Excercices INFO-F-311, Intelligence Artificielle

Axel Abels, Tom Lenaerts, Yannick Molinghen et Pascal Tribel

Années académiques 2022–2024

- Ce fichier contient les solutions pour les exercices du cours INFO-F-311 Intelligence Artificielle organisé en 3$^{\text{ième}}$ année du Bachelier en Informatique à l'Université Libre de Bruxelles (Faculté des Sciences).
- Les solutions pour les questions des examens de l'année académique précédente sont insérées entre les autres exercices.

Contents

1 Introduction IA **3**

2 Algorithmes de recherche simples **9**
2.1 Examen 2022-2023 11

3 Algorithmes de recherche informés **13**
3.1 Examen 2022-2023 15

4 Recherche locale **16**

5 Jeux et recherche d'adversaire **19**
5.1 Examen 2022-2023 21
5.2 Examen 2022-2023 22

6 Les bases de la théorie des probabilités **24**

7 Réseaux Bayésiens **26**
7.1 Examen 2022-2023 34

8 Inférence dans les réseaux bayésiens et *d-separation* **38**
8.1 Examen 2022-2023 40

9 Échantillonnage et réseaux bayésiens **44**
9.1 Examen 2022-2023 47

10 Modèles de Markov **49**

11 Réseaux de décisions **52**
11.1 Examen 2022-2023 55

12 Processus décisionnels de Markov **58**
12.1 Examen 2022-2023 60

13 Apprentissage par renforcement **65**
13.1 Examen 2022-2023 68

14 Apprentissage automatique **70**
14.1 Examen 2022-2023 72

1 Introduction IA

Ex. 1

Définissez ...

1. L'intelligence (artificielle); Les définitions du dictionnaire de l'intelligence parlent de "la capacité d'acquérir et d'appliquer des connaissances" ou "la faculté de penser et de raisonner" ou "la capacité de comprendre et de tirer profit de l'expérience". Ce sont toutes des réponses raisonnables, mais si nous voulons quelque chose de quantifiable, nous utiliserions quelque chose comme "la capacité d'agir avec succès sur un large éventail d'objectifs dans des environnements complexes". Nous définissons l'intelligence artificielle comme l'étude et la construction de programmes agents performants dans une classe d'environnements donnée, pour une architecture agent donnée ; ils font ce qu'il faut. Une partie importante de cela consiste à gérer l'incertitude de l'état actuel, de ce que pourraient être les résultats d'actions possibles et de ce que nous désirons vraiment.

2. L'agent : Nous définissons un agent comme une entité qui agit en réponse aux perceptions d'un environnement.

3. La rationalité: Nous définissons la rationalité comme la propriété d'un système qui fait "ce qu'il faut" compte tenu de ce qu'il sait. Voir la Section 2.2 [1] pour une discussion plus complète. Le concept de base est la rationalité parfaite ; Section 2.2.3 décrit l'impossibilité d'atteindre la rationalité parfaite et propose une définition alternative.

4. Le raisonnement : Nous définissons le raisonnement comme un processus consistant à dériver de nouvelles phrases à partir d'anciennes, de sorte que les nouvelles phrases sont nécessairement vraies si les anciennes sont vraies. (Remarque qui ne fait référence à aucune syntaxe ou langage formel spécifique, mais qui nécessite une notion bien définie de la vérité.)

Ex. 2

Dans quelle mesure les systèmes informatiques suivants sont-ils des instances d'intelligence artificielle :

1. Scanners de codes-barres de supermarché. Bien que la lecture de codes à barres soit en quelque sorte une application de vision par ordinateur, ce ne sont pas des systèmes d'IA. Le problème de la lecture d'un code à barres est une forme d'interprétation visuelle extrêmement limitée et artificielle, et il a été soigneusement conçu pour être aussi simple que possible, compte tenu du matériel.

2. Moteurs de recherche Web. Le problème de la détermination de la pertinence d'une page Web pour une requête est un problème de compréhension du langage naturel, et les techniques sont liées à celles aux derniers chapitres du livre [1]. Les moteurs de recherche utilisent également des techniques de regroupement. De même, d'autres fonctionnalités fournies par un moteur de recherche utilisent des techniques intelligentes ; par exemple, le correcteur orthographique utilise une forme d'exploration de données basée

sur l'observation des corrections des utilisateurs de leurs propres fautes d'orthographe. D'un autre côté, le problème de l'indexation de milliards de pages Web d'une manière qui permet une récupération en quelques secondes est un problème de conception de bases de données, pas d'intelligence artificielle.

3. Fonctions de correction orthographique et grammaticale dans Word. Pas autant. La fonctionnalité de correction orthographique ici est effectuée par comparaison de chaînes à un dictionnaire fixe. La correction grammaticale est plus sophistiquée car elle nécessite l'utilisation d'un ensemble de règles assez complexes reflétant la structure du langage naturel, mais il s'agit toujours d'une tâche très limitée et fixe. Les correcteurs orthographiques des moteurs de recherche seraient considérés comme des instances beaucoup plus proches de l'IA que le correcteur orthographique Word, d'abord parce que la tâche est beaucoup plus dynamique - les correcteurs orthographiques des moteurs de recherche traitent très efficacement les noms propres, qui sont détectés dynamiquement à partir des requêtes des utilisateurs. - et, deuxièmement, en raison de la technique utilisée - l'exploration de données à partir de requêtes d'utilisateurs par rapport à la correspondance de chaînes.

Ex. 3

"Certes, les ordinateurs ne peuvent pas être intelligents, ils ne peuvent faire que ce que leur disent leurs programmeurs." Cette dernière affirmation est-elle vraie et implique-t-elle la première ?

Cela dépend de votre définition de "intelligent" et de "dire". Dans un sens, les ordinateurs ne font que ce que les programmeurs leur ordonnent de faire, mais dans un autre sens, ce que les programmeurs disent consciemment à l'ordinateur de faire a souvent très peu à voir avec ce que l'ordinateur fait réellement. Quiconque a écrit un programme avec un bogue gênant le sait, tout comme quiconque a écrit un programme d'apprentissage automatique réussi. Donc, dans un sens, Samuel "a dit" à l'ordinateur "d'apprendre à jouer aux dames mieux que moi, puis de jouer de cette façon", mais dans un autre sens, il a dit à l'ordinateur "de suivre cet algorithme d'apprentissage" et il a appris à jouer. Nous nous retrouvons donc dans la situation où vous pouvez ou non considérer qu'apprendre à jouer aux dames est un signe d'intelligence (ou vous pouvez penser qu'apprendre à jouer de la bonne manière nécessite de l'intelligence, mais pas de cette manière), et vous pouvez penser que l'intelligence réside dans le programmeur ou dans l'ordinateur.

Ex. 4

Pour chacun des agents suivants, précisez les capteurs (sensors), les actionneurs (actuators) et l'environnement :

1. four à micro-ondes, Ici, il est important de considérer le four comme l'agent, et non l'humain qui l'utilise ; cela signifie que les boutons sont des capteurs, pas des actionneurs ! Il peut sembler qu'il n'y a pas de décisions à prendre, mais la logique du contrôleur interne peut être assez complexe, en particulier s'il y a des capteurs supplémentaires. Pour des

raisons de sécurité, aucune alimentation ne doit être fournie si la porte est ouverte !

- Capteurs : État du bouton, capteur d'ouverture de porte. Certains micro-ondes ont des capteurs d'humidité, des capteurs de température ou des capteurs acoustiques (pour le pop-corn).
- Actionneurs : Allumer/éteindre/régler la puissance du micro-ondes ; alarme sonore d'achèvement ; indiquer l'état illégal du bouton.
- Environnement : Le contenu du four et l'état de la porte (ouverte/fermée).

2. programme d'échecs,

   - Capteurs : un programme typique accepte les entrées au clavier, à la souris ou à l'écran tactile indiquant le mouvement, l'abandon, la demande d'indice, etc. de l'adversaire.
   - Actionneurs : Affichez un coup légal, ou le plateau résultant du coup.
   - Environnement : Au minimum, l'environnement comprend une représentation de l'état du conseil. Il y a quelques problèmes subtils liés à la question de savoir si cela est identique à la représentation interne du programme ou au tableau affiché accessible à l'adversaire. Un programme plus sophistiqué pourrait inclure l'adversaire humain dans l'environnement et essayer de construire un modèle du style de jeu et de la capacité de cet humain afin d'améliorer ses chances de gagner ou de rendre le jeu plus efficace.

3. avion de livraison de ravitaillement autonome. Les réponses ici sont assez similaires à celles pour le taxi autonome. Des détails peuvent être trouvés dans de nombreuses sources en ligne décrivant les systèmes de livraison autonomes actuels.

   - Capteurs : GPS, vitesse de l'air (par exemple, tube de Pitot), altimètre (en fait un capteur de pression atmosphérique, l'altitude signalée dépend donc des conditions météorologiques), capteurs signalant l'état de chaque actionneur, accéléromètre 3 axes, éventuellement une caméra.
   - Actionneurs : alimentation de l'hélice, des ailerons, des gouvernes de profondeur, du gouvernail ; déposer la charge utile.
   - Environnement : zone géographique de vol, conditions atmosphériques.

Ex. 5

Examinons la rationalité des différentes fonctions de l'agent aspirateur (Figure 1).

1. Montrer que la fonction d'agent aspirateur simple est bien rationnelle (Maximise-t-elle son utilité espérée) ? Il suffit de montrer que pour tous les environnements réels possibles (c'est-à-dire toutes les distributions de saleté et les emplacements initiaux), cet agent nettoie les carrés au moins aussi vite que n'importe quel autre agent. C'est trivialement vrai lorsqu'il

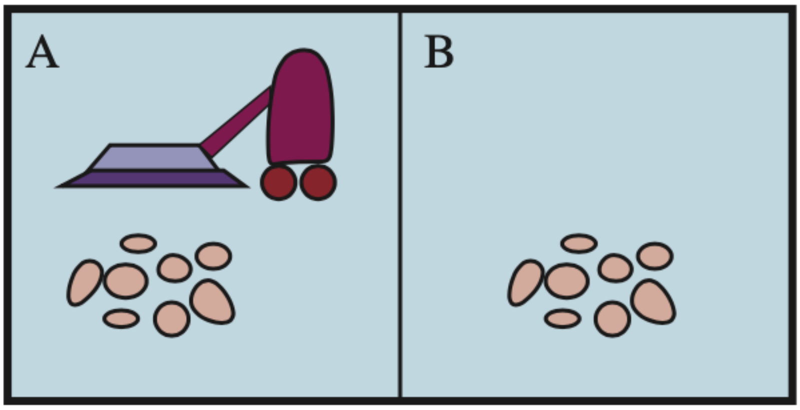


Figure 1: Un monde d'aspirateurs avec seulement deux emplacements. Chaque emplacement peut être propre ou sale, et l'agent peut se déplacer à gauche ou à droite et peut nettoyer la case qu'il occupe. Différentes versions du monde d'aspirateurs imposent différentes règles sur ce que l'agent peut percevoir, si ses actions réussissent toujours, etc. Image de [1].

n'y a pas de saleté. Lorsqu'il y a de la saleté à l'emplacement initial et aucune à l'autre emplacement, le monde est propre après une étape ; aucun agent ne peut faire mieux. Lorsqu'il n'y a pas de saleté à l'emplacement initial mais de la saleté à l'autre, le monde est propre après deux étapes ; aucun agent ne peut faire mieux. Lorsqu'il y a de la saleté aux deux endroits, le monde est propre après trois étapes ; aucun agent ne peut faire mieux. (Remarque : en général, la condition énoncée dans la première phrase de cette réponse est beaucoup plus stricte que nécessaire pour qu'un agent soit rationnel.)

2. Décrivez une fonction d'agent rationnel pour le cas où chaque mouvement coûte un point. Le programme d'agent correspondant nécessite-t-il un état interne ? L'agent en (a) continue d'avancer et de reculer même après que le monde soit propre. Il vaut mieux faire NoOp une fois que le monde est propre (le chapitre le dit). Maintenant, puisque le percept de l'agent ne dit pas si l'autre carré est propre, il semblerait que l'agent doive avoir une certaine mémoire pour dire si l'autre carré a déjà été nettoyé. Rendre cet argument rigoureux est plus difficile - par exemple, l'agent pourrait-il s'arranger pour qu'il ne se trouve que dans un carré gauche propre alors que le carré droit était déjà propre ? En tant que stratégie générale, un agent peut utiliser l'environnement lui-même comme une forme de mémoire externe - une technique courante pour les humains qui utilisent des choses comme les calendriers de rendez-vous et les nœuds dans les mouchoirs. Dans ce cas particulier, cependant, ce n'est pas possible. Considérez les actions réflexes pour [A, Nettoyer] et [B, Nettoyer]. Si l'un d'eux est NoOp, alors l'agent échouera dans le cas où c'est le percept initial mais l'autre carré est sale ; par conséquent, aucun des deux ne peut

être NoOp et, par conséquent, le simple agent réflexe est voué à continuer à bouger. En général, le problème avec les agents réflexes est qu'ils doivent faire la même chose dans des situations qui se ressemblent, même si les situations sont en fait assez différentes. Dans le monde du vide, c'est un gros handicap, car chaque carré intérieur (sauf la maison) ressemble soit à un carré avec de la terre, soit à un carré sans terre.

3. Discutez les conceptions d'agents possibles pour les cas dans lesquels des carrés propres peuvent devenir sales et la géographie de l'environnement est inconnue. Est-il logique que l'agent apprenne de son expérience dans ces cas ? Si oui, que doit-il apprendre ? Si non, pourquoi pas ? Si nous considérons des durées de vie asymptotiquement longues, alors il est clair que l'apprentissage d'une carte (sous une forme ou une autre) confère un avantage car cela signifie que l'agent peut éviter de se heurter à des murs. Il peut également apprendre où la saleté est le plus susceptible de s'accumuler et peut concevoir une stratégie d'inspection optimale. Les détails précis de la méthode d'exploration nécessaire pour construire une carte complète figurent au chapitre 4; les méthodes pour dériver une stratégie d'inspection/nettoyage optimale sont au chapitre 22.

Ex. 6

Pour chacune des affirmations suivantes, indiquez si elle est vraie ou fausse et étayez votre réponse par des exemples ou des contre-exemples, le cas échéant.

1. Un agent qui ne détecte que des informations partielles sur l'état ne peut pas être parfaitement rationnel. Faux. La rationalité parfaite fait référence à la capacité de prendre de bonnes décisions compte tenu des informations de capteur reçues. De plus, dans certains environnements, un agent peut délibérément ignorer une partie des informations de ses capteurs et être tout de même parfaitement rationnel. Par exemple, un agent d'échecs physique peut ignorer les taches de poussière sur l'échiquier et à peu près tout ce qui n'est pas sur l'échiquier ou l'horloge.

2. Il existe des environnements de tâches dans lesquels aucun agent réflexe pur ne peut se comporter rationnellement. Vrai. Un agent réflexe pur ignore les percepts précédents, il ne peut donc pas obtenir une estimation d'état optimale dans un environnement partiellement observable. Par exemple, les échecs par correspondance se jouent en envoyant des coups ; si le coup de l'autre joueur est le percept actuel, un agent réflexe ne pourrait pas suivre l'état du plateau et devrait répondre à, disons, "a4" de la même manière quelle que soit la position dans laquelle il a été joué.

3. Il existe un environnement de tâche dans lequel chaque agent est rationnel. Vrai. Par exemple, dans un environnement avec un seul état, tel que toutes les actions ont la même récompense, peu importe quelle action est entreprise. Plus généralement, tout environnement qui est invariant par la récompense par permutation des actions satisfera cette propriété.

4. L'entrée d'un programme d'agent est la même que l'entrée de la fonction d'agent. Faux. La fonction d'agent, théoriquement parlant, prend

en entrée toute la séquence de percepts jusqu'à ce point, alors que le programme d'agent ne prend que le percept courant.

5. Chaque fonction d'agent peut être mise en œuvre par une combinaison programme/machine. Faux. Par exemple, l'environnement peut contenir des machines de Turing et des bandes d'entrée et le travail de l'agent est de résoudre le problème d'arrêt ; il existe une fonction agent qui spécifie les bonnes réponses, mais aucun programme agent ne peut l'implémenter. Un autre exemple serait une fonction d'agent qui nécessite de résoudre des instances de problèmes insolubles de taille arbitraire en temps constant.

Ex. 7

Donner une description PEAS pour:

1. Acheter des livres d'IA d'occasion sur Internet Partiellement observable, déterministe, séquentiel, statique, discret, agent unique. Cela peut être multi-agents et dynamique si nous achetons des livres aux enchères, ou dynamique si nous achetons sur une échelle suffisamment longue pour que les offres de livres changent.

2. Pratiquer le tennis contre un mur Entièrement observable, stochastique, séquentiel, dynamique, continu, agent unique.

3. Enchérir pour un article à une vente aux enchères Entièrement observable, stratégique, séquentiel, statique, discret, multi-agent.

Caractérisez-les en fonction des propriétés sur la diapositive 41.

Ex. 8

Lisez l'article d'Alan Turing (regardez sur UV).

Quelles objections a-t-il formulées pour la recherche en IA et pour son test ? Il a prédit qu'un ordinateur aurait 30 % de chances de réussir un test de Turing de 5 minutes. Quelle chance aurait un ordinateur aujourd'hui ? Dans 25 ans ?

Ex. 9

Trouvez au moins trois ensembles de principes proposés pour la gouvernance de l'IA. Que proposent-ils et pensez-vous qu'ils sont réalistes ?

Ex. 10

L'apprentissage automatique peut produire des classificateurs avec des préjugés raciaux, sexistes et autres. Trouvez un exemple sur le Web et expliquez quel est le problème de biais dans cette application.

2 Algorithmes de recherche simples

Ex. 1

Donnez une formulation complète du problème pour chacun des problèmes suivants (Session 2 slide 8). Choisir une formulation suffisamment précise pour être mise en œuvre.

1. Il y a six boîtes en verre alignées, chacune avec une serrure. Chacune des cinq premières boîtes contient une clé déverrouillant la boîte qui la suit; la dernière boîte contient une banane. Vous avez la clé de la première boîte et vous voulez la banane.

   - État initial : tel que décrit dans la question.
   - Objectif test : vous avez la banane.
   - Fonction successeur : ouvrez n'importe quelle boîte dont vous avez la clé, obtenez le contenu de n'importe quelle boîte ouverte.
   - Fonction de coût : nombre d'actions.

2. Vous commencez par la séquence ABABAECCEC, ou en général toute séquence composée de A, B, C et E. Vous pouvez transformer cette séquence en utilisant les égalités suivantes : AC = E, AB = BC, BB = E et E$x = x$ pour tout x. Par exemple, ABBC peut être transformé en AEC, puis AC, puis E. Votre objectif est de produire la séquence E.

   - Etat initial : ABABAECCEC.
   - Test de but : est l'état E
   - Fonction successeur : applique une égalité en substituant une sous-séquence par une autre.
   - Fonction de coût : nombre de transformations.

Ex. 2

Vous avez une grille de 9×9 carrés, chacun pouvant être coloré en rouge ou en bleu. La grille est initialement colorée en bleu, mais vous pouvez changer la couleur de n'importe quel carré autant de fois que vous le souhaitez. En imaginant la grille divisée en neuf sous-carrés 3×3, vous voulez que chaque sous-carré soit d'une seule couleur, mais que les sous-carrés voisins soient de couleurs différentes.

1. Formulez ce problème de manière directe. Calculer la taille de l'espace d'état.

   - Etat initial : tous les carrés colorés en bleu.
   - Test de but : chaque sous-carré de couleur identique et les sous-carrés adjacents de couleur différente.
   - Fonction successeur : colorier un carré en rouge ou en bleu.
   - Fonction de coût : nombre de fois que les carrés sont colorés.

Il y a 2^{81} états (2 couleurs possibles pour chacun des $9 \times 9 = 81$ carrés).

2. Vous n'avez besoin de colorer un carré qu'une seule fois. Reformulez et calculez la taille de l'espace d'état. La recherche de graphes en largeur serait-elle plus rapide sur ce problème que sur celui de (1) ? Que diriez-vous d'une recherche arborescente itérative approfondie ? Pour éviter de recolorer les carrés, nous pourrions soit enregistrer les carrés qui ont été colorés, et ne pas leur permettre d'être colorés à nouveau, soit colorer tous les carrés dans un ordre fixe. Nous décrivons ce dernier :

   - Etat initial : tous les carrés sont colorés en bleu, le carré en haut à gauche est le prochain à être coloré.
   - Test de but : chaque sous-carré de couleur identique et les sous-carrés adjacents de couleur différente.
   - Fonction successeur : si nous n'avons pas encore colorié tous les carrés, colorer le carré actuel en rouge ou en bleu (cela met automatiquement à jour le prochain carré à colorier).
   - Fonction de coût : nombre de carrés colorés.

   Il y a $82 \cdot 2^{81}$ états, puisque nous devons enregistrer s'il reste des carrés à colorier (81 booléens), et si oui quel carré est le suivant (1 identifiant).

3. Compte tenu de l'objectif, il nous suffit de ne considérer que les colorations où chaque sous-carré est uniformément coloré. Reformuler le problème et calculer la taille de l'espace d'état.

