

Directives pédagogiques:

1. Toute documentation est autorisée.
 2. La durée de cet examen est de 2h30 (y compris le temps pour faire la remise). Vous devez remettre vos réponses dans un fichier électronique sur Studium au plus tard à 15:00. Le fichier électronique peut être dans un format de votre choix : texte, Word, PDF, ou photo (vous pouvez répondre sur des feuilles et prendre des photos).
 3. Le plagiat est strictement interdit et puni.
-

Question 1. (20%) Logique et inférence

1. Quelle est la substitution la plus générale respectivement pour unifier les paires d'expressions suivantes (ou échec si les expressions ne sont pas unifiables)?
 - a. $P(F(y), w, G(z, y)) = P(x, x, G(z, A))$
 - b. $P(A, x) = P(z, F(y))$
 - c. $Q(x, x) = Q(y, F(y))$
 - d. $F(x, G(F(a), u)) = F(G(u, v), x)$

2. Preuve avec la règle de résolution

Nous avons une base de connaissances contenant les phrases suivantes :

- a) Une personne qui aime les animaux est aimée par quelqu'un.
- b) Une personne qui tue un animal n'est aimée par personne.
- c) Un animal – Tuna, a été tué, soit par Jacques, soit par Marie.
- d) Jacques aime Marie.

Conclusion à prouver :

- e) Jacques a tué Tuna.

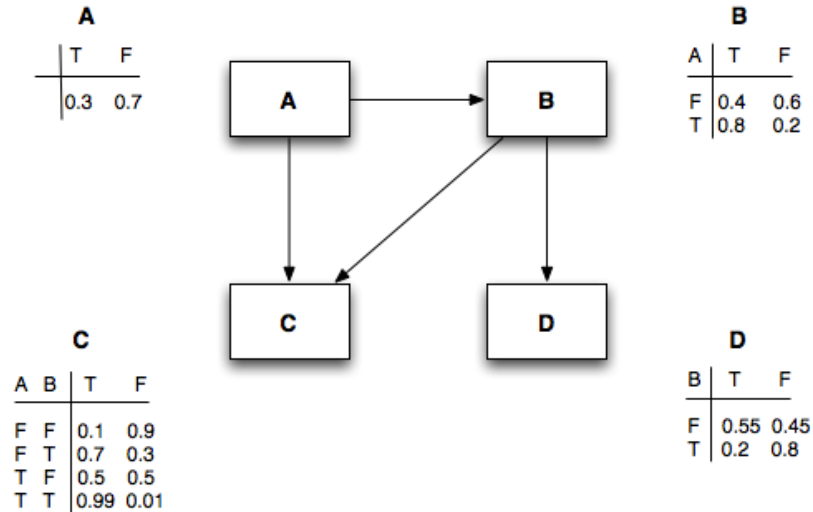
Pour cette question, vous devez d'abord traduire les phrases en logique du premier ordre. Ensuite, vous devez les convertir en forme normale conjonctive. Finalement, vous appliquez la règle de résolution pour prouver que la négation de la conclusion n'est pas compatible avec la base de connaissances.

Question 2. (20%) probabilité

1. Pour une question de test quiz avec 4 réponses au choix, on suppose qu'un étudiant a une probabilité de 0,8 de connaître la vraie réponse, et une probabilité de 0,2 ne pas connaître la réponse. Quand il ne connaît pas la réponse, il choisit au hasard une des 4 réponses. Supposons qu'il a choisi la bonne réponse pour cette question, quelle est la probabilité que l'étudiant connaît réellement la réponse?
2. Supposons que nous avons 3 pièces de monnaie normales. Quelles soit la situation, 2 des pièces vont se montrer pareilles (pile ou face toutes les deux). Pour la troisième pièce. La chance de face ou pile sera 1/2. Ainsi, la probabilité que les 3 pièces se montrent pareilles (face ou pile pour toutes les 3) est 1/2. Est-ce que ce raisonnement est correct? Si non, montrez le bon raisonnement et la bonne probabilité.

