Intelligence Artificielle et Traitement de l'Incertain
Contrôle Terminal du Mercredi 13 Octobre 2021.

Le barème est donné à titre indicatif. La compréhension du sujet faisant partie de l'épreuve, **on ne répondra à aucune question**. Si vous rencontrez des ambiguïtés, vous expliquerez **sur votre copie** comment vous les interprétez.
Durée 2h. **Documents autorisés.**

⚠ Les trois exercices sont indépendants.

I - Réseaux bayésiens (13 points)

Énoncé : Trois lignes de bus (27,44,115) desservent l'Université Paul Sabatier. La fréquence sur la ligne 27 est plus élevée que celle de la ligne 44, elle-même plus élevée que celle de la line 115 : il y a en moyenne 5 bus sur la ligne 27 pour 3 sur la ligne 44 et 2 sur la ligne 115.

Sur la ligne 27, en journée, 9 bus sur 10 sont simples et la nuit tous les bus sont simples. La ligne 44 n'a que des bus doubles en journée et la nuit seulement 1 bus sur 10 est simple. Sur la ligne 115, 8 bus sur 10 sont des bus doubles, de nuit et de jour.

70% des bus doubles ont des écrans d'information alors que seuls 20% des bus simples en ont.

Les bus de la ligne 27 qui circulent dans l'université affichent la direction : Rangueil sauf pour 1 bus sur 10 de cette même ligne qui affiche la direction : Toulouse-centre.
De même les bus des lignes 44 et 115 qui circulent à l'Université affichent la direction : Toulouse-centre sauf pour 1 bus sur 10 de chacune de ces lignes qui affiche la direction : Rangueil.

La plupart (75%) des bus en direction de Toulouse-centre diffusent des messages sonores. En direction de Rangueil aucun bus ne diffusent de message sonore.

Le régime journée (resp. nuit) commence à 6h (resp. 18h) et se termine à 18h (resp. 24h)

On s'intéresse à un bus en train de circuler sur l'université.

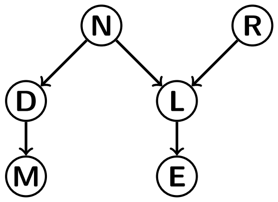
Variables : On propose d'utiliser les 6 variables suivantes :

VARIABLE	valeurs	signification
D	r, t	direction (Rangueil ou Toulouse-centre)
E	e, ne	écran d'information (écran, non écran)
L	s, d	longueur du bus (simple ou double)
M	m, nm	message sonore (message, non message)
N	$l27$, $l44$, $l115$	numéro de la ligne du bus
R	j, n	régime (journée ou nuit)

1) Construire un réseau causal probabiliste (noté RB dans la suite) en utilisant les variables ci-dessus et les probabilités définies dans l'énoncé. *Vous dessinerez le réseau et donnerez le contenu initial de toutes les tables.*

Corrigé

2 pts



Pour compléter le réseau, il faut donner pour chaque variable, la table de probabilité conditionnelle de la variable sachant ses parents.

	l27	l44	l115
P(N)	0.5	0.3	0.2

	j	n
P(R)	2/3	1/3

P(D|N)	l27	l44	l115
r	0.9	0.1	0.1
t	0.1	0.9	0.9

P(L|N,R)	l27		l44		l115	
	j	n	j	n	j	n
s	0.9	1	0	0.1	0.2	0.2
d	0.1	0	1	0.9	0.8	0.8

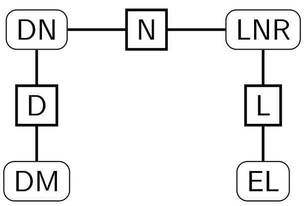
P(M|D)	t	r
m	0.75	0
nm	0.25	1

P(E|L)	s	d
e	0.2	0.7
ne	0.8	0.3

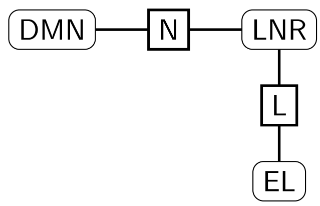
2) Proposez un arbre de groupes de jonction avec des groupes de taille maximum 3. Expliquez comment il est initialisé : vous donnerez les formules permettant de calculer le contenu des tables à l'initialisation sans effectuer les calculs. (Cet arbre sera utilisé dans les questions suivantes.)

