conditionnellement à U , I et T .
- I et H sont indépendants conditionnellement à E .
- I et H sont indépendants conditionnellement à T .
- T et H sont indépendants.
- T et H sont indépendants conditionnellement à E .
- T et H sont indépendants conditionnellement à E et U .

1. Inspiré de CSCI 4150 Introduction to Artificial Intelligence, 2014 Spring, Homework 1 : Probability and Bayesian networks, Rensselaer Polytechnic Institute, Lirong Xia

2. (5 points) **Inférence probabiliste**

- (a) (2 points) Peut-on appliquer l'algorithme de Message Passing à ce réseau ? Pourquoi ?
- (b) (3 points) Construire l'arbre de jonction correspondant à ce réseau bayésien.

3. (10 points) **Apprentissage**

- (a) (4 points) Les tables de probabilité de la figure 1 ont été déterminées à partir de données provenant d'étudiants et d'informations expertes provenant d'enseignants. Quelle méthode d'estimation des paramètres a pu être utilisée pour prendre en compte ces deux types d'information ? Donnez la formule mathématique correspondante. Donnez un exemple de données et d'avis d'experts qui auraient pu mener à aux probabilités $P(T|I)$ données dans la figure 1.
- (b) (3 points) Imaginons que la structure de ce réseau bayésien ait été obtenue par un algorithme d'apprentissage à partir de scores de type recherche gloutonne (*Greedy Search*). Quelle est la dimension de cette structure ? Quel autre graphe équivalent aurait-on pu obtenir avec exactement le même score ? Quelle est la dimension de cette structure équivalente ?
- (c) (3 points) Quels sont les graphes "voisins" du graphe de la figure 1 ? Quelle est la dimension de chacun de ses voisins ?

4. (3 points) **(Bonus) Extensions**

Proposez de manière détaillée une extension possible du réseau bayésien précédent à partir des différentes "extensions" de ces modèles vues en cours.