   - Etat initial : tous les sous-carrés colorés en bleu, le sous-carré supérieur gauche est le prochain à être coloré.
   - Test de but : sous-carrés adjacents colorés différemment.
   - Fonction successeur : si nous n'avons pas encore colorié tous les sous-carrés, colorer le sous-carré actuel en rouge ou en bleu.
   - Fonction de coût : nombre de sous-carrés colorés.

   L'espace d'état a $10 \cdot 2^9$ nœuds: 9 booléens pour indiquer les sous-carrés à colorer, 1 identifiant du prochain sous-carré à colorer, et 2^9 différentes colorations des sous-carrés possibles.

4. Combien de solutions ce problème a-t-il ? Ce problème a 2 solutions : dès que vous choisissez la couleur d'un carré, la couleur de tout autre carré est déterminé.

5. Les parties (2) et (3) ont successivement résumé le problème original (1). Pouvez-vous donner une traduction des solutions du problème (3) en solutions du problème (2) et des solutions du problème (2) en solutions du problème (1) ? Etant donné une solution pour le problème (c), déterminez d'abord la coloration finale des carrés, après que toutes les actions aient été prises. Ensuite, coloriez chacun des carrés de la couleur requise dans l'ordre. Étant donné une solution pour le problème (b), déterminez d'abord la coloration finale des carrés. Nous savons que chaque sous-carré sera coloré d'une seule couleur, nous pouvons donc les colorer afin d'atteindre le même état d'objectif.

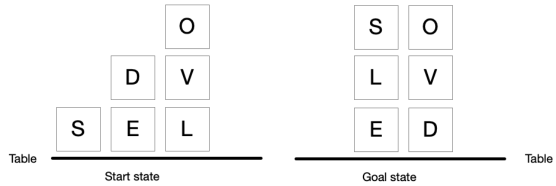


Figure 2: Dans ce problème de blocs, un bras mécanique doit déplacer des blocs pour atteindre l'état cible. Les blocs peuvent être posés sur la table ou les uns sur les autres.

2.1 Examen 2022-2023

Le problème du monde des blocs tel qu'illustré à la Figure 2 demande d'organiser un certain nombre de blocs dans un ordre spécifique. Les blocs peuvent être posés sur la table ou les uns sur les autres. A chaque étape, vous ne pouvez déplacer qu'un seul bloc à la fois. Notez aussi que déplacer un bloc d'un emplacement sur la table à un autre emplacement sur la table n'est pas un mouvement valide. Il y a également suffisamment d'espace sur la table pour placer chaque bloc indépendamment

1. Donnez la description du problème de recherche visualisé dans la Figure 1. (2 points) Un problème de recherche est identifié par 6 éléments : l'espace d'états, un état initial s_0, un ensemble d'actions pour chaque état $A(s)$, un modèle de transition $T(s, a)$, une fonction de test pour l'objectif $G(s)$ et une fonction de coût d'action $c(s, a, s')$. Pour le problème actuel, l'espace d'états est défini par n'importe quelle configuration des six blocs. L'état de départ est donné à la Figure 2 et la fonction de test d'objectif $G(s)$ vérifiera si l'état est le même que l'état d'objectif illustré à la Figure 2. Les actions sont effectuées par le bras mécanique, et elles consistent de la combinaison du ramassage d'un bloc accessible et de son déplacement ailleurs (sur un autre bloc ou sur la table). La fonction de coût est simple : $c(s, a, s') = 1$ ou quelque chose d'équivalent.

2. Quel est le facteur de branchement à la première étape du processus de recherche ? Décrivez les mouvements possibles. (1 point) A l'état initial il y a 8 actions possibles : déplacer le bloc S sur D ou O, déplacer le bloc D sur S, O ou sur la table, déplacer le bloc O sur D, S ou sur la table.

3. Est-ce qu'une recherche en profondeur (*depth-first search*, DFS) sur un graphe donne la garantie de renvoyer une solution en un minimum d'étapes, donc une solution optimale ? Expliquez votre réponse. (1 point) Non, un recherche en profondeur peut être bloquée dans une cycle, ou lorsque ceux-ci sont exclus, fournit simplement le résultat le plus à gauche, indépendamment de la profondeur ou du coût.

Ex. 3

Le problème des *missionnaires et cannibales* est généralement énoncé comme suit. Trois missionnaires et trois cannibales sont d'un côté d'une rivière, avec un bateau qui peut contenir une ou deux personnes. Trouvez un moyen d'amener tout le monde de l'autre côté sans jamais laisser un groupe de missionnaires à un endroit dans lequel ils seraient dépassés en nombre par les cannibales. Ce problème est célèbre en IA parce qu'il a fait l'objet du premier article qui a abordé la formulation du problème d'un point de vue analytique (voyez l'article d'Amarel 1968 sur UV).

1. Formuler précisément le problème, en ne faisant que les distinctions nécessaires pour assurer une solution valable. Dessinez un diagramme de l'espace d'état complet. Voici une représentation possible : un état est un six-tuplet d'entiers indiquant le nombre de missionnaires, de cannibales et de bateaux sur le premier côté, puis sur le deuxième côté de la rivière. Le but est un état avec 3 missionnaires et 3 cannibales du second côté. La fonction de coût est de un par action, et les successeurs d'un état sont tous les états qui déplacent 1 ou 2 personnes et 1 bateau d'un côté à l'autre.

2. Implémenter et résoudre le problème de manière optimale en utilisant un algorithme de recherche approprié. Est-ce une bonne idée de vérifier les états répétés ? L'espace de recherche est petit, donc tout algorithme optimal fonctionne. Il suffit d'éliminer les mouvements qui reviennent à l'état que l'on vient de visiter. De tous les états sauf le premier et le dernier, il n'y a qu'un seul autre choix.

3. Pourquoi pensez-vous que les gens ont du mal à résoudre ce casse-tête, étant donné que l'espace d'état est si simple ? Il n'est pas évident que presque tous les mouvements soient illégaux ou reviennent à l'état précédent. Il y a un sentiment d'un grand facteur de ramification, et aucune manière claire de procéder.

3 Algorithmes de recherche informés

Ex. 1

Définissez dans vos propres mots les termes suivants :

1. état, Un état est une situation dans laquelle un agent peut se trouver. Nous distinguons deux types d'états : les états du monde (les situations concrètes réelles dans le monde réel) et les états représentationnels (les descriptions abstraites du monde réel qui sont utilisées par l'agent dans délibérer sur ce qu'il faut faire).

2. espace d'états, Un espace d'états est un graphe dont les nœuds sont l'ensemble de tous les états, et dont les liens sont des actions qui transforment un état en un autre.

3. arbre de recherche, Un arbre de recherche est un arbre (un graphe sans boucles non dirigées) dans lequel le nœud racine est l'état de départ et l'ensemble d'enfants pour chaque nœud est constitué des états accessibles en effectuant n'importe quelle action.

4. nœud de recherche, Un nœud de recherche est un nœud dans l'arbre de recherche.

5. objectif/but, Un objectif est un état que l'agent essaie d'atteindre.

6. action, Une action est quelque chose que l'agent peut choisir de faire.

7. modèle de transition, Une modèle de transition/fonction successeur décrit les options de l'agent : étant donné un état, elle renvoie un ensemble de paires (action, état), où chaque état est l'état accessible en effectuant l'action.

8. facteur de branchement, Le facteur de branchement dans un arbre de recherche est le nombre d'actions disponibles à l'agent.

Ex.2

Tracez l'opération de recherche A* appliquée au problème de se rendre à Bucarest depuis Lugoj en utilisant l'heuristique de distance en ligne droite (slide 3). Autrement dit, montrez la séquence de nœuds que l'algorithme prendra en compte et les scores f, g et h pour chaque nœud.

L[0+244=244]
M[70+241=311], T[111+329=440]
L[140+244=384], D[145+242=387], T[111+329=440]
D[145+242=387], T[111+329=440], M[210+241=451], T[251+329=580]
C[265+160=425], T[111+329=440], M[210+241=451], M[220+241=461], T[251+329=580]
T[111+329=440], M[210+241=451], M[220+241=461], P[403+100=503], T[251+329=580], R[411+193=604], D[385+242=627]
M[210+241=451], M[220+241=461], L[222+244=466], P[403+100=503], T[251+329=580], A[229+366=595], R[411+193=604], D[385+242=627]
M[220+241=461], L[222+244=466], P[403+100=503], L[280+244=524], D[285+242=527], T[251+329=580], A[229+366=595], R[411+193=604], D[385+242=627]

L[222+244=466], P[403+100=503], L[280+244=524], D[285+242=527], L[290+244=534], D[295+242=537], T[251+329=580], A[229+366=595], R[411+193=604], D[385+242=627]
P[403+100=503], L[280+244=524], D[285+242=527], M[292+241=533], L[290+244=534], D[295+242=537], T[251+329=580], A[229+366=595], R[411+193=604], D[385+242=627], T[333+329=662]
B[504+0=504], L[280+244=524], D[285+242=527], M[292+241=533], L[290+244=534], D[295+242=537], T[251+329=580], A[229+366=595], R[411+193=604], D[385+242=627], T[333+329=662], R[500+193=693], C[541+160=701]

Ex. 3

Considérons le problème du déplacement de k chevaliers à partir de k cases de départ $s_1, \cdots, s_k$ à k carrés $g_1, \cdots, g_k$, sur un échiquier illimité, sous réserve de la règle selon laquelle deux cavaliers ne peuvent pas atterrir sur la même case en même temps.

Chaque action consiste à déplacer jusqu'à k chevaliers simultanément. Nous aimerions terminer la manœuvre dans le plus petit nombre d'actions.

1. Quel est le facteur de branchement maximal dans cet espace d'états, exprimé en fonction de k ? 9^k. Pour chacun des k chevaliers, nous avons un des 8 coups en plus de la possibilité de ne pas bouger du tout ; chaque action exécute l'un de ces 9 choix pour chaque chevalier (sauf si certains choix entrent en conflit en atterrissant sur la même case), il y a donc 9^k actions.

2. Supposons que h_i soit une heuristique *admissible* pour le problème du déplacement du cavalier i vers le but g_i par lui-même. Parmi les heuristiques suivantes, lesquelles sont *admissibles* pour le problème k-knight ? De ceux-là, quel est le meilleur?

   (i) $\min\{h1, \cdots, h_k\}$

   (ii) $\max\{h1, \cdots, h_k\}$

   (iii) $\sum_{i=1}^{k} h_i$

   (i) et (ii) sont *admissibles*. Si h_i était exact, (ii) serait exact pour le problème relaxé où les cavaliers peuvent atterrir sur la même case. Les h_i sont *admissibles* et donc pas plus grands que les valeurs exactes, donc (ii) est *admissible*. (i) n'est pas plus grand que (ii), donc (i) est *admissible*. (iii) n'est pas recevable. Par exemple, si chaque g_i est à un mouvement de son s_i, alors (iii) renvoie k alors que le coût optimal de la solution est 1. (ii) domine (i) donc il doit être aussi bon ou meilleur. (iii) est probablement de peu de valeur car il n'est pas admissible et ignore complètement la possibilité de mouvements parallèles.

3. Répétez (2) pour le cas où vous n'êtes autorisé à déplacer qu'un seul cavalier à la fois. Dans ce cas, toutes sont *admissibles*. (i) et (ii) car le problème de la partie (2) est une relaxation du problème de cette partie, et (i) et (ii) sont *admissibles* pour le problème relaxé. (iii) est *admissible* car exacte pour le problème relaxé où les cavaliers peuvent atterrir sur la même case. (iii) est le meilleur puisqu'il domine le reste et est maintenant *admissible*.

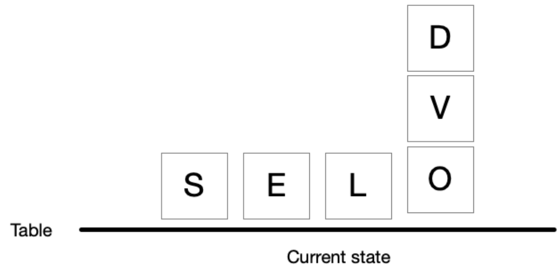


Figure 3: L'état de départ pour la question 2 ici.

3.1 Examen 2022-2023

Considérons l'utilisation de la recherche A* sur le problème visualisé dans la Figure 2. Dans ce contexte, nous utiliserions une fonction heuristique $h(n)$ pour estimer la qualité d'un état par rapport à l'état cible. Imaginez l'heuristique suivante : l'heuristique $h(n)$ attribue à chaque bloc mal placé un score de $(-1 * d)$, d étant égal à la distance de la position finale, et les blocs correctement positionnés obtiennent une note de (+1). Ainsi, l'état de départ de la Figure 2 a $h(n) = -1$ (+1 pour E, V et O; -1 pour D et L; -2 pour S) et le score de l'état de final est $h(n) = 6$ (+1 pour chaque block).

1. Quand une heuristique est-elle admissible ? Est-ce le cas pour l'heuristique proposée ? Expliquez. Si elle n'est pas admissible, proposez une alternative. (2 points) Une heuristique est admissible lorsque le coût qu'elle calcule est inférieur ou égal au coût réel de réalisation de la solution et supérieur ou égal à 0 ($0 \leq h(n) \leq h^*(n)$). L'heuristique actuelle peut être négative. C'est donc inacceptable (*not admissible*). Une solution serait de compter -1 pour chaque bloc égaré et $+1$ pour chaque bloc correctement positionné mais aussi de prendre le maximum entre et ce nombre pour éviter une valeur heuristique négative.

2. Tracez deux étapes de l'opération de A* au problème du monde de blocs en utilisant l'heuristique décrite dans l'énoncé. Soyez clair dans votre description et dans l'utilisation de l'espace d'états et des fonctions $f(n)$, $h(n)$ et $g(n)$. Partez d'un état avec les blocs S, E et L seuls sur la table et les blocs O, V et D empilés les uns sur les autres, avec O en bas, comme illustré dans la Figure 3. Le coût jusqu'à cet état est $g(n) = 5$. (3 points)

3. Quand une heuristique est-elle cohérente (*consistent*) ? Est-ce le cas pour l'heuristique proposée ici ? (1 points) Une heuristique est cohérente lorsque les coûts heuristiques d'un arc sont inférieurs ou égaux au coût réel : $h(a) - h(c) \leq c(a, c)$ ou $h(a) \leq c(a, c) + h(c)$. La conséquence est que le f-cost le long d'un chemin ne diminue jamais et la recherche de graphe A* est donc optimale. Dans l'exemple (b) vous pouvez voir que l'heuristique n'est pas cohérente si on suppose que le coût de la manipulation d'un bloc $c(s, a, s') = 1$.

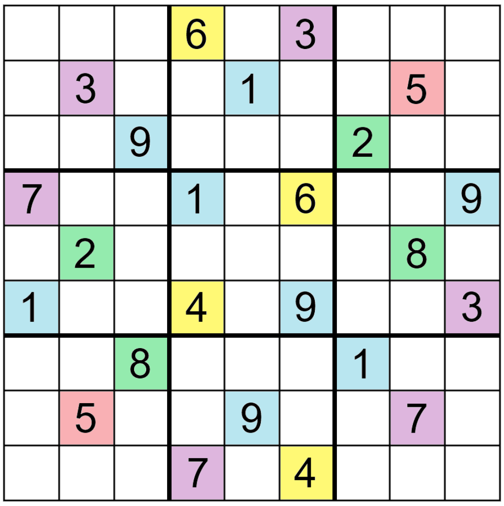


Figure 4: Image de Wikipedia.

4 Recherche locale

Ex. 1

Donnez le nom de l'algorithme qui résulte de chacun des cas particuliers suivants :

1. *Local beam search* avec $k = 1$. *Local beam search* avec $k = 1$ est une recherche par *hill climbing*

2. *Local beam search* avec un état initial et aucune limite sur le nombre d'états retenus. *Local beam search* avec un état initial et pas de limite sur le nombre d'états retenus, ressemble à *breadth-first search* en ce qu'elle ajoute une couche complète de nœuds avant d'ajouter la couche suivante. À partir d'un état, l'algorithme serait essentiellement identique à *breadth-first search*, sauf que chaque couche est générée en une seule fois.

3. *Simulated annealing* avec $T = 0$ à tout moment (et en omettant le test de terminaison). *Simulated annealing* avec $T = 0$ à chaque instant : en ignorant le fait que l'étape de terminaison serait déclenchée immédiatement, la recherche serait identique à *greedy hill climbing* car tout successeur descendant serait rejeté avec probabilité 1.

4. *Simulated annealing* avec $T = \infty$ à tout moment. Le *Simulated annealing* avec $T = \infty$ à tout moment est une recherche de marche aléatoire (*random-walk search*) : il accepte toujours un nouvel état.

5. Algorithme génétique avec taille de population $N = 1$. Algorithme génétique avec taille de population $N = 1$: si la taille de la population est 1, alors les deux parents sélectionnés seront le même individu ; le croisement donne

une copie exacte de l'individu; il y a cependant une petite chance de mutation. Ainsi, l'algorithme exécute une marche aléatoire (*random walk*) dans l'espace des individus.

Ex. 2

Définir les éléments (but, états, actions, fonction de valeur heuristique) du problème du Sudoku (Figure 4) afin qu'il puisse être résolu par:

Les définitions ne sont que deux parmi un certain nombre d'autres possibilités. Assurez-vous que ce que vous proposez a du sens compte tenu de ce que vous savez de la technique.

1. Hill climbing

    - But: Résolvez le puzzle de manière à ce que toutes les contraintes du Sudoku soient satisfaites (Chaque carré doit contenir un seul chiffre, de 1 à 9, le même chiffre ne peut pas apparaître deux fois dans la même ligne et le même chiffre ne peut pas apparaître deux fois dans la même colonne .)
    - Etats: Chaque configuration du puzzle, c'est-à-dire partiellement remplie avec des nombres compris entre 0 et 9, ne violant pas les 3 contraintes.
    - Actions: Ajouter un nombre dans $\{0, ..., 9\}$ dans une position vide afin qu'il ne viole pas les contraintes.
    - Fonction de valeur heuristique: Le nombre de positions non vides dans la grille de sudoku

2. Un algorithme génétique

    - But: Résolvez le puzzle de manière à ce que toutes les contraintes du Sudoku soient satisfaites (Chaque carré doit contenir un seul chiffre, de 1 à 9, le même chiffre ne peut pas apparaître deux fois dans la même ligne et le même chiffre ne peut pas apparaître deux fois dans la même colonne .)
    - Etats: Chaque configuration du puzzle complètement rempli de chiffres. Il y a un nombre limité de chaque nombre qui peut apparaître dans la grille afin que nous puissions nous assurer que cette contrainte est satisfaite dans chaque état.
    - Actions: Prenez deux solutions de la population en fonction de leur qualité (nombre de non-violations) et coupez-les horizontalement ou verticalement à n'importe quelle colonne ou rangée et recombinez les morceaux. Des opérations supplémentaires peuvent être imaginées qui échangent des sous-carrés de taille 3x3, etc. En tant que mutation, on peut imaginer changer un nombre à un endroit aléatoire quelque part dans la grille.
    - Fonction de valeur heuristique: Le nombre de positions qui ne violent aucune des contraintes

Pourquoi l'algorithme génétique est-il meilleur pour ce problème ? *Hill climbing* peut se coincer plus facilement dans un optimum local. L'algorithme génétique peut traverser de plus grandes zones de l'espace de recherche grâce au croisement, ce qui facilite la recherche d'un maximum global.

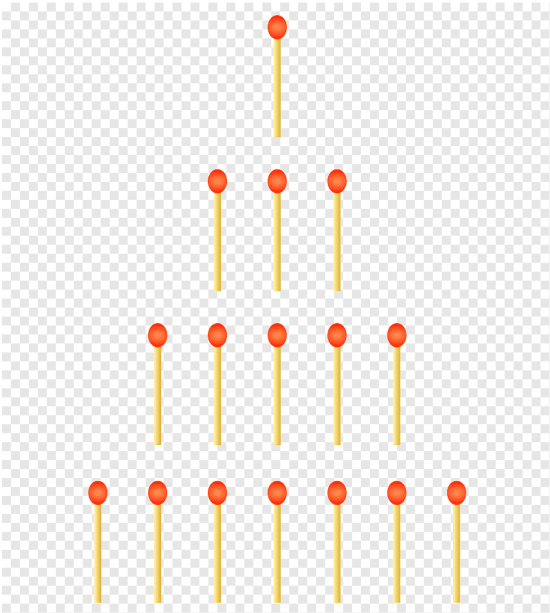


Figure 5: Image de pngwing.com.

5 Jeux et recherche d'adversaire

Ex. 1

Nim est un jeu à deux joueurs joué avec plusieurs tas de pierres. Vous pouvez utiliser autant de tas et autant de pierres dans chaque tas que vous le souhaitez. Les deux joueurs retirent à tour de rôle les pierres du jeu. À chaque tour, le joueur qui enlève des pierres ne peut prendre que des pierres d'un tas, mais il peut retirer autant de pierres de ce tas qu'il le souhaite. S'ils le souhaitent, ils peuvent même retirer tout le tas du jeu! Le gagnant est le joueur qui enlève la dernière pierre

1. Définissez le jeu en utilisant la formulation du jeu sur la slide 6 (états, joueurs, actions, modèle de transition ...)