Corrigé

2 pts



ou



c'est bien un arbre de groupe car toute variable apparait au moins dans un groupe avec ses parents. Il est de jonction car ici toutes les intersection de deux groupes non adjacents sont vides.

V1 : On initialise $t_{DM} = P(M|D)$, $t_{DN} = P(D|N)$, $t_{LNR} = P(N) * P(R) * P(L|N,R)$, $t_{EL} = P(E|L)$ et $t_D = t_N = t_L = \mathbb{1}$.

V2 : on initialise t_{LNR} et t_{EL} comme V1, $t_{DMN} = P(M|D) * P(D|N)$, $t_L = t_N = \mathbb{1}$.

3) Donnez les formules permettant d'obtenir des tables qui contiennent les probabilités conjointes pour tous les groupes de l'arbre de groupes. *Détaillez les raisonnements pour les obtenir, en justifiant les indépendances entre variables, lorsque c'est nécessaire, n'effectuez aucun calcul.*

Corrigé

3 pts On a donc $t_{LNR} = P(L, N, R)$ car $P(L, N, R) = P(N) * P(R|N) * P(L|N, R)$ et N et R sont indépendantes lorsqu'on ne connaît rien sur L ni ses descendants car la liaison est convergente en L donc $P(R|N) = P(R)$.
V1 : Pour que t_{DN} contienne la proba conjointe $P(D, N)$ on la multiplie par $P(N)$ c'est un envoi de message de t_{LNR} vers t_{DN}, au passage $t_N = P(N)$. Donc $t_{DN} = P(N)*P(D|N) = P(D, N)$.
On marginalise sur D la table t_{DN} pour obtenir $t_D = P(D) = \sum_N P(D, N)$ puis on obtient une proba conjointe dans $t_{DM} = P(D) * P(M|D) = P(D, M)$.
Pour t_{EL} on fait passer un message de t_{LNR} à t_{EL}, en passant par le séparateur pour lequel on marginalise sur NR, $t^*_L = P(L) = \sum_{NR} P(L, N, R)$, on obtient ensuite $t^*_{EL} = (t_{EL} * t^*_L)/\mathbb{1} = P(L) * P(E|L) = P(E, L)$
V2 : pour la table t_{DMN} on obtient la proba conjointe par envoi de message de t_{LNR} vers t_{DMN} en passant par la marginalisation de t_{LNR} sur LR pour obtenir $P(N)$. Donc $t_{DMN} = P(D, M, N) = P(N)*P(D|N)*P(M|DN)$ en utilisant le fait que M est indépendant de N sachant D car c'est une liaison série en D donc D d-sépare M et N, ainsi : $P(M|D, N) = P(M|D)$.

4) Un bus de la ligne 44 circule dans l'université. Est-il simple ou double ? *Détaillez votre raisonnement et donnez le résultat final avec 2 décimales.*

Corrigé

4 pts On cherche $P(L = s|N = l44)$, c'est-à-dire $P(L|N) = P(L, N)/P(N)$, on connait déjà $P(N)$, pour calculer $P(L, N)$ on marginalise $P(L, N, R)$ sur R.

t_{LNR} = P(L,N,R)	l27		l44		l155	
	j	n	j	n	j	n
s	2/3*0.5*0.9 = 0.9/3	1/3*0.5*1 =0.5/3	0 0	1/3*0.3*0.1 =0.03/3	2/3*0.2*0.2 =0.08/3	1/3*0.2*0.2 =0.04/3
d	2/3*.0.5*0.1 =0.1/3	0 =0	2/3*0.3*1 =0.6/3	1/3*0.3*0.9 =0.27/3	2/3*0.2*0.8 =0.32/3	1/3*0.2*0.8 =0.16/3
P(N=l44,L=s)	...	...	0.03/3		...	...
P(N=l44,L=d)	...	...	0.6/3+0.27/3 =0.87/3		...	...
P(N=l44)	...	...	0.9/3		...	...