Question 3. (20%) Réseau bayésien

Considérez le réseau bayésien suivant où les variables A, B, C D sont binaires. Les tables de probabilités conditionnelles sont montrées à côté des nœuds.



Déterminez les probabilités suivantes :

- 1) $P(A=T|C=T,D=T)$
- 2) $P(B=T|C=T)$

Pour cette question, si vous n'avez pas le temps de faire le calcul final exactement, il vous suffit de dériver les formules jusqu'à mettre les chiffres dans les formules.

Question 4. (20%) Apprentissage

Nous avons observé qu'un lecteur a choisi à lire ou non (passer) un livre qui lui est présenté, selon si le livre est écrit par un auteur connu ou non, si le sujet du livre est nouveau ou non (suite), et si le livre est long. Tous les attributs observés sont du type binaire. Les 10 cas observés sont listés dans la table suivante :

Exemple	Auteur	Sujet	Longueur	Action
e1	connu	nouveau	long	passer
e2	inconnu	nouveau	court	lire
e3	inconnu	suite	long	passer
e4	connu	suite	long	passer
e5	connu	nouveau	court	lire
e6	connu	suite	long	passer
e7	inconnu	suite	court	passer
e8	inconnu	nouveau	court	lire
e9	connu	suite	long	passer
e10	connu	nouveau	long	passer

1. Pour construire un arbre de décision pour modéliser les décisions de ce lecteur, on doit calculer le gain d'information de chaque attribut (Auteur, Sujet, Longueur). Ici, on vous demande de calculer le gain d'information de l'attribut **Auteur** seulement.

2. Si on utilise la classification bayésienne naïve (Naïve Bayes), on doit déterminer une série de probabilités de base afin de pouvoir calculer $P(\text{Action} \mid \text{Auteur, Sujet, Longueur})$. Déterminez ces probabilités que vous devez estimer. Ensuite, utilisez ces probabilités pour calculer la probabilité que le lecteur choisit à lire le livre pour le cas **e11** suivant:

e11	inconnu	nouveau	long	?
-----	---------	---------	------	---

Si vous n'avez pas eu le temps de déterminer la valeur exacte de probabilités, montrez la formule avec suffisamment de détails.

3. On veut utiliser un réseau de neurones pour prédire l'action du lecteur. Ce réseau utilise une couche cachée de 3 neurones. Dessinez l'architecture correspondant à ce problème. Notamment, vous devez prendre soin de dessiner le nombre de neurones à la couche d'entrée, la couche cachée et la couche de sortie.

Expliquez (en mots) quelles sont les valeurs que vous devez utiliser pour les entrées, quelle sera la valeur que vous obtenez à la sortie, et comment vous utiliserez la sortie pour prendre une décision.

Question 5. (20%) Langue naturelle

1. Définissez une grammaire hors contexte simple (mais assez générale) pour accepter les phrases marquées par '+' et rejeter les phrases marquées par '-' :

marie mange une pomme rouge. (+)
 marie mange une pomme. (+)
 une pomme pousse. (+)
 marie pousse la télévision. (+)
 marie mange pomme. (-)
 marie la télévision. (-)
 marie pousse mange. (-)
 mange une pomme. (-)

Votre grammaire doit contenir les symboles de catégories suivants (en plus des mots) : P (phrase), GN (groupe nominal), GV (groupe verbal), ADJ (adjectif), NOM (nom), PRON (pronom), ART (article) et VERBE (verbe). On ne tient pas compte des accords en nombre et en genre dans cette grammaire, ni si la phrase fait du sens.

2. Supposons qu'on veut apprendre une grammaire hors contexte probabiliste, c'est-à-dire déterminer la probabilité pour chacune des règles que vous avez établies pour la question précédente. On suppose que nous avons un mini corpus pour déterminer ces probabilités, contenant les 4 phrases en haut marquées par +. Pouvez-vous déterminer ces probabilités?
3. Avec votre grammaire hors contexte probabiliste, quelle sera la probabilité de la phrase suivante?

marie mange une pomme.

(Note : si vous n'avez pas pu établir les probabilités pour les règles, mettez les probabilités au hasard, et calculez la probabilité de la phrase avec ces probabilités.)