   - État initial: p.ex. Figure 5.
   - Joueurs: deux joueurs, la fonction `Joueur(E)` déterminera à qui appartient le tour du jeu
   - Actions: Prenez un certain nombre de pièces, entre 1 ou toutes, d'un seul tas
   - Modèle de transition: Les transitions sont déterministes, la suppression d'un certain nombre de pièces dans un état donné ne conduira qu'à un seul état. De plus, les états antérieurs ne peuvent pas être visités car aucune pièce n'est réintroduite dans le jeu.
   - Teste de fin : Aucune pièce restante dans aucun des tas
   - Utilité: $+1$ quand on gagne et -1 quand on perd.

2. Est-ce un jeu à somme nulle ou un jeu à somme générale ? Jeu à somme nulle

3. Imaginez une configuration de 3 tas, avec une pierre dans le premièr tas, 2 dans le second et 1 dans le troisième tas. Dessinez l'arbre de jeu (regardez Figure 6) et montrez avec une analyse minimax quelles sont les meilleures options pour le premier joueur.

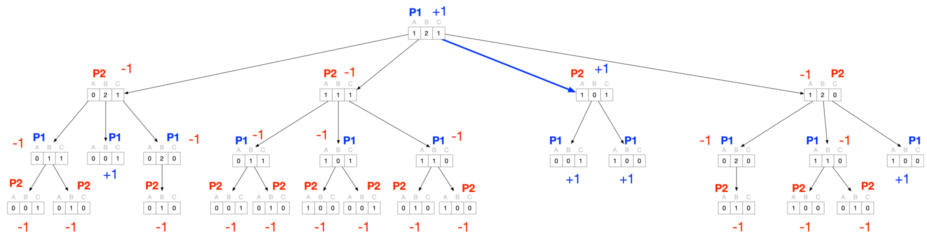


Figure 6: Arbre Minimax pour le jeu Nim

Analyse Minimax : Dans cet arbre de jeu, la seule action qui garantit le succès est de retirer les 2 pièces du tas central.

4. Peut-on appliquer $\alpha - \beta$ *pruning* ? Comment cela fonctionnerait-il ? oui, dès qu'un sous-arbre commençant à un nœud du joueur 2 donne -1, il est certain que le joueur 2 choisira celui-là. Toutes les autres actions peuvent ainsi être élaguées (Figure 7).

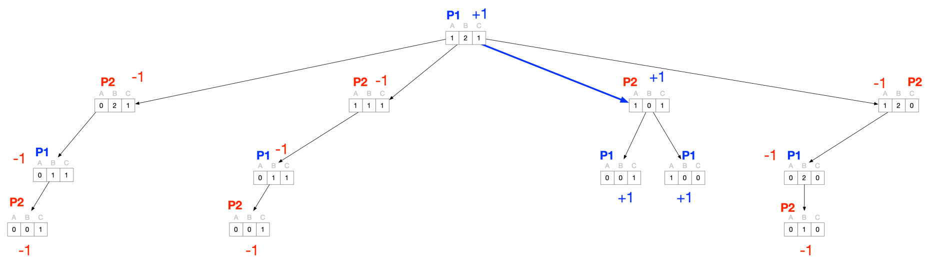


Figure 7: Arbre Minimax pour le jeu Nim avec $\alpha - \beta$ pruning.

5. Quelque chose changera-t-il dans l'analyse précédente si nous jouons NIM sous une forme misère (le joueur qui prend le dernier objet perd)

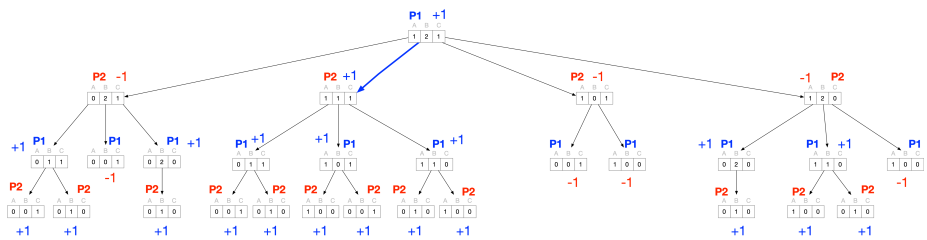


Figure 8: Arbre Minimax pour le jeu Nim sous forme de misère

Il y a maintenant beaucoup plus de résultats qui ont une utilité positive pour le premier joueur. la meilleure action est maintenant de retirer seulement 1 pièce du deuxième tas car cela garantit une victoire pour le premier joueur.

6. Imaginez qu'on a dix tas dont le 1er contient 3, le 2ème 4, le 3ème 5 objets, etc. Le nombre d'actions possibles augmente rapidement et nous ne pourrons peut-être pas calculer comme avant les feuilles de l'arbre de recherche. Quelle pourrait être une fonction d'évaluation utile pour permettre une anticipation limitée ?

   Une fonction d'évaluation potentielle pourrait évaluer si l'on pouvait gagner en supprimant des tas complets, alors qu'en réalité on pourrait également supprimer un élément à la fois. Il existe bien sûr d'autres fonctions, alors assurez-vous d'expliquer votre fonction afin qu'elle ait un sens pour un lecteur.

7. Comment MCTS (*Monte-Carlo Tree search*) pourrait-il aider à résoudre ce jeu ? Supposons que la politique de déploiement (*roll-out policy*) supprime simplement une pièce au hasard dans n'importe quel tas.

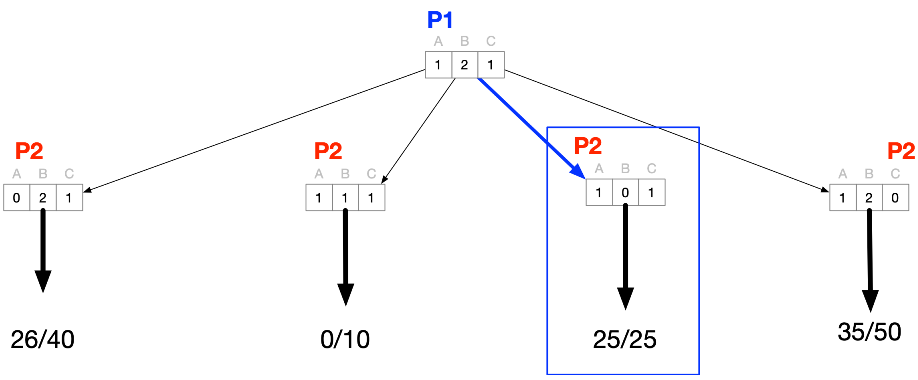


Figure 9: MCTS pour le jeu Nim

   Comme on peut le voir dans la Figure 9, nous pouvons faire un certain nombre de déploiements et voir à quelle fréquence nous gagnons le jeu pour chacun des 4 actions du joueur 1. Étant donné qu'il y a toujours une victoire pour le troisième, nous pouvons choisir celui-là, mais gardez à l'esprit que tout dépend du nombre de tests que on peut faire (ici 125 en total).

5.1 Examen 2022-2023

Considérez l'arbre *minimax* de la Figure 10, répondez aux questions suivantes:

1. Quelle est la valeur *minimax* du nœud A dans l'arbre ? expliquez votre réponse. (2 points) La valeur de $A = 12$. En regardant les feuilles à la fin de l'action de gauche, le joueur 1 préférera le score 21, le deuxième joueur le sait et préférera donc effectuer l'action qui mène au score de 12 car c'est le minimum de 12 et 21. L'autre branche au milieu conduira le deuxième

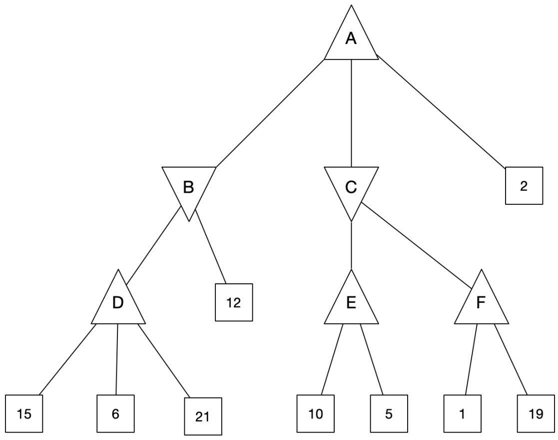


Figure 10

joueur à aller chercher la branche E et donc un score de 10. Le joueur 1 doit donc choisir entre une action qui produit 12, 10 ou 2, où 12 est donc préféré.

2. Enlevez les nœuds qui sont élagués par l'*alpha-beta pruning*. Supposons le parcours standard de gauche à droite de l'arbre. Si un état non terminal (A, B, C, D, E ou F) est élagué, enlevez tout le sous-arbre. (2 points) Toute la branche F est supprimée.

5.2 Examen 2022-2023

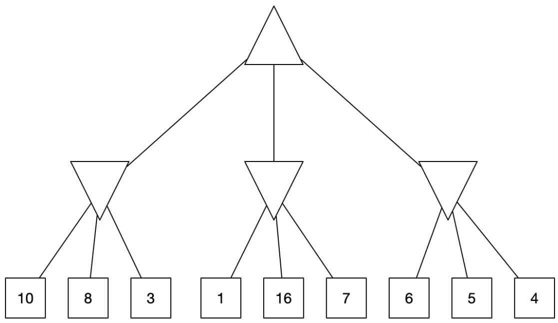


Figure 11

1. Considérez l'arbre de jeu à somme nulle dans la Figure 11. Les triangles qui pointent vers le haut, comme au nœud supérieur (racine), représentent des choix pour le joueur qui maximise ; les triangles qui pointent vers le bas représentent des choix pour le joueur qui minimise. En supposant que les deux joueurs agissent de manière optimale, remplissez la valeur *minimax* de chaque nœud. Y a-t-il des nœuds qui seront supprimés par $\alpha - \beta$ *pruning* ? Expliquez votre réponse. (2 points) Les valeurs dans les noeud du joueur adverse sont respectivement 3,1 et 4. et pour le joueur maximisant, le score dans le nœud est 4. Le joueur doit donc jouer l'action

associée avec la branche droite. Les noeud avec les valeurs 16 et 7 seront supprimé par α-β *pruning*.

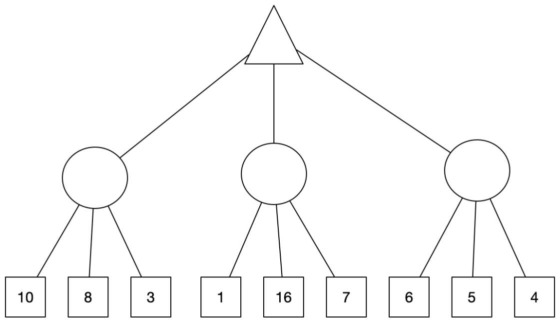


Figure 12

2. Considérons le même arbre de jeu (Figure 12) à somme nulle, sauf que maintenant, au lieu d'un joueur minimisant, nous avons un nœud aléatoire qui sélectionnera une valeur uniformément au hasard. Remplissez la valeur *expectimax* de chaque nœud. Y a-t-il des nœuds qui seront supprimés par $\alpha - \beta$ *pruning*? Expliquez votre réponse. (2 points) Aucun noeud sera supprimé par α-β *pruning*. Il y aura toujours la possibilité qu'une feuille non visitée du noeud aléatoire ait une valeur très élevée, ce qui augmente la moyenne globale de ce noeud. Par exemple, lorsque nous voyons que la feuille 4 vaut 1, ce qui est bien inférieur à la valeur du noeud de chance de gauche, 7, à ce stade, nous ne pouvons faire aucune hypothèse sur la façon dont la valeur du noeud de chance du milieu sera finalement supérieure ou égale à 1. Il s'avère que la feuille 5 a une valeur de 16, ce qui porte la valeur attendue du noeud de chance du milieu à 8.5, ce qui est supérieur à la valeur du noeud de chance de gauche. Dans le cas où il existe une limite supérieure sur la valeur d'un noeud feuille, il existe une possibilité d'élagage.

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Figure 13: Image de [1].

A	$P(A)$
t	0.8
f	0.2

A	B	$P(B\|A)$
t	t	0.9
t	f	0.1
f	t	0.6
f	f	0.4

B	C	$P(C\|B)$
t	t	0.8
t	f	0.2
f	t	0.8
f	f	0.2

C	D	$P(D\|C)$
t	t	0.25
t	f	0.75
f	t	0.5
f	f	0.5

Figure 14: Image de [1].

6 Les bases de la théorie des probabilités

Ex. 1

Compte tenu de la distribution conjointe complète illustrée dans la Figure 13, calculez:

1. $P(+\mathsf{toothache})$ On demande la probabilité que Toothache soit vrai: $\mathsf{P}(+\mathsf{toothache}) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

2. $P(\mathsf{Cavity})$ On demande le vecteur des valeurs de probabilité pour la variable Cavity. Il a deux valeurs, que nous listons dans l'ordre $\langle \mathsf{true}, \mathsf{false} \rangle$. Additionnez d'abord $0.108 + 0.012 + 0.072 + 0.008 = 0.2$. Ensuite nous avons $P(\mathsf{Cavity}) = \langle 0.2, 0.8 \rangle$.

3. $P(\mathsf{Toothache}|\mathsf{cavity})$ Ici on demande le vecteur des valeurs de probabilité pour Toothache, étant donné que Cavity est vrai: $P(Toothache|+cavity) = \langle (.108 + .012)/0.2, (0.072 + 0.008)/0.2 \rangle = \langle 0.6, 0.4 \rangle$

4. $P(\mathsf{Cavity}|\mathsf{toothache} \vee \mathsf{catch})$ On demande le vecteur des valeurs de probabilité pour Cavity, étant donné que Toothache ou Catch est vrai. Calculez d'abord $P(+\mathsf{tootache} \vee +\mathsf{catch}) = 0,108 + 0,012 + 0,016 + 0,064 + 0,072 + 0,144 = 0,416$. Alors $P(\mathsf{Cavity}|\mathsf{toothache} \vee \mathsf{catch}) = \langle (0.108 + 0.012 + 0.072)/0.416, (0.016 + 0.064 + 0.144)/0.416 \rangle = \langle 0.4615, 0.5384 \rangle$

Ex. 2

Considérez les distributions de probabilité de la Figure 14.

Étant donné uniquement ces tableaux et aucune hypothèse d'indépendance, calculez les probabilités suivantes, en montrant votre travail. S'il est impossible de calculer sans plus d'hypothèses d'indépendance, spécifiez plutôt un ensemble minimal d'hypothèses d'indépendance qui vous permettrait de répondre à la question.

1. $P(+a,-b)$ $P(+a,-b) = P(+a)P(-b|+a) = 0.8 \times 0.1 = 0.08$.

2. $P(+b)$ $P(+b) = \sum_a P(+b|a)P(a) = P(+b|+a)P(+a) + P(+b|-a)P(-a) = (0.9 \times 0.8) + (0.6 \times 0.2) = 0.84$.

3. $P(-a,-b,+c)$ pas possible, il faut avoir que C est indépendant de A étant donné B.

4. Supposons maintenant que C est indépendant de A étant donné B et D est indépendant de A et B étant donné C. Calculez $P(+a,-b,+c,+d)$ Utilisez *le chain rule*:

$$\begin{aligned}
P(+a,-b,+c,+d) &= P(-b,+c,+d|+a)P(+a) \\
&= P(+c,+d|+a,-b)P(-b|+a)P(+a) \\
&= P(+d|+a,-b,+c)P(+c|+a,-b)P(-b|+a)P(+a) \\
&= P(+d|+c)P(+c|-b)P(-b|+a)P(+a) \\
&= 0.25 \times 0.8 \times 0.1 \times 0.8 = 0.016
\end{aligned}$$

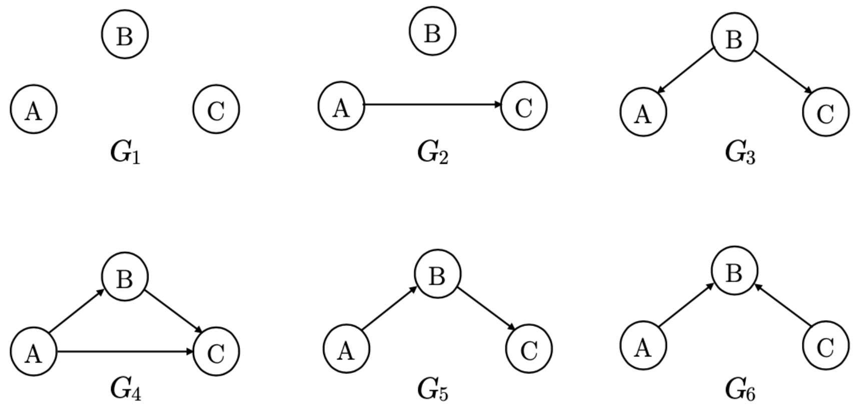


Figure 15: Image de [1].

7 Réseaux Bayésiens

Ex. 1

Rappelons que chaque graphe orienté acyclique G a une famille de distributions de probabilités associée, qui se compose de toutes les distributions de probabilité qui peuvent être représentées par un réseau de Bayes avec la structure G. Considérez les six graphiques acycliques orientés de la Figure 15.

1. Supposons que tout ce que nous savons de la distribution conjointe $P(A, B, C)$ est qu'elle peut être représentée par le produit $P(A|BC)P(B|C)P(A)$. Parmi les six graphes, lesquels sont garantis capables de représenter $P(A, B, C)$? Dans ce cas, la décomposition de $P(A, B, C)$ est juste celle donnée par le *chain rule*, donc aucune indépendance conditionnelle n'est impliquée. $G4$ (Figure 15) est entièrement connecté, et est donc capable de représenter n'importe quelle distribution conjointe. Les autres affirment des indépendances conditionnelles et peuvent donc ne pas être en mesure de représenter $P(A, B, C)$.

2. Supposons maintenant que tout ce que nous savons de la distribution conjointe $P(A, B, C)$ est qu'elle peut être représentée par le produit $P(C|B)P(C|A)P(A)$. Parmi les six graphes, lesquels sont garantis capables de représenter $P(A, B, C)$? $G3$, $G4$ et $G5$ peuvent représenter la distribution. $G1$ suppose que toutes les variables sont indépendantes ; $G2$ affirme que B est indépendant des autres ; et $G6$ affirme $A \perp C$.

Ex. 2

Considérons un réseau de Bayes sur les variables aléatoires A, B, C, D et E avec la structure montrée dans la Figure 16, avec une distribution conjointe complète $P(A, B, C, D, E)$.

1. Considérons la distribution marginale $P(A, B, D, E) = \sum_c P(A, B, c, D, E)$ où C a été éliminé. Dessinez le réseau de Bayes minimal qui est garanti

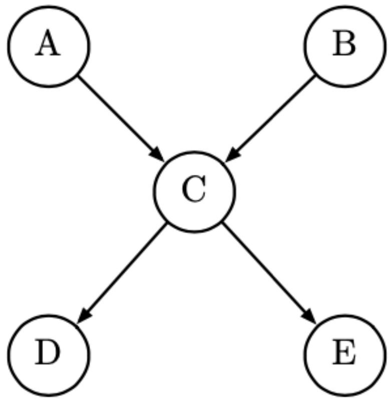


Figure 16: Image de [1].

pour pouvoir représenter cette distribution. Pour $P(A, B, D, E)$ on a la

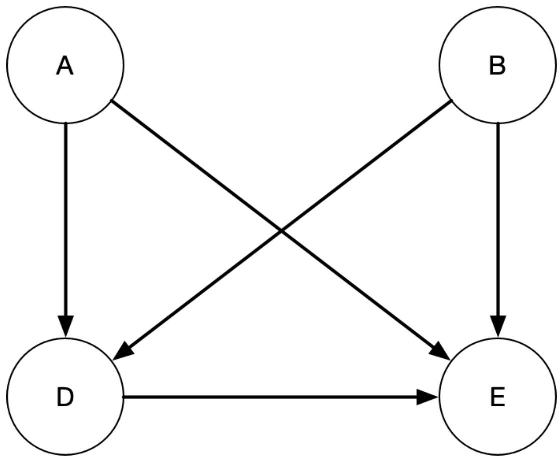


Figure 17: Image de [1].

Figure 17. La vue générale pour ce type de problèmes est que le graphe résultant doit être capable d'encoder les mêmes hypothèses d'indépendance conditionnelle à partir du réseau Bayes initial. Dans le BN ci-dessus, nous avons les hypothèses d'indépendance suivantes :

- $A \perp\!\!\!\perp B$
- $A \perp\!\!\!\perp D|C$
- $B \perp\!\!\!\perp D|C$
- $A \perp\!\!\!\perp E|C$
- $B \perp\!\!\!\perp E|C$
- $D \perp\!\!\!\perp E|C$

Lorsque nous marginalisons C, nous supprimons C du graphe. Les hypothèses d'indépendance conditionnelle impliquant C n'ont plus d'importance, il suffit donc de préserver : $A \perp\!\!\!\perp B$. Le graphe présenté le garantit. La flèche entre D et E peut aller dans les deux sens.

