Représentation de Connaissances Incertaines

Exercices

Exercice 1 *On considère le problème classique suivant. On utilisera les notations suivantes :* `FN` *pour "il fait froid en Norvege",* `ZO` *pour "la cage des cygnes est ouverte au Zoo",* `CP` *pour "il y a un cygne sauvage sur ma pelouse", et enfin* `SA` *pour "le prix du saumon a augmenté". Les dépendances directes entre ces variables nous permettent de donner la structure suivante pour le réseau Bayesien.*

On dispose en outre des probabilités conditionnelles significatives (celles qui doivent être associées au noeuds du réseau) suivantes :

$P(fn)$	0.95
$P(zo)$	0.7
$P(sa \mid fn)$	0.95
$P(sa \mid \neg fn)$	0.1
$P(cp \mid fn \wedge zo)$	0.9
$P(cp \mid fn \wedge \neg zo)$	0.05
$P(cp \mid \neg fn \wedge zo)$	0
$P(cp \mid \neg fn \wedge \neg zo)$	0

Vous répondrez aux requêtes suivantes :

- Quelle est la probabilité pour qu'il y ait un cygne sur ma pelouse sachant que le Zoo a laissé une cage ouverte ?

  Corrigé : Commençons par l'approche "intuitive". On cherche à calculer $P(cp|zo)$, donc

  $$\frac{P(cp \wedge zo)}{P(zo)}$$

  Or (par marginalisation), on obtient $P(cp \wedge zo) =$

  $$P(cp \wedge zo \wedge fn \wedge sa) + P(cp \wedge zo \wedge \neg fn \wedge sa) + P(cp \wedge zo \wedge fn \wedge \neg sa) + P(cp \wedge zo \wedge \neg fn \wedge \neg sa)$$

  Observons que SA n'est ancêtre ni de la variable requête, ni de la variable observation. On aurait donc pu l'éliminer, et il suffit par conséquent de marginaliser sur FN (selon le principe de pertinence). Ce qui donne simplement

  $$P(cp \wedge zo \wedge fn) + P(cp \wedge zo \wedge \neg fn)$$

puis, après décomposition selon la structure du réseau

$$P(cp|fn \wedge zo) \times P(fn|zo) \times P(zo) + P(cp|\neg fn \wedge zo) \times P(\neg fn|zo) \times P(zo)$$

On observe que le dénominateur est $P(zo)$, ce qui permet d'encore simplifier pour arriver à

$$P(cp|fn \wedge zo) \times P(fn|zo) + P(cp|\neg fn \wedge zo) \times P(\neg fn|zo) = 0.855$$

On peut aussi appliquer l'algorithme d'élimination de variables, que nous détaillons ici (pour simplifier l'écriture je pose α pour désigner la valeur du dénominateur).

$$P(CP|zo) = \alpha \sum_{SA} \sum_{FN} P(CP|FN \wedge zo) \times P(SA|FN) \times P(FN) \times P(zo)$$

$$P(CP|zo) = \alpha \sum_{SA} \sum_{FN} P(CP|FN \wedge zo) \times P(SA|FN) \times P(FN) \times P(zo)$$

$$P(CP|zo) = \alpha P(zo) \times \sum_{FN} P(CP|FN \wedge zo) \times P(FN) \times \sum_{SA} P(SA|FN)$$

On observe bien que la somme sur SA va donner 1 (ce qui explique le test de pertinence). Nous allons à présent calculer les facteurs correspondants.

$$P(CP|zo) = \underbrace{P(zo)}_{f_1} \times \sum_{FN} \underbrace{P(CP|FN \wedge zo)}_{f_2} \times \underbrace{P(FN)}_{f_3}$$

$f_3(FN) =$

FN	val
t	0.95
f	0.05

et $f_2(CP, FN) =$

CP	FN	val
t	t	0.9
t	f	0
f	t	0.1
f	f	1

Ce qui donne

$(f_2 \times f_3)(CP, FN) =$

CP	FN	val
t	t	0.855
t	f	0
f	t	0.095
f	f	0.05

et $\sum_{FN}(f_2 \times f_3)(CP, FN) =$

CP	val
t	0.855
f	0.145

Il n'est pas nécessaire d'aller plus loin puisque le facteur f_1 s'élimine avec le dénominateur.

– Quelle est la probabilité pour la cage ne soit pas ouverte au Zoo étant donné qu'il ne fait pas froid en Norvège ?

  Corrigé : Les évènements sont indépendants... Il s'agit donc simplement de la probabilité pour que la cage ne soit pas ouverte ($P(\neg zo|\neg fn) = P(\neg zo)$). Quant à la proba des évènements conjoints, ce serait évidemment le produit.

– Quelle est la probabilité pour la cage soit ouverte au Zoo étant donné qu'il n'y a pas de cycgne sur ma pelouse ?

  Corrigé : Il s'agit ici d'un diagnostic. On cherche à calculer $P(ZO|\neg cp)$. Utilisons l'algorithme de l'élimination de variables. Après élimination de SA et décomposition, on obtient

$$P(ZO|\neg cp) = \alpha \underbrace{P(ZO)}_{f_1} \times \sum_{FN} \underbrace{P(\neg cp|FN \wedge ZO)}_{f_2} \times \underbrace{P(FN)}_{f_3}$$

$f_3(FN) =$

FN	val
t	0.95
f	0.05

et $f_2(FN, ZO) =$

FN	ZO	val
t	t	0.1
t	f	0.95
f	t	1
f	f	1

Ce qui donne

$(f_2 \times f_3)(FN, ZO) =$

FN	ZO	val
t	t	0.095
t	f	0.9025
f	t	0.05
f	f	0.05

et $\sum_{FN}(f_2 \times f_3)(FN, ZO) =$

ZO	val
t	0.145
f	0.9525

Reste à multiplier par

$f_1(ZO) =$

ZO	val
t	0.7
f	0.3

On obtient alors le facteur résultat

$f_r(ZO) =$

ZO	val
t	0.1015
f	0.28575

Notez que la somme de ces probas ne fait pas 1 ! Pour normaliser, il suffit alors de sommer ce facteur résultat sur la variable requête (ZO, ce qui donne 0.38275) pour considérer toutes les valeurs que peut prendre cette variable, et de diviser par cette valeur. On aura alors :

$$P(zo|\neg cp) = 0.26 \text{ et } P(\neg zo|\neg cp) = 0.74$$

– Quelle est la probabilité pour la cage ne soit pas ouverte au Zoo étant donné qu'il ne fait pas froid en Norvège et que le prix du saumon n'a pas augmenté ?

  CORRIGÉ : Ici les variables seraient à nouveau indépendantes, donc $P(zo|\neg fn \wedge \neg sa) = 0.7$, pas très intéressant. Pour ceux qui voudraient s'entraîner un peu plus, quelques autres résultats, par exemple : $P(zo|\neg cp \wedge \neg sa) = 0.56$, ou encore $P(fn|\neg sa \wedge \neg cp) = 0.27$.