Donc $P(L = s|N = l44) = 0.03/0.9 \simeq 0.033$ et $P(L = d|N = l44) = 0.87/0.9 \simeq 0.967$, il y a une probabilité de 96.7% que ce soit un bus double.

5) Il est 20h30. Un bus simple circule dans l'université. Quelle est la probabilité que sa direction soit Rangueil ? *Pour cette question vous détaillerez votre méthode (les évidences à intégrer, les envois de messages à faire) et vous donnerez le résultat final avec 2 décimales, vous pouvez éviter les calculs des cases inutiles des tables.*

Corrigé

2 pts On cherche $P(D = r|L = s, R = n)$. Il faut donc considérer l'événement $e_1 = (L = s, R = n)$ calculer son impact sur t_{LNR} et transférer cette information sur la table t_{DN} (pour la version V1, sur la table t_{DMN} pour la version V2).

Corrigé

On calcule d'abord t^*_{LNR} après l'arrivée de $e_1 = (L = s, R = n)$. Puis on marginalise sur N pour envoyer le message au groupe TN (resp. DMN pour la V2).

$t_{LNR} =$	l27		l44		l155		$P(e_1)$
$P(L, N, R, e_1)$	j	n	j	n	j	n	
s	0	0.5/3	0	0.03/3	0	0.04/3	
d	0	0	0	0	0	0	
$P(N, e_1)$	0.5/3		0.03/3		0.04/3		0.57/3

V1 : On envoie ensuite le message au groupe DN : $t^*_{DN} = t_{DN} * t^*_N / t_N$ ainsi on obtient $t^*_{DN} = P(D, N, e_1) = P(N)*P(D|N)*P(N, e_1)/P(N) = P(D|N)*P(N, e_1)$ qu'on marginalise pour connaître la direction : $P(D, e_1)$

$t^*_{DN} = P(D, N, e_1)$	l27	l44	l115	$P(D, e_1)$
r	0.9*0.5/3	0.1*0.03/3	0.1*0.04/3	(0.45+0.003+0.004)/3
t	0.1*0.5/3	0.9*0.03/3	0.9*0.04/3	...

Donc la proba de direction Rangueil sachant qu'il est 20h30 et que le bus est simple est $P(D = r|e_1) = P(D, e_1)/P(e_1) = 0.457/0.57 \simeq 0.8017$.
Il y a donc 80.17% de chances qu'il soit en direction de Rangueil.

II - Raisonnement en logique possibiliste (7 points)

A) Raisonnement avec les axiomes possibilistes

On dispose des connaissances possibilistes suivantes :

K1. $N(p \to m) \geq 0.6$

K2. $\Pi(o \to p) \geq 0.8$

K3. $N(o \to d) \geq 0.4$

K4. $N(d \to \neg m) \geq 0.9$

K5. $N(o) = 1$

Signifiant respectivement qu'il est 0.6-certain que manger un oeuf pourri rend malade, il est 0.8-possible que lorsqu'on a un oeuf alors il est pourri, il est 0.4-certain que lorsqu'on a un oeuf alors il est dur, il est 0.9-certain que manger un oeuf dur ne rend pas malade. On dispose d'un oeuf.

1) *En justifiant chaque étape du raisonnement par les noms des axiomes sur les mesures de nécessité ou possibilité utilisés*, donnez une information sur le degré de certitude associé à la vérité de chaque proposition suivante ou de sa négation.

   (a) p (oeuf pourri)

   (b) d (oeuf dur)

Corrigé

$\boxed{\textbf{1 pt}}$

K6. $\Pi(o \to p) = \Pi(\neg o \vee p) = \max(\Pi(\neg o), \Pi(p)) \geq 0.8$	K2 et Π mesure de possibilité.
K7. $\Pi(\neg o) = 0$	dualité N/P et K5
K8. $\Pi(p) \geq 0.8$	K7 et K6
K9. $N(\neg p) \leq 0.2$	dualité N/P

Corrigé

$\boxed{\textbf{0.5 pt}}$

K10. $(o \to d) \wedge o \vdash d$ donc $N((o \to d) \wedge o) \leq N(d)$	monotonie de N
K11. $\min(N(o \to d), N(o)) \leq N(d)$	N nécessité
K12. $N(d) \geq 0.4$	K3,K5 et K12

2) Donnez deux informations sur le degré de nécessité de non malade en utilisant deux raisonnements différents. Que peut-on en conclure ?