2. Supposons qu'on nous donne une observation : $A = a$. Dessinez le réseau de Bayes minimal qui est garanti pouvoir représenter la distribution $P'(B, C, D, E) = P(B, C, D, E|A = a)$. Pour $P'(B, C, D, E)$ on a

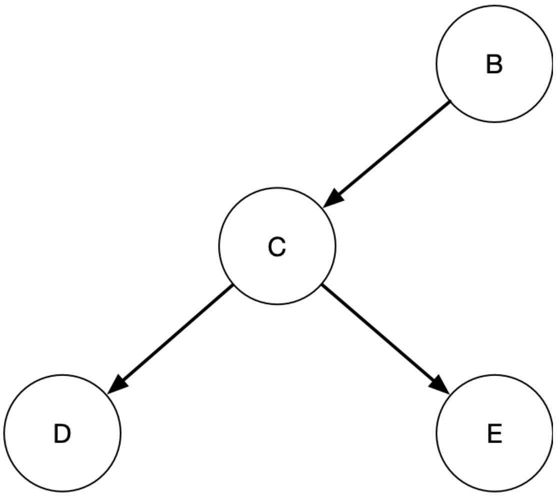


Figure 18: Image de [1].

la Figure 18. L'observation de A fixe sa valeur et la retire du réseau de Bayes. Par *d-séparation* aucune autre dépendance n'est introduite.

Ex. 3

Nous avons un sac de trois pièces biaisées a, b et c avec des probabilités de sortir face de 30 %, 60 % et 75 %, respectivement. Une pièce est tirée au hasard du sac (avec une probabilité égale de tirer chacune des trois pièces), puis la pièce est lancée trois fois pour générer les résultats X_1, X_2 et X_3.

1. Dessinez le réseau Bayésien correspondant à cette configuration et définissez les tables de probabilités conditionnelles nécessaires.

   Avec la variable aléatoire C indiquant quelle pièce $\{a, b, c\}$ nous avons tirée, le réseau a C à la racine et X_1, X_2 et X_3 comme enfants. Le CPT pour C est :

C	$P(C)$
a	1/3
b	1/3
c	1/3

   Les CPT pour X_i étant donné C sont les mêmes et égaux à :

C	X_i	$P(X_i\|C)$
a	heads	0.3
b	heads	0.6
c	heads	0.75

2. Calculez quelle pièce a été la plus susceptible d'avoir été tirée du sac si les *flips* observés sortent face deux fois et pile une fois. La pièce la plus susceptible d'avoir été tirée du sac compte tenu de cette séquence est la valeur de C avec la plus grande probabilité a posteriori $P(C|2\ \mathsf{faces}, 1\ \mathsf{pile})$. Donc,

$$\begin{aligned} P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C)P(C)/P(2\ \mathsf{faces}, 1\ \mathsf{pile}) &\propto P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C)P(C) \\ &\propto P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C) \end{aligned}$$

où dans la deuxième partie on observe que la constante de proportionnalité $1/P(2\ \mathsf{faces}, 1\ \mathsf{pile})$ est indépendante de C, et dans la dernière on observe que $P(C)$ est également indépendant de la valeur de C puisqu'il est, par hypothèse, égal à $1/3$.

À partir du réseau Bayésien, nous pouvons voir que X_1, X_2 et X_3 sont conditionnellement indépendants étant donné C, donc par exemple

$$\begin{aligned} &P(X_1 = \mathsf{pile}, X_2 = \mathsf{face}, X_3 = \mathsf{face}|C = a) \\ &= P(X_1 = \mathsf{pile}|C = a)P(X_2 = \mathsf{face}|C = a)P(X_3 = \mathsf{face}|C = a) \\ &= 0.7 \times 0.3 \times 0.3 = 0.063 \end{aligned}$$

Notez que puisque les CPT pour chaque pièce sont les mêmes, nous obtiendrions la même probabilité ci-dessus pour tout ordre de 2 faces et 1 pile. Puisqu'il existe trois ordres de ce type, nous avons $P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C = a) = 3 \times 0.063 = 0.189$. Des calculs similaires à ceux ci-dessus montrent que

- $P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C = b) = 0.432$.
- $P(2\ \mathsf{faces}, 1\ \mathsf{pile}|C = c) = 0.422$.

montrant que la pièce b est la plus susceptible d'avoir été tirée.

Ex. 4

Considérez le réseau Bayésien dans la Figure 19.

1. Si aucune preuve n'est observée, le Burglary et le EarthQuake sont-ils indépendants ? Démontrez ceci à partir de la sémantique numérique et de la sémantique topologique. Oui. De la sémantique numérique, nous

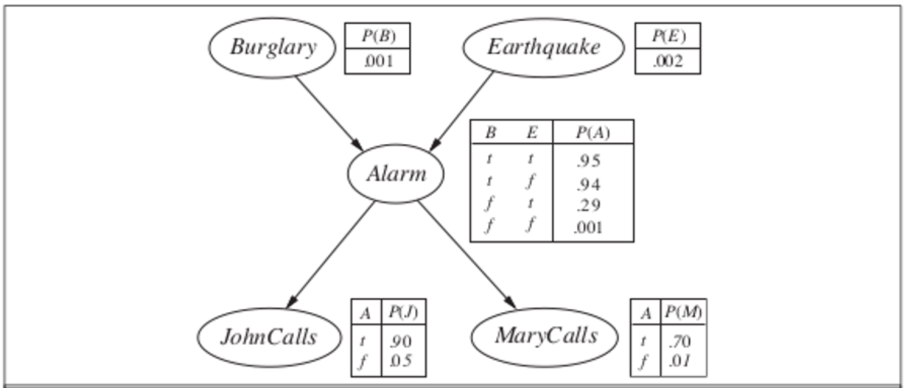


Figure 14.2 A typical Bayesian network, showing both the topology and the conditional probability tables (CPTs). In the CPTs, the letters B, E, A, J, and M stand for *Burglary*, *Earthquake*, *Alarm*, *JohnCalls*, and *MaryCalls*, respectively.

Figure 19: Image de [1].

avons:

$$\begin{aligned}
P(B,E) &= \sum_{a,j,m} P(B,E,a,j,m) = \sum_{a,j,m} P(B)P(E)P(a|B,E)P(j|a)P(m|a) \\
&= P(B)P(E) \sum_{a,j,m} P(a|B,E)P(j|a)P(m|a) \\
&= P(B)P(E) \sum_{a} P(a|B,E) \left(\sum_{j} P(j|a) \right) \left(\sum_{m} P(m|a) \right) \\
&= P(B)P(E)
\end{aligned}$$

en utilisant la contrainte *sum-to-1* pour les distributions conditionnelles. Topologiquement, E est indépendant de ses non-descendants (c'est-à-dire B) étant donné ses parents (c'est-à-dire l'ensemble vide), donc B et E sont absolument indépendants.

2. Si nous observons $\mathsf{Alarm} = \mathsf{true}$, $\mathsf{Burglary}$ et $\mathsf{EarthQuake}$ sont-ils indépendants ? Justifiez votre réponse en calculant si les probabilités en jeu satisfont à la définition de l'indépendance conditionnelle. Nous vérifions si $P(B,E|a) =$

X	$P(X)$
0	0.75
1	0.25

X	W	$P(W\|X)$
0	0	0.4
0	1	0.6
1	0	0.4
1	1	0.6

X	Y	$P(Y\|X)$
0	0	0.3
0	1	0.7
1	0	0.1
1	1	0.9

Z	W	$P(W\|Z)$
0	0	0.2
0	1	0.8
1	0	0.8
1	1	0.2

Figure 20: Image de [1].

$P(B|a)P(E|a)$. En premier lieu, calculez $P(B,E|a)$:

$$
\begin{aligned}
P(B,E|a) &= \alpha P(a|B,E)P(B,E) \\
&= \alpha \begin{cases} .95 \times 0.001 \times 0.002 & \text{si } B = b \text{ et } E = e \\ .94 \times 0.001 \times 0.998 & \text{si } B = b \text{ et } E = -e \\ .29 \times 0.999 \times 0.002 & \text{si } B = -b \text{ et } E = e \\ .001 \times 0.999 \times 0.998 & \text{si } B = -b \text{ et } E = -e \end{cases} \\
&= \alpha \begin{cases} 0.0008 & \text{si } B = b \text{ et } E = e \\ 0.3728 & \text{si } B = b \text{ et } E = -e \\ 0.2303 & \text{si } B = -b \text{ et } E = e \\ 0.3962 & \text{si } B = -b \text{ et } E = -e \end{cases}
\end{aligned}
$$

où α est une constante de normalisation. En vérifiant si $P(b,e|a) = P(b|a)P(e|a)$ on trouve

$$P(b,e|a) = 0.0008 \neq 0.0863 = 0.3736 \times 0.2311 = P(b|a)P(e|a)$$

montrant que B et E ne sont pas conditionnellement indépendants étant donné A

Ex. 5

On vous donne les distributions conditionnelles qui relient les variables binaires dans la Figure 20.

Lequel des réseaux bayésiens de la Figure 21 peut représenter une distribution conjointe cohérente avec ces distributions conditionnelles ? Parmi ceux-ci, lesquels sont minimes, dans le sens où aucun arc ne peut être supprimé tout en conservant la propriété ?

Rappelez-vous que la topologie d'un réseau Bayésien revendique une indépendance conditionnelle par l'absence de flèches. Les trois premiers réseaux affirment que W et Z sont absolument indépendants, ce qui n'est pas le cas quand on regarde les distributions conditionnelles dans la Figure 20. En examinant la distribution $P(W|X)$, nous voyons que la probabilité de W ne dépend pas réellement de X, donc W et X sont absolument indépendants. Les seules dépendances requises sont entre X et Y et entre W et Z. Les trois réseaux de la deuxième ligne capturent ces dépendances. (La direction de la flèche n'a pas d'importance.) Le quatrième réseau a une flèche superflue de W à X, donc les cinquième et sixième réseaux sont le configurations minimales.

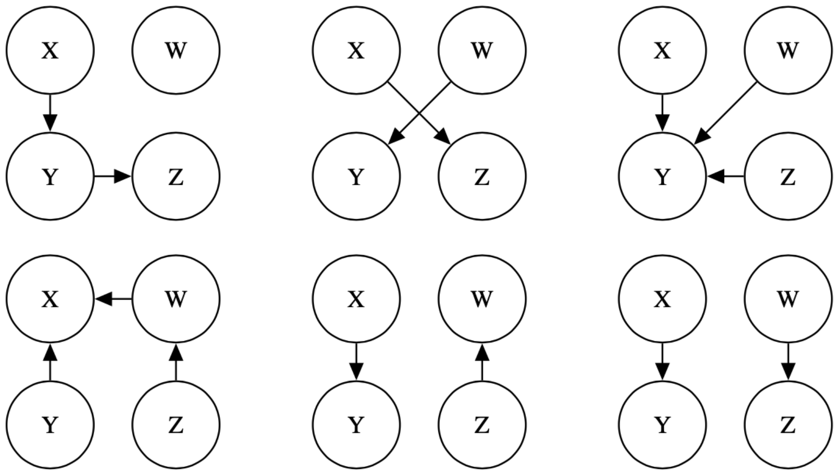
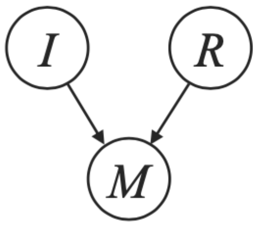


Figure 21: Image de [1].

I	$P(I)$
$+i$	0.8
$-i$	0.2

R	$P(R)$
$+r$	0.4
$-r$	0.6



I	R	M	$P(M\|I,R)$
$+i$	$+r$	$+m$	0
$+i$	$+r$	$-m$	1.0
$+i$	$-r$	$+m$	0
$+i$	$-r$	$-m$	1.0
$-i$	$+r$	$+m$	0.7
$-i$	$+r$	$-m$	0.3
$-i$	$-r$	$+m$	0.2
$-i$	$-r$	$-m$	0.8

Figure 22: Image de [1].

Ex. 6

Il y a eu une épidémie d'oreillons dans votre collège. Vous vous sentez bien, mais vous êtes inquiets que vous puissiez déjà être infecté. Vous décidez d'utiliser des reseaux Bayésiens pour analyser la probabilité que vous ayez contracté les oreillons. Pensez d'abord aux deux facteurs suivants :

1. Vous pensez être immunisé contre les oreillons ($+i$) car vous avez été vacciné récemment, mais le vaccin n'est pas totalement efficace, vous n'êtes donc peut-être pas immunisé ($-i$).

2. Votre colocataire ne se sentait pas bien hier, et bien que vous n'en soyez pas encore sûr, vous soupçonnez qu'il pourrait avoir les oreillons ($+r$).

Notons ces variables aléatoires par I et R. Soit la variable aléatoire M prendra la valeur $+m$ si vous avez les oreillons, et $-m$ si vous ne les avez pas. Le réseau Bayésien est décrit par la Figure 22. Elle vous donne la possibilité de décrire vos chances d'être malade :

I	R	M	$P(I,R,M)$
$+i$	$+r$	$+m$	0
$+i$	$+r$	$-m$	
$+i$	$-r$	$+m$	0
$+i$	$-r$	$-m$	
$-i$	$+r$	$+m$	0.056
$-i$	$+r$	$-m$	0.024
$-i$	$-r$	$+m$	0.024
$-i$	$-r$	$-m$	0.096

Figure 23: Image de [1].

1. Remplissez le tableau dans la Figure 23 avec la distribution conjointe sur I, M et R: $P(I, M, R)$.

I	R	M	$P(I,R,M)$
+i	+r	+m	0
+i	+r	-m	**0.32**
+i	-r	+m	0
+i	-r	-m	**0.48**
-i	+r	+m	0.056
-i	+r	-m	0.024
-i	-r	+m	0.024
-i	-r	-m	0.096

2. Quelle est la probabilité marginale $P(+m)$ que vous ayez les oreillons ?

$$\begin{aligned} P(+m) &= \sum_{i,r} P(i,r,+m) \\ &= P(+i,+r,+m) + P(+i,-r,+m) + P(-i,+r,+m) + P(-i,-r,+m) \\ &= 0 + 0 + 0.056 + 0.024 = 0.08 \end{aligned}$$

3. En supposant que vous ayez les oreillons, vous craignez que votre colocataire soit également atteint par la maladie. Quelle est la probabilité $P(+r|+m)$ que votre colocataire ait les oreillons sachant que vous avez les oreillons ? Notez que vous ne savez toujours pas si oui ou non vous avez l'immunité.

$$P(+r|+m) = \frac{P(+r,+m)}{P(+m)} = \frac{\sum_i P(i,+r,+m)}{P(+m)} = \frac{0+0.056}{0.080} = 0.143$$

Vous n'êtes toujours pas sûr d'avoir suffisamment d'informations sur vos chances d'avoir les oreillons, alors vous décidez d'inclure deux nouvelles variables

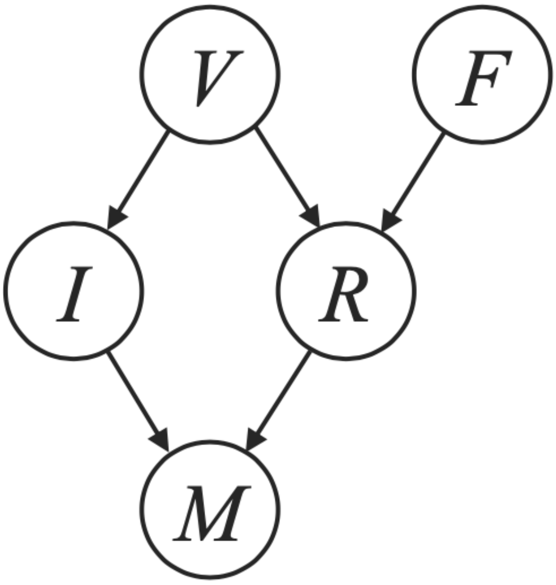


Figure 24: Image de [1].

dans le réseau Bayésien. Votre colocataire est allé à une fête ce week-end, et il y a des chances qu'une autre personne à la fête ait eu les oreillons ($+f$). De plus, vous et votre colocataire avez été vaccinés dans une clinique qui a signalé une confusion de vaccins. Que vous ayez reçu ou non le bon vaccin ($+v$ ou $-v$) a des ramifications à la fois sur votre immunité (I) et sur la probabilité que votre colocataire ait depuis contracté la maladie (R). En tenant compte de ceci, vous dessinez le réseau de Bayes modifié illustré dans la nouvelle Figure 24.

Parmi les affirmations suivantes, lesquelles sont garanties vraies pour ce réseau bayésien ?

1. $V \perp\!\!\!\perp M | I, R$
2. $V \perp\!\!\!\perp M | R$
3. $M \perp\!\!\!\perp F | R$
4. $V \perp\!\!\!\perp F$
5. $V \perp\!\!\!\perp F | M$
6. $V \perp\!\!\!\perp F | I$

Les assertions (1), (4), et (6) sont garanties d'être vraies.

7.1 Examen 2022-2023

Le réseau Bayésien dans la Figure 25 peut être utilisé pour diagnostiquer si un patient souffre d'un simple rhume (C) et/ou de la plus dangereuse grippe martienne (M), sur la base des symptômes du patient - que le patient ait ou non le nez qui coule (R), que le patient ait mal à la tête ou non (H), et que

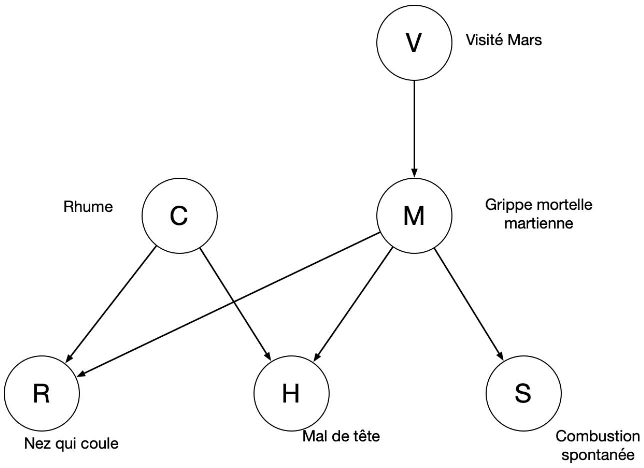


Figure 25: Réseau Bayésien composé de 5 nœuds, c'est-à-dire que vous ayez visité Mars ou non (V), que vous ayez ou non un rhume (C), que vous ayez la grippe martienne mortelle (M), que vous ayez le nez qui coule (R), mal de tête (H) ou que vous ayez pris feu spontanément (S) à un moment donné. Les arcs modélisent les dépendances conditionnelles entre les variables (V, M, C, R, H et S).

le patient s'enflamme spontanément ou non (S) – ainsi que des informations contextuelles pertinentes, à savoir s'il a déjà visité Mars ou non (V). Toutes les variables (V, M, C, R, H, S) sont binaires et sont associées aux tables de probabilités conditionnelles dans le Figure 26.

V	P(V)
+v	0.9999
-v	0.0001

C	P(C)
+c	0.95
-c	0.05

V	M	P(M|V)
+v	+m	0.001
-v	+m	0.0
+v	-m	0.999
-v	-m	1.0

M	S	P(S|M)
+m	+s	0.8
-m	+s	0.0
+m	-s	0.2
-m	-s	1.0

C	M	R	P(R|M,C)
+c	+m	+r	0.98
+c	+m	-r	0.02
+c	-m	+r	0.9
+c	-m	-r	0.1
-c	+m	+r	0.5
-c	+m	-r	0.5
-c	-m	+r	0.05
-c	-m	-r	0.95

C	M	H	P(H|M,C)
+c	+m	+h	0.99
+c	+m	-h	0.01
+c	-m	+h	0.6
+c	-m	-h	0.4
-c	+m	+h	0.98
-c	+m	-h	0.02
-c	-m	+h	0.07
-c	-m	-h	0.93

Figure 26

Résolvez maintenant les questions suivantes. Montrez chaque fois les formules que vous avez utilisées pour faire le calcul.

1. Calculez l'entrée suivante à partir de la distribution jointe : $P(+r, +h, -s, -c, +m, +v)$. (1 point) $P(+r, +h, -s, -c, +m, +v) = P(+v)P(-c)P(+m| + v)P(-s| +$

$m)P(+r|-c,+m)P(+h|-c,+m) = 0.9999 \times 0.05 \times 0.001 \times 0.2 \times 0.5 \times 0.98 = 0.0000048995$

2. Calculez l'entrée suivante à partir de la distribution jointe : $P(-r,+h,-s,+c,+m,+v)$. (1 point) $P(-r,+h,-s,+c,+m,+v) = P(+v)P(+c)P(+m|+v)P(-s|+m)P(-r|+c,+m)P(+h|+c,+m) = 0.9999 \times 0.95 \times 0.001 \times 0.2 \times 0.02 \times 0.99 = 0.0000037616238$

3. Quelle est la probabilité qu'une personne n'ayant pas de rhume(-c) et ayant visité Mars (+v) ait mal à la tête (+h)? (3 points)

$$
\begin{aligned}
P(+h|+v,-c) &= \frac{P(+h,+v,-c)}{P(+h,+v,-c)+P(-h,+v,-c)} \\
P(+h,+v,-c) &= \sum_m P(+h,-c,m,+v) \\
&= \sum_m P(+h|-c,m)P(-c)P(m|+v)P(+v) \\
&= P(+v)P(-c)\sum_m P(m|+v)P(+h|-c,m) \\
&= P(+v)P(-c)\left(P(+m|+v)P(+h|-c,+m)+P(-m|+v)P(+h|-c,-m)\right) \\
&= 0.9999 \times 0.05 \times (0.001 \times 0.98 + 0.999 \times 0.07) \\
&= 0.9999 \times 0.05 \times (0.00098 + 0.06993) \\
&= 0.9999 \times 0.05 \times 0.07091 \\
&= 0.003545145...
\end{aligned}
$$

vous devez faire le même calcul pour $P(-h,+v,-c)$ puis l'utiliser pour normaliser les valeurs.

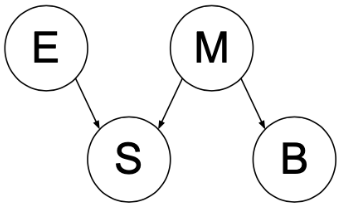
4. Quelle est la probabilité qu'une personne ayant visité Mars $(+v)$ ait mal à la tête $(+h)$? (3 points)

$$
\begin{aligned}
P(+h|+v) &= \frac{P(+h,+v)}{P(+h,+v)+P(-h,+v)} \\
P(+h,+v) &= \sum_{c,m} P(+h,c,m,+v) \\
&= \sum_{c,m} P(+h|c,m)P(c)P(m|+v)P(+v) \\
&= P(+v)\sum_{m} P(m|+v)\sum_{c} P(+h|c,m)P(c) \\
&= P(+v)\Bigg(P(+m|+v)\sum_{c} P(+h|c,+m)P(c)+ \\
&\quad P(-m|+v)\sum_{c} P(+h|c,-m)P(c)\Bigg) \\
\sum_{c} P(+h|c,+m)P(c) &= P(+h|+c,+m)P(+c)+P(+h|-c,+m)P(-c) \\
&= 0.99\times 0.95+0.98\times 0.05 = 0.9895 \\
\sum_{c} P(+h|c,+m)P(c) &= P(+h|+c,+m)P(+c)+P(+h|-c,+m)P(-c) \\
&= 0.6\times 0.95+0.07\times 0.05 = 0.5735 \\
P(+h,+v) &= P(+v)\left(P(+m|+v)0.9895+P(-m|+v)0.5735\right) \\
&= 0.9999\left(0.001\times 0.9895+0.999\times 0.5735\right) \\
&= 0.5738...
\end{aligned}
$$

vous devez faire le même calcul pour $P(-h,+v)$ puis l'utiliser pour normaliser les valeurs.

$P(E)$	
$+e$	0.4
$-e$	0.6

$P(S \mid E, M)$			
$+e$	$+m$	$+s$	1.0
$+e$	$+m$	$-s$	0.0
$+e$	$-m$	$+s$	0.8
$+e$	$-m$	$-s$	0.2
$-e$	$+m$	$+s$	0.3
$-e$	$+m$	$-s$	0.7
$-e$	$-m$	$+s$	0.1
$-e$	$-m$	$-s$	0.9



$P(M)$	
$+m$	0.1
$-m$	0.9

$P(B \mid M)$		
$+m$	$+b$	1.0
$+m$	$-b$	0.0
$-m$	$+b$	0.1
$-m$	$-b$	0.9

Figure 27: Image de [1].

8 Inférence dans les réseaux bayésiens et *d-separation*

Ex. 1

Une odeur de soufre (S) peut être causée par des œufs pourris (E), ou être un signe de malheur apporté par l'apocalypse Maya (M). L'apocalypse Maya fait également bouillir les océans (B). Le réseau Bayésien et les tables de probabilités conditionnelles correspondantes pour cette situation sont présentés dans la Figure 27.

1. Calculez la probabilité jointe $P(-e, -s, -m, -b)$. $P(-e, -s, -m, -b) = P(-e)P(-m)P(-s|-e, -m)P(-b|-m) = (0.6)(0.9)(0.9)(0.9) = 0.4374.$

2. Quelle est la probabilité que les océans bouillonnent ? $P(+b) = P(+b|+m)P(+m) + P(+b|-m)P(-m) = (1.0)(0.1) + (0.1)(0.9) = 0.19.$

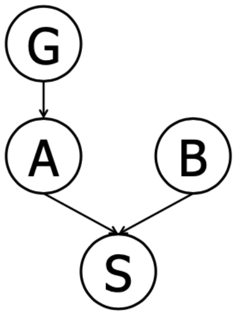
3. Quelle est la probabilité que l'apocalypse Maya se produise, étant donné que les océans sont en ébullition ? $P(+m|+b) = \frac{P(+b|+m)}{P(+m)} = \frac{(1.0)(0.1)}{0.19} \approx 0.5263$

4. Quelle est la probabilité que l'apocalypse Maya se produise, étant donné qu'il y a une odeur de soufre, que les océans sont en ébullition et qu'il y a des œufs pourris ?

$$\begin{aligned}
P(+m|+s,+b,+e) &= \frac{P(+m,+s,+b,+e)}{\sum_m P(m,+s,+b,+e)} \\
&= \frac{P(+e)P(+m)P(+s|+m,+e)P(+b|+m)}{\sum_m P(+e)P(m)P(+s|m,+e)P(+b|m)} \\
&= \frac{(0.4)(0.1)(1.0)(1.0)}{(0.4)(0.1)(1.0)(1.0) + (0.4)(0.9)(0.8)(0.1)} \\
&= \frac{0.04}{0.04 + 0.0288} \approx 0.5814
\end{aligned}$$

$\mathbf{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbf{P}(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$\mathbf{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbf{P}(S \mid A, B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

Figure 28: Image de [1].

5. Quelle est la probabilité que des œufs pourris soient présents, étant donné que l'apocalypse Maya se produit ? $P(+e|+m) = P(+e) = 0.4$ parce que E est indépendant de M.

Ex. 2

Supposons qu'un patient puisse avoir un symptôme (S) pouvant être causé par deux maladies différentes (A et B). On sait que la variation du gène G joue un rôle important dans l'apparition de la maladie A. Le réseau de Bayes et les tables de probabilités conditionnelles correspondantes pour cette situation sont dans la Figure 28.

1. Calculer la probabilité conjointe $P(+g, +a, +b, +s)$ $P(+g, +a, +b, +s) = P(+g)P(+a|+g)P(+b)P(+s|+b, +a) = (0.1)(1.0)(0.4)(1.0) = 0.04$

2. Quelle est la probabilité qu'un patient ait la maladie A? $P(+a) = P(+a|+g)P(+g) + P(+a|-g)P(-g) = (1.0)(0.1) + (0.1)(0.9) = 0.19$

3. Quelle est la probabilité qu'un patient ait la maladie A sachant qu'il a la maladie B ? $P(+a|+b) = P(+a) = 0.19$ parce qu' A et B sont absolument indépendants.

4. Quelle est la probabilité qu'un patient ait la maladie A sachant qu'il a le

symptôme S et la maladie B ?

$$\begin{aligned}
P(+a|+s,+b) &= \frac{P(+a,+s,+b)}{P(+a,+s,+b)+P(-a,+s,+b)} \\
&= \frac{P(+a)P(+b)P(+s|+a,+b)}{P(+a)P(+b)P(+s|+a,+b)+P(-a)P(+b)P(+s|-a,+b)} \\
&= \frac{(0.19)(0.4)(1.0)}{(0.19)(0.4)(1.0)+(0.81)(0.4)(0.8)} \\
&= \frac{0.076}{0.076+0.2592} \approx 0.2267
\end{aligned}$$

5. Quelle est la probabilité qu'un patient ait la variation du gène porteur de la maladie G étant donné qu'il a la maladie A ?

$$\begin{aligned}
P(+g|+a) &= \frac{P(+g)P(+a|+g)}{P(+g)P(+a|+g)+P(-g)P(+a|-g)} \quad = \frac{(0.1)(1.0)}{(0.1)(1.0)+(0.9)(0.1)} \\
&= \frac{0.1+}{0.1+0.09} \approx 0.5263
\end{aligned}$$

6. Quelle est la probabilité qu'un patient ait la variation du gène porteur de la maladie G étant donné qu'il a la maladie B ? $P(+g|+b) = P(+g) = 0.1$,parce que $G \perp\!\!\!\perp B$.

8.1 Examen 2022-2023

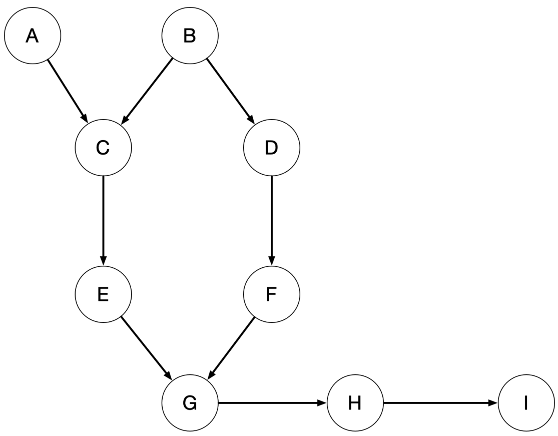


Figure 29: Réseaux bayésien composé de 9 nœuds. Les arcs modélisent les dépendances conditionnelles entre les variables.

1. Parmi les déclarations d'indépendance conditionnelles suivantes, lesquelles sont garanties vraies par la structure du réseau Bayésien dans la Figure 29 ? (3 points)

(a) $A \perp\!\!\!\perp D|C$ non

(b) $A \perp\!\!\!\perp D|B$ oui

(c) $A \perp\!\!\!\perp I$ oui

(d) $E \perp\!\!\!\perp I|H$ oui

(e) $D \perp\!\!\!\perp E|G$ non

2. Quelles sont les couvertures de Markov (*Markov Blanket*) des variables E et H ? (2 points) E :C, G, F et H : G, I

Ex. 3

PacLabs vient de créer un nouveau type de mini-pastille de puissance suffisamment petite pour que Pacman puisse l'emporter avec lui lorsqu'il parcourt des labyrinthes. Malheureusement, ces mini-pastilles ne garantissent pas que Pacman gagnera tous ses combats avec des fantômes, et ils ressemblent aux points réguliers que Pacman transporte pour grignoter.

Pacman vient de manger une collation (P), qui était soit une mini-pastille ($+p$), soit un point normal ($-p$), et est sur le point de se battre (W). Il peut resortir de son combat gagnant ($+w$) ou perdant ($-w$). Ces deux variables sont inconnues, mais heureusement, Pacman est un maître des probabilités. Il sait que son sac de collations contient 5 mini-pastilles et 15 pastilles régulières. Il sait aussi que s'il a mangé une mini-pastille, il a 70% de chance de gagner, mais s'il a mangé une pastille régulière, il n'a que 20% de chance.

1. Quelle est $P(+w)$, la probabilité marginale que Pacman gagne ?

$$\begin{aligned} P(+w) &= P(+w,+p) + P(+w,-p) = P(+w|+p)P(+p) + P(+w|-p)P(-p) \\ &= \frac{7}{10} \times \frac{1}{4} + \frac{2}{10} \times \frac{3}{4} = \frac{13}{40} = 0.325 \end{aligned}$$

2. Pacman a gagné ! Hourra ! Quelle est la probabilité conditionnelle $P(+p|+w)$ que l'aliment qu'il a mangé soit une mini-pastille, étant donné qu'il a gagné ?

$$\begin{aligned} P(+p|+w) &= \frac{P(+p,+w)}{P(+w)} = \frac{P(+w|+p)P(+p)}{P(+w)} \\ &= \frac{\frac{7}{10} \times \frac{1}{4}}{\frac{13}{40}} = \frac{7}{13} \approx 0.538 \end{aligned}$$

Pacman peut faire de meilleures estimations de probabilité s'il prend en compte plus d'informations. Premièrement, l'haleine de Pacman, B, peut être mauvaise ($+b$) ou fraîche ($-b$). Deuxièmement, il existe deux types de fantôme (M) : moyen ($+m$) et gentil ($-m$). Pacman a encodé ses connaissances sur la situation dans le réseau de Bayes dans la Figure 30.

3. Quelle est la probabilité de l'événement atomique $P(-m,+p,+w,-b)$, où Pacman mange une mini-pastille et a une haleine fraîche avant de gagner un combat contre un gentil fantôme ? $P(-m,+p,+w,-b) = P(-m)P(+p)P(+w|-m,+p)P(-b|+p) = \frac{1}{2} \times \frac{1}{4} \times \frac{4}{5} \times \frac{1}{5} = \frac{1}{50} = 0.02$.

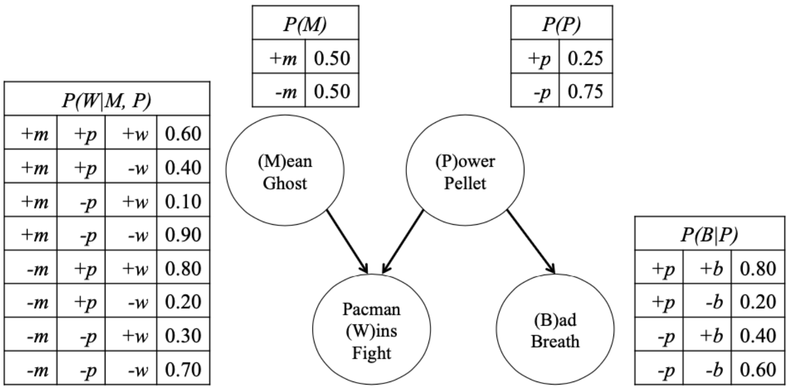


Figure 30: Image de [1].

$P(M, P, W, B)$				
$+m$	$+p$	$+w$	$+b$	0.0800
$+m$	$+p$	$+w$	$-b$	0.0150
$+m$	$+p$	$-w$	$+b$	0.0400
$+m$	$+p$	$-w$	$-b$	0.0100
$+m$	$-p$	$+w$	$+b$	0.0150
$+m$	$-p$	$+w$	$-b$	0.0225
$+m$	$-p$	$-w$	$+b$	0.1350
$+m$	$-p$	$-w$	$-b$	0.2025

Figure 31: Image de [1].

4. Parmi les déclarations d'indépendance conditionnelle suivantes, lesquelles sont garanties vraies par la structure du réseau de Bayes ?

   (a) $W \perp\!\!\!\perp B$

   (b) $W \perp\!\!\!\perp B | P$

   (c) $M \perp\!\!\!\perp P$

   (d) $M \perp\!\!\!\perp P | W$

   (e) $M \perp\!\!\!\perp B$

   (f) $M \perp\!\!\!\perp B | P$

   (g) $M \perp\!\!\!\perp B | W$

   Les déclarations d'indépendance conditionnelle (b), (c), (e) et (f) sont correct.

Pour le reste de cette question, utilisez la moitié de la table de probabilité conjointe qui a été calculée pour vous dans la Figure 31.

5. Quelle est la probabilité marginale, $P(+m,+b)$ que Pacman rencontre un fantôme méchant et ait mauvaise haleine ? $P(+m,+b) = 0.08 + 0.04 + 0.015 + 0.135 = 0.27$.

6. Pacman constate qu'il a mauvaise haleine et que le fantôme auquel il est confronté est méchant. Quelle est la probabilité conditionnelle, $P(+w|+m,+b)$, qu'il gagne le combat, compte tenu de son observation ?

$$\begin{aligned} P(+w|+m,+b) &= \frac{P(+w,+m,+b)}{P(+m,+b)} \\ &= \frac{0.08+0.015}{0.27} = \frac{19}{54} \approx 0.352 \end{aligned}$$

7. L'utilité de Pacman est de $+10$ pour un combat gagné, -5 pour un combat perdu et -1 s'il court loin d'un combat. Pacman veut maximiser son utilité espérée. Étant donné qu'il a mauvaise haleine et qu'il fait face à un méchant fantôme, doit-il rester et se battre, ou s'enfuir ? Justifiez votre réponse numériquement ! Soit U_f l'utilité du combat et U_r l'utilité de la fuite.

$$\begin{aligned} E(U_f|+m,+b) &= 10 \times P(w+|+m,+b) + (-5) \times P(-w|+m,+b) \\ &\approx 10 \times 0.352 - 5 \times 0.648 \\ &= 0.28 > -1 = U_r \end{aligned}$$

Puisque $E(U_f|+m,+b) > E(U_r|+m,+b)$, Pacman doit rester et se battre.

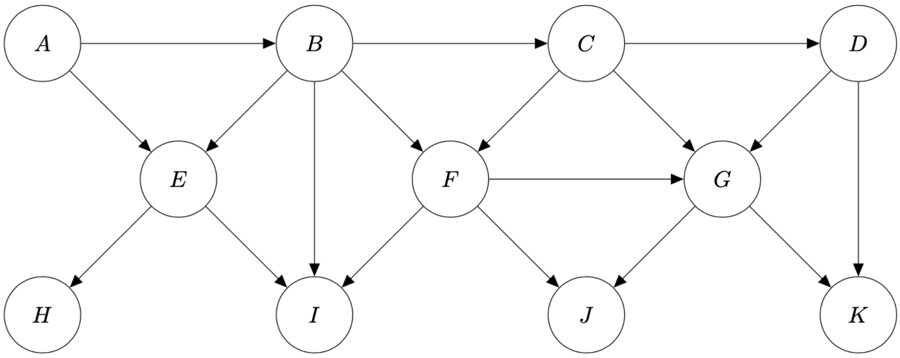
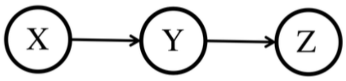


Figure 32: Image de [1].



X	$\mathbf{P}(X)$
$+x$	$2/5$
$-x$	$3/5$

Y	X	$\mathbf{P}(Y \mid X)$
$+y$	$+x$	2/3
$-y$	$+x$	1/3
$+y$	$-x$	3/4
$-y$	$-x$	1/4

Z	Y	$\mathbf{P}(Z \mid Y)$
$+z$	$+y$	1/3
$-z$	$+y$	2/3
$+z$	$-y$	1/5
$-z$	$-y$	4/5

Figure 33: Image de [1].

9 Échantillonnage et réseaux bayésiens

Ex. 1

Identifier la couverture de Markov (*Markov Blanket*) de chaque variable dans la Figure 32. La couverture de Markov sont les parents, les enfants et les parents des enfants d'une variable.

A: B,E; **B**: A,E,F,C; **C**: B,F,G,D; **D**: C,G,K; **E**: A,B,H,I; **F**: B,C,G,I,J; **G**: C,D,F,J,K; **H**: E; **I**: B,E,F; **J**: F,G; **K**: D,G

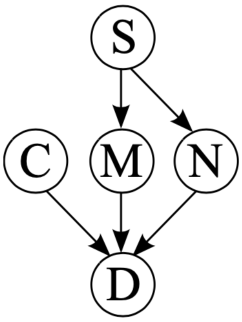
Ex. 2

Supposons le réseau Bayésien et les distributions correspondantes sur les variables du réseau de la Figure 33.

1. Votre tâche consiste maintenant à estimer $P(+y|+x,+z)$ en utilisant le *rejection sampling*. Vous trouverez ci-dessous quelques échantillons qui ont été produits par *prior sampling* (c'est-à-dire que l'étape de rejet dans l'échantillonnage n'a pas encore eu lieu). Lequel des échantillons suivants serait rejeté par le *rejection sampling*?

   (a) $+x$, $+y$, $+z$

   (b) $-x$, $+y$, $+z$



S	C	M	N	D
$-s$	$+c$	$+m$	$-n$	$+d$
$+s$	$+c$	$-m$	$-n$	$-d$
$+s$	$-c$	$-m$	$+n$	$-d$
$+s$	$+c$	$+m$	$-n$	$+d$
$+s$	$+c$	$-m$	$+n$	$+d$
$+s$	$-c$	$-m$	$+n$	$-d$
$+s$	$-c$	$-m$	$-n$	$-d$
$+s$	$+c$	$+m$	$+n$	$+d$
$+s$	$+c$	$-m$	$+n$	$-d$
$-s$	$-c$	$-m$	$-n$	$-d$

S	C	M	N	D	weight
$-s$	$+c$	$+m$	$+n$	$-d$	
$-s$	$+c$	$-m$	$-n$	$-d$	
$-s$	$-c$	$-m$	$+n$	$-d$	
$-s$	$+c$	$+m$	$+n$	$-d$	

Figure 34: Image de [1].

(c) $-x$, $-y$, $+z$

(d) $+x$, $-y$, $-z$

(e) $+x$, $-y$, $+z$

Les échantillons (b), (c), et (d) seront refusés.

2. À l'aide d'un *rejection sampling*, donnez une estimation de $P(+y|+x, +z)$ à partir des cinq échantillons précédents, ou indiquez pourquoi il ne peut pas être calculé.

   Sur les deux échantillons restants, l'un a $+y$, donc l'estimation est 1/2.

3. En utilisant les échantillons suivants (qui ont été générés en utilisant le *Likelihood weighting*), estimez $P(+y| + x, +z)$ en utilisant le *likelihood weighting*, ou indiquez pourquoi il ne peut pas être calculé.