Corrigé

$\boxed{\textbf{1.5 pts}}$

K13. $(d \to \neg m) \wedge d \vdash \neg m$ donc $N((d \to \neg m) \wedge d) \leq N(\neg m)$	monotonie de N
K14. $\min(N(d \to \neg m), N(d)) \leq N(\neg m)$	N nécessité
K15. $N(\neg m) \geq 0.4$	K12,K4 et K14

Deuxième raisonnement :

K16. $(p \to m) \wedge \neg m \vdash \neg p$ donc $N((p \to m) \wedge \neg m) \leq N(\neg p)$	monotonie de N
K17. $\min(N(p \to m), N(\neg m)) \leq N(\neg p)$	N nécessité
K18. $N(\neg m) \leq 0.2$	K9,K2 et K17

Conclusion : les connaissances possibilistes sont incohérentes, on ne peut rien conclure sur la nécessité de non malade.

B) Résolution possibiliste

Les connaissances disponibles concernent la réalisation d'un projet. On les accompagne de leur degré de certitude.

F1. On est certain à un degré au moins 0.6 que si l'ingénieur 2 est malade alors le projet ne sera pas réalisé.

F2. On est un peu certain (à un degré au moins 0.5) que si l'ingénieur 1 n'est pas malade et que le matériel est disponible alors le projet sera réalisé.

F3. On est certain à un degré au moins 0.6 que si le matériel n'est pas disponible alors le projet ne sera pas réalisé.

F4. On est faiblement certain (à un degré 0.4) que si le précédent projet est en retard alors le matériel ne sera pas disponible.

F5. On est quasiment sûr (à un degré au moins 0.8) que l'ingénieur 1 n'est pas malade

F6. On a quelques certitude (à un degré au moins 0.2) que l'ingénieur 2 est malade.

F7. On pense (avec une certitude d'au moins 0.3) que le projet précédent est en retard.

On utilisera les variables suivantes :

VARIABLE	signification
MD	Matériel disponible
MI1	Malade Ingénieur 1
MI2	Malade Ingénieur 2
PR	Projet réalisé
RPP	Retard projet précédent

1) En utilisant les variables ci-dessus, écrire les connaissances sous la forme d'une base de clauses possibilistes BP (numérotez les clauses), on rappelle qu'une clause possibiliste se note (C, α) et signifie qu'on est certain de C à un degré supérieur ou égal à α c'est-à-dire $N(C) \geq \alpha$ où N est une mesure de nécessité.

Corrigé

1 pt

$$
\begin{array}{ll}
C1 : \neg MI2 \vee \neg PR & 0.6 \\
C2 : MI1 \vee \neg MD \vee PR & 0.5 \\
C3 : MD \vee \neg PR & 0.6 \\
C4 : \neg RPP \vee \neg MD & 0.4 \\
C5 : \neg MI1 & 0.8 \\
C6 : MI2 & 0.2 \\
C7 : RPP & 0.3
\end{array}
$$

2) BP est-elle consistante ? sinon quel est son degré d'inconsistance.

Corrigé

1 pt Elle est consistante : un modèle est RPP, $MI2$, $\neg MI1$, $\neg PR$, $\neg MD$.

3) Que peut-on conclure sur $N(PR)$ et $N(\neg PR)$? *Vous détaillerez les résolutions effectuées.*

Corrigé

2 pts

$$C8 = C1 + C6 : \neg PR \quad 0.2$$

Donc on peut conclure $N(\neg PR) \geq 0.2$

$$
\begin{array}{ll}
C9 = C4 + C7 : \neg MD & 0.3 \\
C10 = C3 + C9 : \neg PR & 0.3
\end{array}
$$

Donc on peut conclure $N(\neg PR) \geq 0.3$ ce qui est plus précis.
La base étant consistante ($N(\perp) = 0$), puisqu'on a quelque certitude sur $\neg PR$ alors $N(PR) = 0$ (car $\min(N(A), N(\neg A)) = N(\perp) = 0$)