   (a) $+x$, $+y$, $+z$

   (b) $+x$, $+y$, $+z$

   (c) $+x$, $+y$, $+z$

   Les trois poids sont les suivants : $w_1 = w_3 = P(+x)P(+z| + y) = 2/5 \times 1/3 = 2/15$ et $w_2 = P(+x)P(+z| - y) = 2/5 \times 1/5 = 2/25$. Donc $P(+y| + x, +z) \approx (2w_1)/(2w_1 + w_2) = 10/13$.

Ex. 3

Vous êtes un exobiologiste, étudiant le large éventail de la vie dans l'univers. Vous êtes également un danseur passionné et avez un excellent modèle de la façon dont les espèces inventent la danse. Les variables clés sont :

- Détection du son (S) : si une espèce a ou non la capacité de détecter le son
- Climat froid (C) : si la planète natale de l'espèce connaît ou non un climat froid
- Musique (M) : Si oui ou non l'espèce a inventé la musique

C	M	N	D	P(D | C , M , N)
$+c$	$+m$	$+n$	$+d$	0.9
$+c$	$+m$	$+n$	$-d$	0.1
$+c$	$+m$	$-n$	$+d$	0.8
$+c$	$+m$	$-n$	$-d$	0.2
$+c$	$-m$	$+n$	$+d$	0.8
$+c$	$-m$	$+n$	$-d$	0.2
$+c$	$-m$	$-n$	$+d$	0.2
$+c$	$-m$	$-n$	$-d$	0.8
$-c$	$+m$	$+n$	$+d$	0.8
$-c$	$+m$	$+n$	$-d$	0.2
$-c$	$+m$	$-n$	$+d$	0.5
$-c$	$+m$	$-n$	$-d$	0.5
$-c$	$-m$	$+n$	$+d$	0.6
$-c$	$-m$	$+n$	$-d$	0.4
$-c$	$-m$	$-n$	$+d$	0.1
$-c$	$-m$	$-n$	$-d$	0.9

S	M	P(M | S)
$+s$	$+m$	0.8
$+s$	$-m$	0.2
$-s$	$+m$	0.1
$-s$	$-m$	0.9

S	P(S)
$+s$	0.9
$-s$	0.1

S	N	P(N | S)
$+s$	$+n$	0.7
$+s$	$-n$	0.3
$-s$	$+n$	0.9
$-s$	$-n$	0.1

C	P(C)
$+c$	0.5
$-c$	0.5

Figure 35: Image de [1].

- Communication non verbale (N) : si l'espèce a ou non une forme quelconque de communication non verbale.

Vous modélisez les relations entre ces variables et la danse (D) à l'aide du réseau bayésien spécifié dans la Figure 34.

Vous voulez savoir quelle est la probabilité qu'une espèce dansante et sensible au son de la musique, selon le réseau Bayésien dans la Figure 34. Vous décidez de faire de l'inférence par échantillonnage. Vous utilisez "prior sampling" pour tirer les échantillons (voir le tableau dans la Figure 34).

1. Sur base du *rejection sampling* utilisant les échantillons ci-dessus, quelle est la réponse à votre question $P(+m|+d,+s)$?

   L'estimation de la probabilité est de 1/3. Trouvez simplement toutes les lignes du tableau ci-dessus avec $+d$ et $+s$, et comptez le nombre de fois où $-m$ se produit divisé par le nombre total de ces lignes.

Bien que votre méthode d'échantillonnage ait assez bien fonctionné dans de nombreux cas, dans de rares cas (comme les espèces qui ne peuvent pas détecter le son), vos résultats sont moins précis car le *rejection sampling* rejette presque toutes les données. Vous décidez d'utiliser *Likelihood weighting* à la place. Les probabilités conditionnelles du réseau de Bayes sont répertoriées dans le tableau dans la Figure 35.

Vous souhaitez maintenant calculer la probabilité qu'une espèce qui n'a pas de détection sonore ($-s$) ou de danse ($-d$) ait néanmoins de la musique ($+m$), en utilisant cette information. C'est-à-dire que vous voulez calculer $P(+m|-s,-d)$.

2. Vous tirez les échantillons du tableau, en utilisant le *likelihood weighting*. Pour chacun de ces échantillons, indiquez son poids

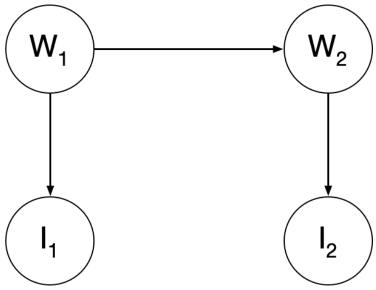


Figure 36: Réseaux Bayésien composé de 4 nœuds. Les arêtes modélisent les dépendances conditionnelles entre les variables.

S	C	M	N	D	weight
-s	+c	+m	+n	-d	0.01
-s	+c	-m	-n	-d	0.08
-s	-c	-m	+n	-d	0.04
-s	+c	+m	+n	-d	0.01

3. Calculez la réponse à votre requête, $P(+m|-s,-d)$, en utilisant la pondération de vraisemblance avec ces échantillons.

   $P(+m|-s,-d) \approx 1/7$; diviser la somme des poids correspondant aux lignes $+m$ par la somme totale des poids.

9.1 Examen 2022-2023

Nous aimerions analyser la consommation de crème glacée les jours ensoleillés et pluvieux. Supposons que nous considérions la météo, ainsi que la consommation de glace d'une personne, sur une période de deux jours. Nous aurons quatre variables aléatoires : W_1 et W_2 représentent le temps qu'il fait les jours 1 et 2, qui peuvent être pluvieux R ou ensoleillés S, et les variables I_1 et I_2 représentent si la personne a mangé ou non de la glace les jours 1 et 2, et prennent les valeurs T (si elle a mangé de la crème glacée) ou F. Nous pouvons modéliser cela comme le réseau Bayésien de la Figure 36. Les variables sont associées aux probabilités suivantes :

W_1	$P(W_1)$
S	0.6
R	0.4

W_1	W_2	$P(W_2\|W_1)$
S	S	0.7
S	R	0.3
R	S	0.5
R	R	0.5

W	I	$P(I\|W)$
S	T	0.9
S	F	0.1
R	T	0.2
R	F	0.8

1. Supposons que nous produisions des échantillons de (W_1, I_1, W_2, I_2) à partir du modèle de crème glacée. Ceux-ci sont donnés dans la table 1

   (a) En utilisant ces échantillons, quelle est l'estimation de $P(W_2 = R)$? (1 point) $5/10 = 0.5$

W_1	I_1	W_2	I_2
R	F	R	F
R	F	R	F
S	F	S	T
S	T	S	T
S	T	R	F
R	F	R	T
S	T	S	T
S	T	S	T
S	T	R	F
R	F	S	T

Table 1: Échantillons

(b) Quels échantillons ci-dessus sont rejetés par *rejection sampling* si nous essayons d'estimer $P(W_2|I_1 = T, I_2 = F)$? (1 point) (R,F,R,F), (R,F,R,F), (S,F,S,T), (S,T,S,T), (R,F,R,T), (S,T,S,T) , (S,T,S,T), (R,F,S,T). Donc tous les échantillons sauf les deux (S,T,R,F).

2. Le *rejection sampling* semble peu efficace, nous décidons donc de passer au *likelihood weighting*. Supposons qu'on génère les six échantillons suivants pour (W_1, I_1, W_2, I_2), étant donné la preuve $I_1 = T$ et $I_2 = F$: $(S, T, R, F), (R, T, R, F), (S, T, R, F), (S, T, S, F), (S, T, S, F), (R, T, S, F)$

(a) Calculez le poids de chaque échantillon. (2 points) Dans ce cas, la preuve est $I_1 = T$, $I_2 = F$. Le poids du premier échantillon est donc $w = P(I_1 = T|W_1 = S) \times P(I_2 = F|W_2 = R) = 0.9 \times 0.8 = 0.72$. De même, on peut calculer les poids des autres échantillons comme suit:

(W_1, I_1, W_2, I_2)	w
S,T,R,F	0.72
R,T,R,F	0.16
S,T,R,F	0.72
S,T,S,F	0.09
S,T,S,F	0.09
R,T,S,F	0.02

(b) Estimez $P(W_2|I_1 = T, I_2 = F)$ en utilisant les poids de vraisemblance (*likelihood weighting*) de la partie précédente. (2 points) Pour calculer l'estimation de probabilité, nous normalisons les poids et trouvons

$$P(W_2 = R|I_1 = T, I_2 = F) = \frac{0.72 + 0.16 + 0.72}{0.72 + 0.16 + 0.72 + 0.09 + 0.09 + 0.02} = 0.889$$

et

$$P(W_2 = S|I_1 = T, I_2 = F) = 1 - 0.889 = 0.111$$

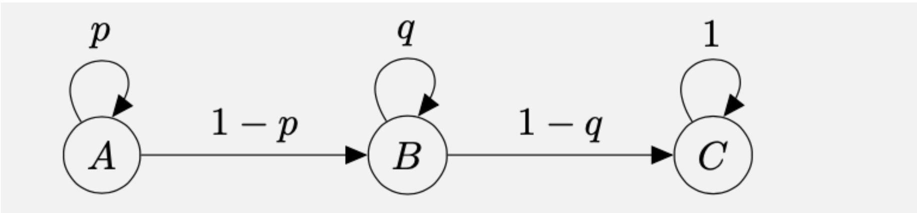


Figure 37: Image de [1].

10 Modèles de Markov

Ex. 1

Supposons qu'un objet se déplace selon le modèle de transition de la Figure 37. Ici, $0 < p < 1$ et $0 < q < 1$ sont des probabilités arbitraires. Au temps 0, l'on sait que l'objet est dans l'état A.

Nous définissons $\beta(n)$ et $\gamma(n)$ comme la probabilité que l'objet atteigne d'abord B et C, respectivement au temps n. On définit $a(n)$, $b(n)$, $c(n)$ comme la probabilité que l'objet soit dans A, B, C, respectivement à l'instant n.

1. Quelle est la probabilité que l'objet soit dans A au temps $n \geq 0$? Pour que l'objet soit dans A à l'instant n, il doit être resté dans A pendant n étapes, ce qui se produit avec probabilité $a(n) \equiv p^n$.

2. Quelle est la probabilité que l'objet atteigne B pour la première fois au temps $n \geq 1$? Pour que l'objet atteigne B pour la première fois au temps n, il doit être resté en A pendant $n-1$ étapes, puis être passé à B. Cela se produit avec la probabilité $\beta(n) \equiv a(n-1) \cdot (1-p) = p^{n-1}(1-p)$.

3. Quelle est la probabilité que l'objet soit dans B au temps $n \geq 1$? Pour que l'objet soit dans B à l'instant n, il doit d'abord avoir atteint B à l'instant i tel que $1 \leq i \leq n$, puis y être resté pendant $n-i$ étapes. La sommation sur toutes les valeurs de i donne

$$\begin{aligned} b(n) &= \sum_{i=1}^{n} \beta(i) \cdot q^{n-1} \\ &= \sum_{i=1}^{n} p^{i-1}(1-p)q^{n-1} \\ &= \frac{(1-p)q^n}{p} \sum_{i=1}^{n} \left(\frac{p}{q}\right)^i \\ &= \frac{(1-p)q^n}{p} \cdot \frac{p}{q} \cdot \frac{1-\left(\frac{p}{q}\right)^n}{1-\frac{p}{q}} \\ &= (1-p)\frac{q^n - p^n}{q-p} \end{aligned}$$

   où les simplifications des deux dernières lignes supposent $p \neq q$.

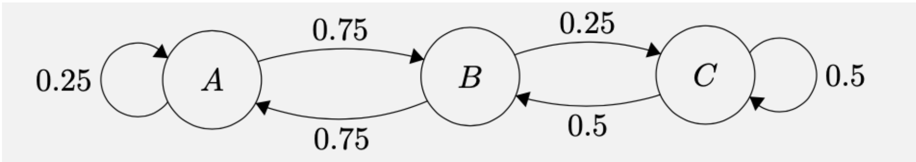


Figure 38: Image de [1].

4. Quelle est la probabilité que l'objet atteigne C pour la première fois au temps $n \geq 2$? Pour que l'objet atteigne pour la première fois C au temps n, il doit avoir été dans B au temps $n-1$, puis être passé à C. Cela se produit avec probabilité :

$$\gamma(n) = b(n-1)(1-q) = (1-p)\frac{q^{n-1} - p^{n-1}}{q-p}(1-q)$$

5. Quelle est la probabilité que l'objet soit dans C au temps $n \geq 2$?

   Pour que l'objet soit dans C à l'instant n, il ne doit pas être dans A ou B à l'instant n. Cela se produit avec probabilité :

$$1 - a(n) - b(n) = 1 - p^n - (1-p)\frac{q^n - p^n}{q-p}$$

Ex. 2

Considérons une chaîne de Markov avec 3 états et probabilités de transition comme dans la Figure 38. Calculer la distribution stationnaire. Autrement dit, calculez $P_\infty(A)$, $P_\infty(B)$ et $P_\infty(C)$.
$P_\infty(A) = 0.4, P_\infty(B) = 0.4, P_\infty(C) = 0.2$

Ex. 3

Considérons un HMM avec des variables d'état $\{X_i\}$ et des variables d'émission $\{Y_i\}$.

1. (vrai ou faux) X_i est toujours conditionnellement indépendant de Y_{i+1} étant donné X_{i+1}. vrai

2. (vrai ou faux) Il existe un HMM où X_i est conditionnellement indépendant de Y_i étant donné $X_i + 1$. vrai

3. (vrai ou faux) Si $Y_i = X_i$ avec probabilité 1, et que l'espace d'état est de taille k, alors l'algorithme le plus efficace pour calculer $P(X_t|y_1, \cdots, y_t)$ prend $O(k)$ ou moins de temps. vrai

Ex. 4

1. (vrai ou faux) Avec un modèle de transition déterministe et un modèle d'observation stochastique, comme le temps va vers l'infini, lors de l'exécution d'un filtre à particules, nous nous retrouverons avec des particules toutes identiques. vrai

2. (vrai ou faux) Avec un modèle d'observation déterministe, toutes les particules pourraient finir par avoir un poids nul. vrai

3. (vrai ou faux) Il est possible d'utiliser le filtrage à particules lorsque l'espace d'état est continu. vrai

4. (vrai ou faux) Le nombre de particules tendant vers l'infini, le filtrage à particules représentera la même distribution de probabilité que vous obtiendriez en utilisant l'inférence exacte. vrai

5. (vrai ou faux) Le filtrage à particules peut représenter une distribution plate (c'est-à-dire uniforme) avec moins de particules qu'il n'en faudrait pour une distribution plus concentrée (c'est-à-dire gaussienne). faux

Ex. 5

Pour quels paramètres le filtrage à particules est-il meilleur que l'inférence HMM exacte?

1. Espaces d'états grands ou petits.

2. Priorité à la durée d'exécution par rapport à la précision.

Le filtrage à particules fonctionne mieux dans les grands espaces d'état où nous souhaitons donner la priorité à la durée d'exécution (*runtime*) plutôt qu'à la précision. L'inférence HMM exacte évolue avec l'espace d'état. Pour le filtrage particulaire, le nombre de particules détermine la complexité temporelle et la qualité de l'approximation. De grands espaces d'états (même continus) peuvent être approchés par un nombre beaucoup plus petit de particules.

Ex. 6

Considérons un HMM avec T pas de temps, variables d'état cachées $X_1, \cdots, X_T$, et variables observées $E_1, \cdots, E_T$. Soit S le nombre d'états possibles pour chaque variable d'état cachée X.

On veut calculer (avec l'algorithme direct) ou estimer (avec filtrage à particules) $P(X_T|E_1 = e_1, \cdots, E_T = e_T)$

Combien de particules, en termes de S et/ou de T, faudrait-il pour que le filtrage particulaire ait la même complexité temporelle que l'algorithme direct ? Vous pouvez supposer que, dans le filtrage à particules, chaque étape d'échantillonnage peut être effectuée en temps constant pour une seule particule (bien que ce ne soit pas nécessairement le cas dans la réalité).

S^2

11 Réseaux de décisions

Ex. 1

Considérez un étudiant qui a le choix d'acheter ou de ne pas acheter un manuel pour un cours.

Modélisez cela comme un problème de décision avec un nœud de décision booléen, B, indiquant si l'agent choisit d'acheter le livre, et deux nœuds de hasard booléens, M, indiquant si l'étudiant a maîtrisé le contenu du livre, et P, indiquant si l'élève réussit le cours. Bien sûr, il existe également un nœud utilitaire, U.

Un certain étudiant, Sam, a une fonction d'utilité additive : 0 si l'achat n'est pas fait et -100\$ si l'achat est fait ; et 2000\$ s'il réussi le cours et 0 s'il ne réussi pas le cours. Les estimations de probabilités conditionnelles de Sam sont les suivantes : $P(+p|+b,+m) = 0.9$; $P(+m|+b) = 0.9$; $P(+p|+b,-m) = 0.5$; $P(m|-b) = 0.7$; $P(+p|-b,+m) = 0.8$; $P(p|-b,-m) = 0.3$ Vous pourriez penser que P serait indépendant de B étant donné M, mais ce cours a un examen à livre ouvert, donc avoir le livre aide.

1. Dessinez le réseau de décision pour ce problème. Regardez la Figure 39.

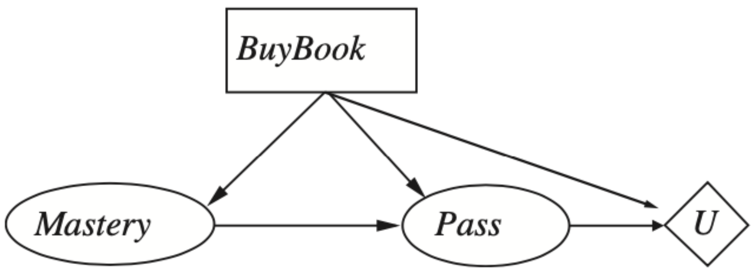


Figure 39: Image de [1].

2. Calculez l'utilité espérée d'acheter le livre et de ne pas l'acheter.

   Pour chacun de $B = +b$ et $B = -b$, nous calculons $P(+p|B)$ et donc $P(-p|B)$ en marginalisant M , puis utilisons ceci pour calculer l'utilité espérée.

$$\begin{aligned} P(+p|+b) &= \sum_m P(+p|+b,m)P(m|+b) \\ &= 0.9 \times 0.9 + 0.5 \times 0.1 \\ &= 0.86 \end{aligned}$$

$$\begin{aligned} P(+p|-b) &= \sum_m P(+p|-b,m)P(m|-b) \\ &= 0.8 \times 0.7 + 0.3 \times 0.3 \\ &= 0.65 \end{aligned}$$

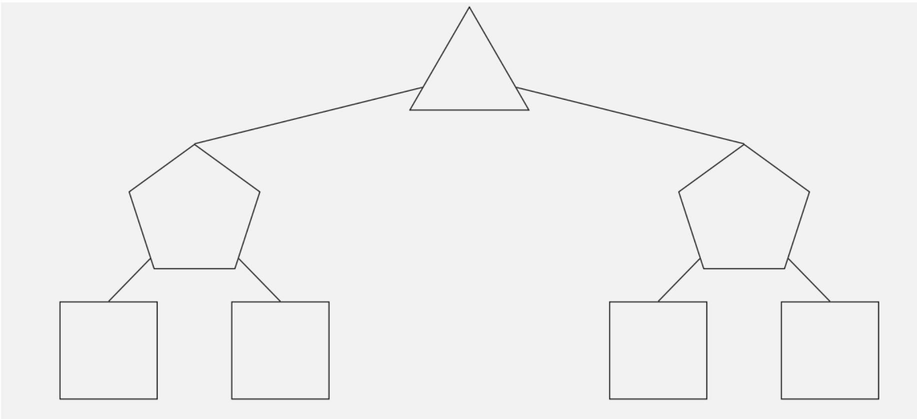


Figure 40: Image de [1].

Les utilités attendues sont :

$$\begin{aligned} EU[+b] &= \sum_p P(p|+b)U(p,+b) \\ &= 0.86(2000-100) + 0.14(-100) \\ &= 1620 \end{aligned}$$

$$\begin{aligned} EU[-b] &= \sum_p P(p|-b)U(p,-b) \\ &= 0.65 \times 2000 + 0.14 \times 0 \\ &= 1300 \end{aligned}$$

3. Que doit faire Sam ? Il faut acheter le livre.

Ex. 2

Valérie vient de trouver un biscuit par terre. Elle craint que le biscuit contienne des raisins secs, ce qu'elle n'aime vraiment pas, mais elle veut quand même manger le biscuit. Si elle mange le biscuit et qu'il contient des raisins secs, elle recevra une utilité de -100 et si le biscuit ne contient pas de raisins secs, elle recevra une utilité de 10. Si elle ne mange pas le biscuit, elle obtiendra 0 utilité. Le biscuit contient des raisins secs avec une probabilité de 0.1.

1. Nous voulons représenter ce réseau de décision comme un arbre de jeu *expectimax*. Remplissez les nœuds de l'arbre donné (Figure 40), le nœud supérieur représentant son choix de maximisation.

   Regardez la Figure 41.

2. Valérie devrait-elle manger le biscuit ? Non

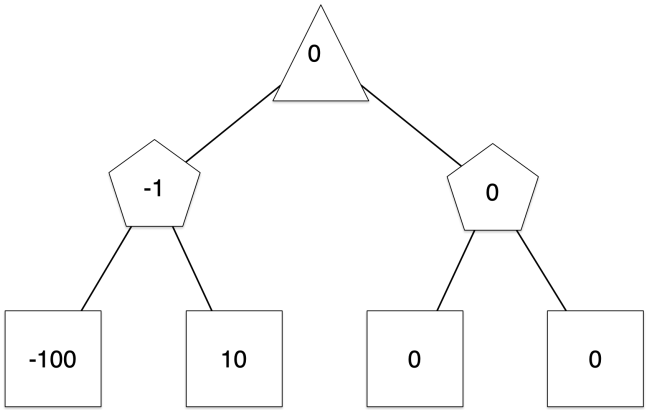


Figure 41: Image de [1].

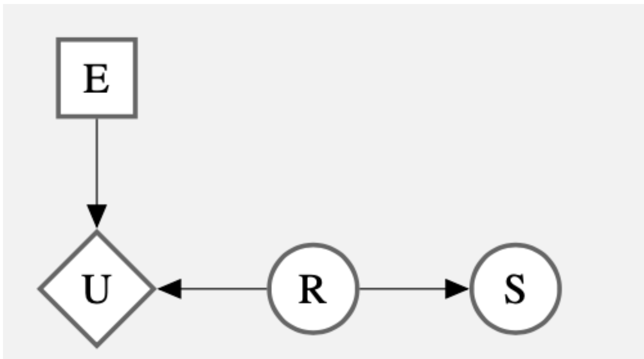


Figure 42: Image de [1].

Valérie peut maintenant sentir le biscuit pour juger s'il contient des raisins secs avant de le manger. Cependant, comme elle n'aime pas les raisins secs, elle n'a pas beaucoup d'expérience avec eux et ne peut pas bien reconnaître leur odeur. En conséquence, elle identifiera incorrectement les raisins secs lorsqu'il n'y a pas de raisins secs avec une probabilité de 0.2 et identifiera incorrectement l'absence de raisins secs lorsqu'il y a des raisins secs avec une probabilité de 0.3. Ce réseau de décision peut être représenté par le diagramme en Figure 42 où E est son choix de manger, U est son utilité gagnée, R est la présence de raisins secs dans le biscuit, et S est sa tentative de sentir.

3. Valérie vient de sentir le biscuit et elle pense qu'il n'y a pas de raisins secs. Donnez la probabilité, X, que le biscuit ait des raisins secs étant donné qu'elle n'a pas senti de raisins secs sous la forme la plus simple d'une fraction ou d'un nombre décimal.

$$P(+r|-s) = \frac{P(-s|+r)P(+r)}{P(-s)}$$
$$= \frac{P(-s|+r)P(+r)}{P(-s|r)P(+r)+P(-s|-r)P(-r)}$$
$$= \frac{(0.3)(0.1)}{(0.3)(0.1)+(0.8)(0.9)} = 0.04$$

4. Quelle est son utilité maximale attendue, Y, étant donné qu'elle n'a pas senti de raisins secs ? Vous pouvez répondre en terme de X ou sous la forme la plus simple d'une fraction ou d'un nombre décimal.

$$MEU(-s) = max(MEU(eating|-s), MEU(noteating|-s))$$
$$= \max(P(+r|-s) \times EU(eating, +r) + P(-r|-s) \times EU(eating, -r), MEU(noteating))$$
$$= max(X(-100) + (1-X)10, 0)$$
$$= 100X + 10(1-X)$$

5. Quelle est la valeur de l'information parfaite (*Value of Perfect Inforamtion*) de sentir le cookie ? Vous pouvez répondre en termes de X et Y ou sous la forme la plus simple d'une fraction ou d'un nombre décimal

- $VPI(S) = MEU(S) - MEU(\{\})$ avec $MEU(S) = P(-s)MEU(-s) + P(+s)MEU(+s)$.
- $P(-s) = .75$ de la question (3).
- $MEU(-s) = Y$
- $MEU(+s) = 0$ parce qu'il était préférable pour elle de ne pas manger le raisin sec sans rien savoir, sentir les raisins secs ne fera que rendre plus probable que le biscuit ait des raisins secs et il sera toujours préférable pour elle de ne pas manger et de gagner une utilité de 0. Notez que cela signifie que nous n'avons pas à calculer $P(+s)$.
- $MEU(\{\}) = 0$

Donc :$VPI(S) = 0.75Y + 0 - 0 = 0.75Y$

11.1 Examen 2022-2023

Pacman erre dans un labyrinthe et rencontre un coffre magique. Il aimerait ouvrir le coffre mais il craint qu'il contienne un fantôme (G). S'il l'ouvre et le fantôme apparaît, il sera tué par le fantôme, ce qui se traduira par un score de -10. Si le coffre contient en revanche une pastille dorée, Pacman obtiendra un score de 50. La probabilité d'obtenir la pastille dorée est de 20%.

- Dessinez le réseau de décision et calculez l'utilité espérée (*expected utility - EU*) de l'ouverture du coffre, c.-à-d. $EU(+open)$? (1 point) Le dessin est le figure 43 mais sans le noeud P. Pour le calcul $EU(+open) = P(+g)U(+g, +open) + P(-g)U(-g, +open) = 0.8 \times (-10) + 0.2 \times 50 = 2$

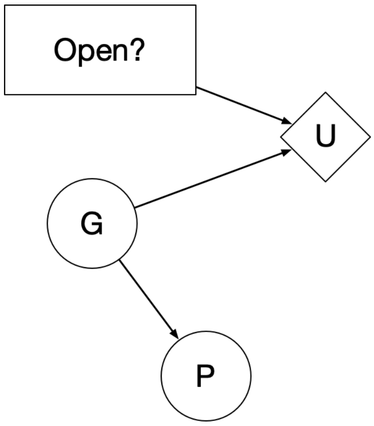


Figure 43: Réseau de décision

- Pacman peut demander l'aide d'un prêtre (P), qui peut vérifier si le coffre contient un fantôme. Le succès du prêtre ($P = +p$) dépendra du fait que le coffre magique contienne ou non un fantôme. Nous savons que $P(+p|+g) = 0.7$ et $P(+p|-g) = 0.2$. Calculez la probabilité que le prêtre réussisse son évaluation (c.-à-d. $P(+p)$ et $P(-p)$) et la probabilité que le fantôme (ou la pastille dorée) apparaisse du coffre à la suite de chaque résultat de l'évaluation du prêtre (c.-à-d. $P(+g|+p)$ et $P(+g|-p)$). Développez également la conception du réseau de décision de (a) avec cet élément supplémentaire. (2 points) Vous devez dessiner la figure 3 et calculer les valeurs des différentes probabilités.

$$
\begin{aligned}
P(+p) &= \sum_g P(+p, g) \\
&= P(+p|+g)P(+g) + P(+p|-g)P(-g) \\
&= 0.7 \times 0.8 + 0.2 \times 0.2 = 0.6 \\
P(-p) &= 1 - 0.6 = 0.4 \\
P(+g|+p) &= \frac{P(+p|+g)P(+g)}{P(+p)} \\
&= \frac{0.7 \times 0.8}{0.6} = 0.933... \approx 0.93 \\
P(+g|-p) &= \frac{P(-p|+g)P(+g)}{P(-p)} \\
&= \frac{0.3 \times 0.8}{0.4} = 0.6
\end{aligned}
$$

- Calculez les décisions optimales (c.-à-d. $+open$ ou $-open$) compte tenu du succès ($P = +p$) ou de l'échec ($P = -p$) du prêtre, et leurs utilités attendues maximale (MEU - *maximal expected utility*). (2 points)

$$\begin{aligned}
EU(+open|+p) &= P(+g|+p)U(+g,+open) + P(-g|+p)U(-g,+open) \\
&= 0.93 \times (-10) + 0.07 \times 50 = -5.8 \\
EU(+open|-p) &= P(+g|-p)U(+g,+open) + P(-g|-p)U(-g,+open) \\
&= 0.6 \times (-10) + 0.4 \times 50 = 14 \\
EU(-open|+p) &= 0 \\
EU(-open|-p) &= 0
\end{aligned}$$

Donc : $MEU(+p) = 0$ (en cas de $-open$) et $MEU(-p) = 14$ (en cas de $+open$).

- Calculez la valeur de l'information (parfaite) du test. Pacman doit-il demander au prêtre de faire le test ? (1 point)

$$VPI(P) = \left(\sum_p P(p)MEU(p)\right) - MEU(\emptyset) = (0.6 \times 0 + 0.4 \times 14) - 2 = 3.6$$

Pacman doit donc demander au prêtre de faire le test.

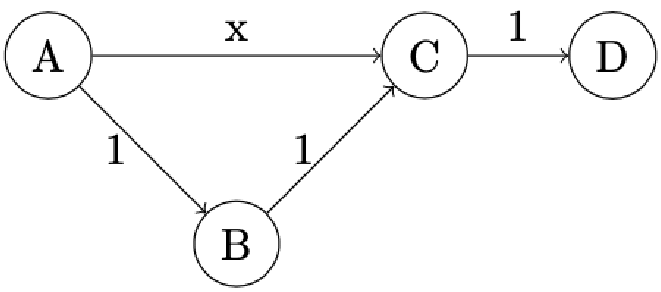


Figure 44: Image de [2].

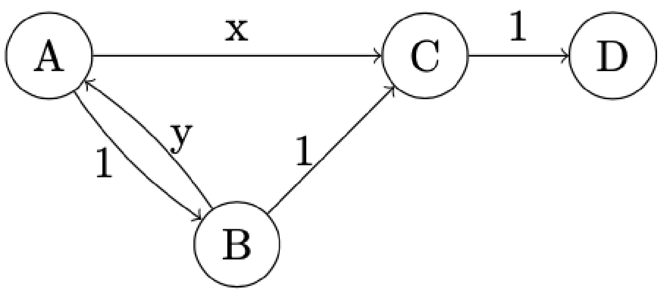


Figure 45: Image de [2].

12 Processus décisionnels de Markov

Ex. 1

Considérons le MDP déterministe dans la Figure 44 avec quatre états A, B, C et D. Les arcs désignent les actions entre les états, les poids sur ces arcs sont les récompenses et le facteur d'actualisation (discount factor) est $\gamma = 1$. Soit k la première itération lors de l'itération de valeur à laquelle la fonction de valeur converge pour certains x pour un état particulier (c'est-à-dire $V^k(s) = V^*(s)$). Pour chaque état A, B, C et D, énumérez toutes les valeurs possibles de k. Dans le cas où une fonction de valeur pour un état particulier ne converge jamais, définissez $k = \infty$ pour cet état.

Pour A $k = 2$ ou $k = 3$, pour B $k = 2$, pour C $k = 1$, et pour D $k = 0$. C aura sa valeur optimale en une itération. B aura sa valeur optimale une itération après cela (2 itérations au total). A aura sa valeur optimale une itération après C (2 itérations au total) ou une itération après B (3 itérations au total), selon la valeur de x.

Ex. 2

Considérons le MDP déterministe dans la Figure 45 avec quatre états A, B, C et D. Les arêtes désignent les actions entre les états, les poids sur ces arêtes sont les récompenses et le facteur d'actualisation est à nouveau $\gamma = 1$. Supposons en outre que $x, y \geq 0$.

Soit k la **première** itération de l'itération de valeur pour certains x et y non négatifs auxquels la fonction de valeur converge pour un état particulier ($V_k(s) = V^*(s)$). Pour chaque état A, B, C et D énumérez **toutes** les valeurs possibles de k. Dans le cas où une valeur pour un état particulier ne converge jamais, définissez $k = \infty$ pour cet état.

Pour A $k = \infty$, pour B $k = \infty$, pour C $k = 1$ et pour D $k = 0$. A et B n'auront jamais leurs valeurs optimales car ils peuvent prendre une valeur infinie. C et D sont les mêmes que dans Ex. 1.

Ex. 3

Considérons que nous effectuons une itération de politique (*policy iteration*) et que k est la **première** itération pour laquelle la politique est optimale pour un état particulier (c'est-à-dire $\pi_k(s) = \pi^*(s)$). En plus de $x, y \geq 0$ supposent également que $x+y < 1$ et que le départage lors de l'amélioration de la politique est alphabétique. Pour chaque état A, B, C et D, trouvez k ; si la politique ne converge jamais, fixez $k = \infty$ pour cet état. La politique initiale est donnée dans le tableau ci-dessous.

État s	Politique $\pi_0(s)$
A	C
B	C
C	D
D	S

pour A $k = 1$, pour B $k = 2$ et pour C et D $k = 0$. D'abord il faut faire l'évaluation de π_0:

$$\begin{aligned}
&V^{\pi_0}(D) = 0 \\
&\Rightarrow V^{\pi_0}(C) = 1 + V^{\pi_0}(D) = 1 \\
&\Rightarrow V^{\pi_0}(B) = 1 + V^{\pi_0}(C) = 2 \\
&\Rightarrow V^{\pi_0}(A) = x + V^{\pi_0}(C) = x + 1
\end{aligned}$$

Effectuez maintenant une amélioration de la politique pour obtenir π_1 :

$$\begin{aligned}
&\pi_1(D) = D \\
&\pi_1(C) = D \\
&\pi_1(B) = \arg\max_{A,C}(A : x + y + 1, C : 2) = C \\
&\pi_1(A) = \arg\max_{B,C}(B : 3, A : x + 1) = B
\end{aligned}$$

Maintenant, évaluez π_1:

$$\begin{aligned}
&V^{\pi_1}(D) = 0 \\
&\Rightarrow V^{\pi_1}(C) = 1 + V^{\pi_1}(D) = 1 \\
&\Rightarrow V^{\pi_1}(B) = 1 + V^{\pi_1}(C) = 2 \\
&\Rightarrow V^{\pi_1}(A) = 1 + V^{\pi_1}(B) = 3
\end{aligned}$$

Effectuez maintenant une amélioration de la politique pour obtenir π_2 :

$$\begin{aligned}
&\pi_2(D) = D \\
&\pi_2(C) = D \\
&\pi_2(B) = \arg\max_{A,C}(A : y + 3, C : 2) = A \\
&\pi_2(A) = \arg\max_{B,C}(B : 3, A : x + 1) = B
\end{aligned}$$

Observez que cette politique est optimale, car la valeur $V^{\pi_2}(A) = V^{\pi_2}(B) = \infty$. Les autres valeurs sont trivialement optimales car l'agent n'a qu'un seul choix d'action.

12.1 Examen 2022-2023

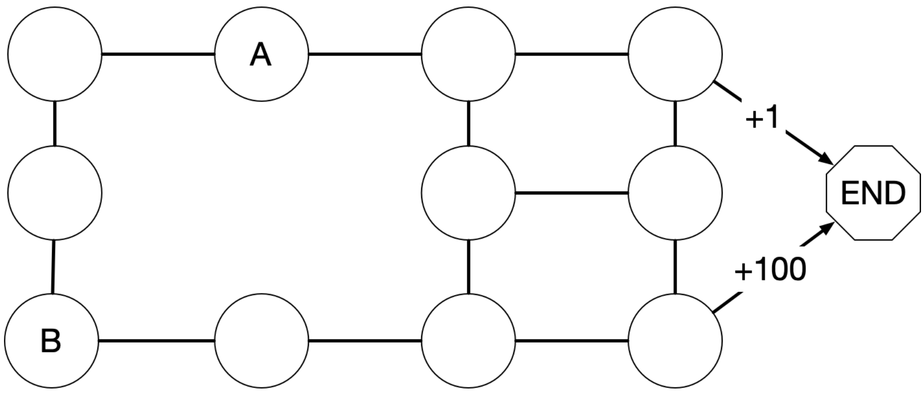


Figure 46: Réseaux MDP. Le réseau est composé de 11 états et un état final avec le label END. Les connections entre les états modélisent les actions autorisées pour un agent. Donc même si un agent peut prendre quatre actions dans chaque état (↑, ↓, ←, →), certaines des actions n'entraîneront pas de mouvement. Dans les états avec un bord dirigé sortant, les agents ont une cinquième action EXIT qui les amène à l'état final, qui sont récompensés par +1 ou +100 selon l'endroit où l'action de sortie est effectuée.

Le problème décrit dans cette question aura lieu dans une instance d'un réseau MDP visualisé dans la Figure 46. Dans tous les états, l'agent dispose des actions ↑, ↓, ← et →. L'exécution d'une action non autorisée entraîne le maintien de l'agent dans son état d'origine. Dans deux états avec une arête dirigée sortant de l'état, l'agent a une action supplémentaire EXIT. Dans le cas où l'action EXIT est effectuée, l'agent reçoit la récompense étiquetée et

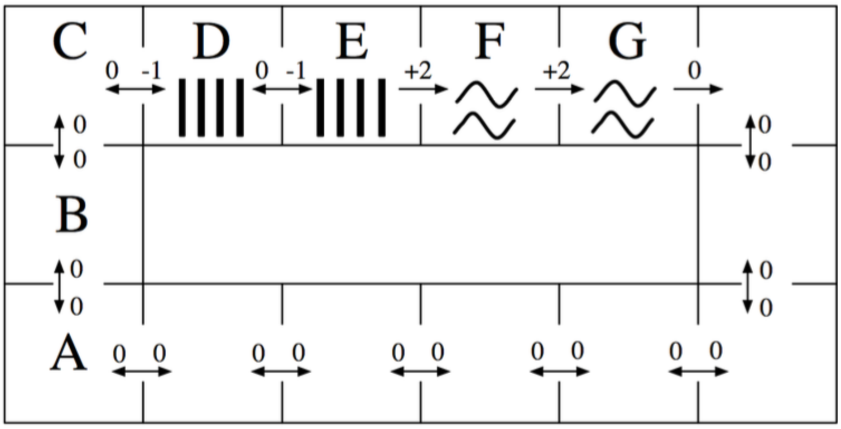


Figure 47: Image de [2].

termine le jeu dans l'état terminal END. Sauf indication contraire, toutes les autres transitions ne reçoivent aucune récompense et toutes les transitions sont déterministes. Pour toutes les parties du problème, supposons que l'itération de valeur (*Value Iteration*) commence avec tous les états initialisés à zéro, c'est-à-dire $V_0(s) = 0 \forall s$. Soit le facteur d'actualisation $\gamma = 1/2$ pour toutes les parties suivantes.

1. Quelles sont les valeurs optimales pour A et B ($V^*(A)$ et $V^*(B)$) dans ce problème ? (2 points) $V^*(A) = \frac{25}{4}$ et $V^*(B) = \frac{25}{2}$

2. Après combien d'itérations k aurons-nous $V_k(s) = V^*(s)$ pour tous les états s ? Si cela ne se produit jamais, écrivez "jamais". (1 point) $k = 6$

3. Supposons que nous voulions reconcevoir la fonction de récompense (*reward function*). Pour laquelle des nouvelles fonctions de récompense suivantes la politique optimale resterait-elle inchangée, soit $R(s, a, s')$ la fonction de récompense original? (2 points)

   (a) $R_1(s, a, s') = 5R(s, a, s') + 17$ ne change pas

   (b) $R_2(s, a, s') = 1 + R(s, a, s')^2$ change, ce n'est pas une transformation affine

   (c) $R_3(s, a, s') = -1$ change, ce n'est pas une transformation affine

Ex. 4

Considérez le MDP dessiné dans la Figure 47. L'espace d'état se compose de toutes les cases d'un parc aquatique du monde de la grille. Il y a un seul toboggan aquatique composé de deux carrés d'échelle et de deux carrés de toboggan (marqués respectivement de barres verticales et de lignes sinueuses). Un agent de ce parc aquatique peut se déplacer de n'importe quelle case vers n'importe quelle case voisine, sauf si la case en cours est un toboggan auquel cas il doit avancer d'une case le long du toboggan. Les actions sont indiquées par des flèches entre les carrés sur la carte et toutes déplacent l'agent de manière déterministe dans la direction donnée. L'agent ne peut pas rester immobile : il doit se déplacer à chaque pas de temps. Les récompenses sont également indiquées ci-dessous : l'agent ressent un grand plaisir en glissant sur le toboggan aquatique (+2), un certain inconfort en gravissant les échelons de l'échelle (−1), et reçoit des récompenses de 0 dans tout autre cas. L'horizon temporel est infini ; ce MDP dure éternellement.

1. Combien de *policies* (déterministes) π sont possibles pour ce MDP ? 2^{11}

2. Remplissez les cellules vides du tableau ci-dessous avec des valeurs correctes pour la fonction, le facteur d'actualisation et l'état correspondants. Astuce : Vous ne devriez pas avoir besoin de faire des calculs substantiels ici.

	γ	$s = A$	$s = B$
$V_3(s)$	1.0	0	4
$V_{10}(s)$	1.0	2	4
$V_{10}(s)$	0.1	0	2.2
$Q_1(s, west)$	1.0	$- - --$	0
$Q_{10}(s, west)$	1.0	$- - --$	3
$V^*(s)$	1.0	∞	∞
$V^*(s)$	0.1	0	2.2

- $V^*_{10}(A)$, $\gamma = 1$: En 10 pas de temps sans actualisation, les récompenses ne décroissent pas donc la stratégie optimale est de monter les deux escaliers (−1 récompense chacun), puis de glisser vers le bas des deux carrés du toboggan (+2 récompenses chacun). Vous n'avez le temps de le faire qu'une seule fois. En résumé, nous obtenons $-1 - 1 + 2 + 2 = 2$.
- $V^*_{10}(E)$, $\gamma = 1$: Pas d'actualisation, donc la stratégie optimale glisse vers le bas du toboggan. Vous n'avez pas le temps de faire plus. Somme des récompenses$= 2 + 2 = 4$.
- $V^*_{10}(A)$, $\gamma = 0.1$: Le taux d'actualisation est de 0.1, ce qui signifie que les récompenses 1 étape plus loin dans le futur sont actualisées par un facteur de 0.1. Supposons qu'à partir de A, nous ayons opté de glisser du toboggan. Ensuite, nous devrions prendre les actions $A \to B$,$B \to C$,$C \to D$,$D \to E$,$E \to F$,$F \to G$. Nous obtenons la première récompense −1 de $C \to D$, actualisée de γ^2 car elle est deux actions dans le futur. $D \to E$ est actualisé par γ^3, $E \to F$ par γ^4 et $F \to G$ par γ^5. Étant donné que γ est faible, les récompenses positives que vous obtenez du toboggan ont moins d'effet que les récompenses négatives plus importantes que vous obtenez en grimpant. Par conséquent, la somme des récompenses de prendre le chemin du toboggan serait négative ; la valeur optimale est 0.
- $V^*_{10}(E)$, $\gamma = 0.1$: Maintenant, vous n'avez plus à faire le travail de monter les escaliers, et vous prenez simplement le toboggan vers le bas. La somme des récompenses serait 2 (pour $E \to F$) +0.2 (pour $F \to G$, réduit de 0.1) $= 2.2$.

- $Q_{10}(E, west)$,$\gamma = 1$: Souvenez-vous qu'un état $Q\ (s, a)$ est le moment où vous partez de l'état s et que vous vous engagez à prendre l'action a. Par conséquent, à partir de E, vous effectuez l'action $west$ et atterrissez en D, en utilisant un pas de temps et en obtenant une récompense immédiate de 0. À partir de D, la stratégie optimale consiste à remonter la volée d'escaliers supérieure, puis à glisser sur le toboggan. Par conséquent, les récompenses seraient $-1(D \to E) + 2(E \to F) + 2(F \to G) = 3$.
- $V^*(s)$,$\gamma = 1$: Jeu infini sans *discount* ? Amusez-vous à glisser du toboggan peu importe où vous êtes.
- $V^*(s)$, $\gamma = 0.1$. Le même raisonnement s'applique à A et E à partir de $V^*_{10}(s)$. Avec *discounting*, les escaliers sont plus coûteux à grimper que la récompense que vous obtenez en glissant sur le toboggan aquatique. Par conséquent, en A, vous ne voudriez pas vous diriger vers le toboggan. À partir de E, puisque vous êtes déjà en haut du tobogan, vous devez simplement glisser vers le bas.

3. Remplissez les cellules vides du deuxième tableau ci-dessous avec les valeurs Q résultant de l'application de la mise à jour Q pour la transition spécifiée sur chaque ligne. Vous pouvez laisser vides les valeurs Q qui ne sont pas affectées par la mise à jour actuelle. Utilisez le discount $\gamma = 1.0$. Supposez que toutes les valeurs Q sont initialisées à 0. (Remarque : les transitions spécifiées ne résulteraient pas d'un seul épisode.)

	$Q(D, west)$	$Q(D, east)$	$Q(E, west)$	$Q(E, east)$
Initial	0	0	0	0
Transition 1: ($s = D$, $a = east$, $r = -1$, $s' = E$)		-0.5		
Transition 1: ($s = D$, $a = east$, $r = -1$, $s' = E$)				1.0
Transition 1: ($s = D$, $a = east$, $r = -1$, $s' = E$)				
Transition 1: ($s = D$, $a = east$, $r = -1$, $s' = E$)		-0.25		

Ex. 5

A	B	C
D	E	F

Pacman utilise les MDP pour maximiser son utilité attendue. Dans chaque environnement :

1. Pacman a les actions standard $\{Nord, Est, Sud, Ouest\}$ à moins qu'il ne soit bloqué par un mur extérieur.

2. Il y a une récompense de 1 point en mangeant le point (par exemple, dans la grille ci-dessus, $R(C, Sud, F) = 1$).

3. Le jeu se termine lorsque le point est mangé.

Considérez la grille ci-dessus où il y a une seule pastille alimentaire dans le coin inférieur droit (F). Le discount factor est 0,5. Il n'y a pas de récompense vivante. Les états sont simplement les emplacements de la grille.

1. Quelle est la politique optimale pour chaque état ?

s	$\pi(s)$
A	*East* or *South*
B	*East* or *South*
C	*South*
D	*East*
E	*East*

2. Quelle est la valeur optimale pour l'état d'être dans le coin supérieur gauche (A) ? Rappel: le facteur d'actualisation γ est de 0.5. $V^*(A) = 0.25$

k	$V(A)$	$V(B)$	$V(C)$	$V(D)$	$V(E)$	$V(F)$
0	0	0	0	0	0	0
1	0	0	1	0	1	0
2	0	0.5	1	0.5	1	0
3	0.25	0.5	1	0.5	1	0
4	0.25	0.5	1	0.5	1	0

3. En utilisant l'itération de valeur avec la valeur de tous les états étant zéro à $k = 0$, pour quelle itération k aura-t-on que $V_k(A) = V^*(A)$? $k = 3$, Comme le tableau ci-dessus l'indique.

13 Apprentissage par renforcement

Ex. 1

Pacman est dans un MDP inconnu à trois états $\{A, B, C\}$ et deux actions $\{Stop, Go\}$. On nous donne les exemples suivants générés à partir de la prise d'actions dans le MDP inconnu. Pour les problèmes suivants, supposons $\gamma = 1$ et $\alpha = 0.5$.

Nous exécutons Q-learning sur les exemples suivants :

s	a	s'	r
A	Go	B	2
C	$Stop$	A	0
B	$Stop$	A	-2
B	Go	C	-6
C	Go	A	2
A	Go	A	-2

1. Quelles sont les estimations des valeurs Q obtenues par Q-learning après cette séquence d'actions ? Toutes les valeurs Q sont initialisées à 0.

   (a) $Q(C, stop) = 0.5$

   (b) $Q(C, Go) = 1.5$

   Pour cela, il suffit de considérer les trois échantillons suivants.

$$Q(A, Go) \leftarrow (1-\alpha)Q(A, Go) + \alpha(r + \gamma \max_a Q(B, a)) = 0.5(0) + 0.5(2) = 1$$
$$Q(C, Stop) \leftarrow (1-\alpha)Q(C, Stop) + \alpha(r + \gamma \max_a Q(A, a)) = 0.5(0) + 0.5(1) = 0.5$$
$$Q(C, Go) \leftarrow (1-\alpha)Q(C, Go) + \alpha(r + \gamma \max_a Q(A, a)) = 0.5(0) + 0.5(3) = 1.5$$

Pour la prochaine partie, nous allons passer à une représentation basée sur les fonctionnalités (*features*). Nous allons utiliser deux fonctionnalités :

- $f_1(s, a) = 1$
- $f_2(s, a) = 1$ if $a = Go$ and $f_2(s, a) = -1$ if $a = Stop$

En partant de poids initiaux de 0, calculez les poids mis à jour après avoir observé les échantillons ci-dessous.

s	a	s'	r
A	Go	B	4
B	$Stop$	A	0

1. Quels sont les poids après la première mise à jour ? (en utilisant le premier échantillon):

   (a) $w_1 = 2$

   (b) $w_2 = 2$

$$Q(A, Go) = w_1 f_1(A, Go) + w_2 f_2(A, Go) = 0$$
$$\text{difference} = [r + \max_a Q(B, a)] - Q(A, Go) = 4$$
$$w_1 = w_1 + \alpha(\text{difference}) f_1 = 2$$
$$w_2 = w_2 + \alpha(\text{difference}) f_2 = 2$$

2. Quels sont les poids après la deuxième mise à jour ? (en utilisant le deuxième échantillon):

   (a) $w_1 = 4$

   (b) $w_2 = 0$

$$Q(B, Stop) = w_1 f_1(B, stop) + w_2 f_2(B, Stop) = 2(1) + 2(-1) = 0$$
$$Q(A, Go) = w_1 f_1(A, Go) + w_2 f_2(A, Go) = 2(1) + 2(1) = 4$$
$$\text{difference} = [r + \max_a Q(A, a)] - Q(B, Stop) = [0 + 4] - 0 = 4$$
$$w_1 = w_1 + \alpha(\text{difference}) f_1 = 4$$
$$w_2 = w_2 + \alpha(\text{difference}) f_2 = 0$$

Ex. 2

Considérons un MDP inconnu à trois états (A, B et C) et deux actions ($\leftarrow$ et $\rightarrow$). Supposons que l'agent choisisse des actions selon une politique π dans le MDP inconnu, en collectant un ensemble de données composé d'échantillons (s,a,s',r) représentant l'action a dans l'état s résultant en une transition vers l'état s' et une récompense de r. Vous pouvez supposer un discount factor de $\gamma = 1$.

s	a	s'	r
A	$\rightarrow$	B	2
C	$\leftarrow$	B	2
B	$\rightarrow$	C	-2
A	$\rightarrow$	B	4

1. Supposons que toutes les valeurs Q soient initialisées à 0 et utilisent un taux d'apprentissage de $\alpha = \frac{1}{2}$. Exécutez Q-learning sur le tableau d'expérience ci-dessus et remplissez les valeurs Q suivantes:

   (a) $Q(A, \rightarrow) = 5/2$

   (b) $Q(B, \rightarrow) = -1/2$

$$
\begin{aligned}
Q_1(A,\rightarrow) &= \frac{1}{2}Q_0(A,\rightarrow) + \frac{1}{2}\left(2+\gamma\max_{a'}Q_0(B,a')\right) = 1\\
Q_1(C,\leftarrow) &= 1\\
Q_1(B,\rightarrow) &= \frac{1}{2}(-2+1) = -\frac{1}{2}\\
Q_2(A,\rightarrow) &= \frac{1}{2}\cdot 1 + \frac{1}{2}\left(4+\gamma\max_{a'}Q_1(B,a')\right)\\
&= \frac{1}{2} + \frac{1}{2}(4+0) = \frac{5}{2}
\end{aligned}
$$

2. Après avoir exécuté Q-learning et produit les valeurs Q ci-dessus, vous construisez une politique π_Q qui maximise la valeur Q dans un état donné. Quelles sont les actions choisies par la politique dans les états A et B ?

   (a) $\pi_Q(A)$ est égal à i) $\leftarrow$, ii) $\rightarrow$, ou iii) indéfini $\pi_Q(A) = \rightarrow$

   (b) $\pi_Q(B)$ est égal à i) $\leftarrow$, ii) $\rightarrow$, ou iii) indéfini $\pi_Q(B) = \leftarrow$

   Remarquez que $Q(B,\leftarrow) = 0 > -\frac{1}{2} = Q(B,\rightarrow)$

3. Utilisez la méthode empirique d'apprentissage par renforcement basée sur un modèle de comptage de fréquence décrite dans les cours pour estimer la fonction de transition $\hat{T}(s,a,s')$ et la fonction de récompense $\hat{R}(s,a,s')$. (N'utilisez pas de pseudo-comptes ; si une transition n'est pas observée, elle a un compte de 0.) Notez les quantités suivantes. Vous pouvez écrire N/A pour des quantités indéfinies.

   $\hat{T}(A,\rightarrow,B) = 1$ $\quad$ $\hat{R}(A,\rightarrow,B) = 3$
   $\hat{T}(B,\rightarrow,A) = 0$ $\quad$ $\hat{R}(B,\rightarrow,A) = N/A$
   $\hat{T}(B,\leftarrow,A) = N/A$ $\quad$ $\hat{R}(B,\leftarrow,A) = N/A$

Ex. 3

Cette question examine les propriétés des algorithmes d'apprentissage par renforcement pour les MDP discrets arbitraires ; vous n'avez pas besoin de vous référer au MDP considéré dans les parties précédentes.

Laquelle des méthodes suivantes, à la convergence, fournit suffisamment d'informations pour obtenir une politique optimale ? (En supposant une exploration adéquate.)

1. Apprentissage basé sur un modèle de $T(s,a,s')$ et $R(s,a,s')$.
2. Évaluation directe pour estimer $V(s)$.
3. Apprentissage de la différence temporelle pour estimer $V(s)$.
4. Q-Learning pour estimer $Q(s,a)$.

1 et 4. Avec suffisamment de données, l'apprentissage basé sur un modèle se rapprochera arbitrairement du véritable modèle de l'environnement, auquel cas la planification (par exemple, *value iteration*) peut être utilisée pour trouver

une politique optimale. De même, le Q-learning est assuré de converger vers les valeurs Q optimales de la politique optimale, point auquel la politique optimale peut être récupérée par $\pi^*(s) = \arg\max_a Q(s,a)$. L'évaluation directe et l'apprentissage par différence temporelle ne récupèrent tous deux qu'une fonction de valeur $V(s)$, ce qui est insuffisant pour choisir entre des actions sans connaître les probabilités de transition.

Ex. 4

Lorsque le nombre de pas tends vers l'infini, sous laquelle des politiques d'exploration suivantes le Q-learning est-il garanti de converger vers les valeurs Q optimales pour tous les états ? (Vous pouvez supposer que le taux d'apprentissage α est choisi de manière appropriée et que le MDP est ergodique : c'est-à-dire que chaque état est accessible à partir de tous les autres états avec une probabilité non nulle.)

1. Une politique fixe prenant des actions uniformément au hasard.
2. Une politique *greedy*.
3. Une politique ϵ-*greedy*.
4. Une politique optimale fixe.

1 et 3. Pour que le Q-learning converge, chaque paire état-action (s,a) doit se produire infiniment souvent. Une politique aléatoire uniforme permettra d'atteindre cet objectif dans un MDP ergodique. Une politique optimale fixe ne prendra aucune action sous-optimale et n'explorera donc pas suffisamment. De même, une politique gourmande (*greedy*) cessera de prendre des mesures que les valeurs Q actuelles suggèrent comme sous-optimales, et ne mettra donc jamais à jour les valeurs Q pour des actions supposées sous-optimales. (Ceci est problématique si, par exemple, une action ne rapporte la plupart du temps aucune récompense, mais occasionnellement une très grande récompense. Après avoir observé l'absence de récompense à quelques reprises, le Q-learning avec une politique gourmande arrêterait de prendre cette action, n'obtenant jamais la haute récompense nécessaire afin de mettre à jour dans la direction de sa vraie valeur.)

13.1 Examen 2022-2023

Considérons un MDP inconnu à 4 états (CC,CA,AC,AA) et deux actions (C et A, signifiant la coopération et l'abstention). Supposons qu'un agent choisisse des actions selon une politique π dans le MDP inconnu, en collectant un ensemble de données composé d'échantillons (s, a, s', r) représentant l'action a dans l'état s résultant en une transition vers l'état s' et une récompense de r. Nous supposons un facteur d'actualisation (*discount factor*) de $\gamma = 1$.

- Supposons que toutes les valeurs Q soient initialisées à 0 et utilisent un taux d'apprentissage de $\alpha = 1/2$. Exécutez Q-learning sur le tableau d'expérience ci-dessus et remplissez les valeurs Q finales pour: (2 points)

s	a	s'	r
CC	C	CA	-1
AC	A	AC	4
AC	C	CA	-1
CA	C	CC	2
AA	A	AC	4

Table 2: Les echantillons

$$\begin{aligned}
Q(CC,C) &= (1-\alpha)Q(CC,C) + \alpha(r + \gamma \max_{a'} Q(CA,a')) \\
&= \frac{1}{2}(0) + \frac{1}{2}(-1+0) = -\frac{1}{2} \\
Q(AC,A) &= (1-\alpha)Q(AC,A) + \alpha(r + \gamma \max_{a'} Q(AC,a')) \\
&= \frac{1}{2}(0) + \frac{1}{2}(4+0) = 2 \\
Q(AC,C) &= (1-\alpha)Q(AC,C) + \alpha(r + \gamma \max_{a'} Q(CA,a')) \\
&= \frac{1}{2}(0) + \frac{1}{2}(-1+0) = -\frac{1}{2} \\
Q(CA,C) &= (1-\alpha)Q(CA,C) + \alpha(r + \gamma \max_{a'} Q(CC,a')) \\
&= \frac{1}{2}(0) + \frac{1}{2}(2+0) = 1 \\
Q(AA,A) &= (1-\alpha)Q(AA,A) + \alpha(r + \gamma \max_{a'} Q(AC,a')) \\
&= \frac{1}{2}(0) + \frac{1}{2}(4+2) = 3
\end{aligned}$$

1. $Q(CC,C) =-\frac{1}{2}$
2. $Q(AC,A) =2$
3. $Q(CA,C) =1$
4. $Q(AA,C) =0$

- Après avoir exécuté Q-learning et produit les valeurs Q ci-dessus, vous construisez une politique π_Q qui maximise la valeur Q dans un état donné. Quelles sont les actions choisies par la politique dans les états AA, AC, CA et CC ? (2 points) AA:A. AC:A, CA:C, CC:A

A	1	1	1	1	0	1	0	1	1	1
B	1	0	0	1	1	1	1	0	1	1
Y	1	1	0	0	0	1	1	0	0	0

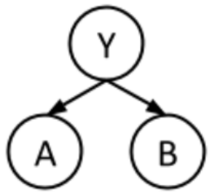


Figure 48: Image de [2].

14 Apprentissage automatique

Ex. 1

Dans cette question, nous allons former un classificateur Naive Bayes pour prédire les étiquettes de classe Y en fonction des caractéristiques d'entrée A et B (Figure 48). Y, A et B sont toutes des variables binaires, avec des domaines 0 et 1. Nous recevons 10 points d'apprentissage à partir desquels nous allons estimer notre distribution.

1. Quelles sont les estimations du maximum de vraisemblance pour les tables $P(Y)$, $P(A|Y)$ et $P(B|Y)$ (Figure 49)?

Y	$P(Y)$
0	3/5
1	2/5

A	Y	$P(A\|Y)$
0	0	1/6
1	0	5/6
0	1	1/4
1	1	3/4

B	Y	$P(B\|Y)$
0	0	1/3
1	0	2/3
0	1	1/4
1	1	3/4

Figure 49: Image de [2].

2. Considérons un nouveau point de données ($A = 1$ et $B = 1$). Quelle étiquette ce classificateur attribuerait-il à cet échantillon ?

$$\begin{aligned} P(Y=0, A=1, B=1) &= P(Y=0)P(A=1|Y=0)P(B=1|Y=0) \\ &= (3/5)(5/6)(2/3) = 1/3 \\ P(Y=1, A=1, B=1) &= P(Y=1)P(A=1|Y=1)P(B=1|Y=1) \\ &= (2/5)(3/4)(3/4) = 9/40 \end{aligned}$$

Le classificateur prédira le label 0.

3. Utilisons *Laplace smoothing* pour lisser notre distribution. Calculez la nouvelle distribution pour $P(A|Y)$ étant donné le *Laplace smoothing* de $k = 2$ (Figure 50).

A	Y	$P(A\|Y)$
0	0	3/10
1	0	7/10
0	1	3/8
1	1	5/8

Figure 50: Image de [2].

Ex. 2

Vous voulez prédire si les films seront rentables en fonction de leurs scénarios. Vous engagez deux critiques A et B pour lire un scénario que vous avez et le noter sur une échelle de 1 à 4. Les critiques ne sont pas parfaites ; Ci-dessous cinq points de données, y compris les notes des critiques et la performance du film.

#	Movie Name	A	B	Profit?
1	Pellet Power	1	1	-
2	Ghosts!	3	2	+
3	Pac is Bac	2	4	+
4	Not a Pizza	3	4	+
5	Endless Maze	2	3	-

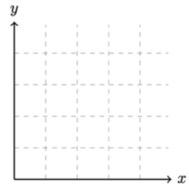


Figure 51: Image de [2].

1. Tout d'abord, vous souhaitez examiner la séparabilité linéaire des données. Tracez les données sur le plan 2D (Figure 51) ; étiquetez les films rentables avec $+$ et les films non rentables avec $-$ et déterminez si les données sont linéairement séparables.

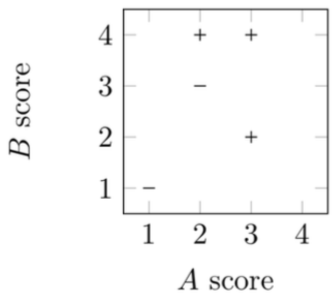


Figure 52: Image de [2].

Les données sont linéairement séparables, regardez Figure 52.

2. Vous décidez maintenant d'utiliser un perceptron pour classer vos données. Supposons que vous utilisiez directement les scores donnés ci-dessus comme caractéristiques, avec une caractéristique de biais. Soit $f_0 = 1$, $f_1 =$ score donné par A et $f_2 =$ score donné par B. Exécutez un passage à travers les données avec l'algorithme perceptron, en remplissant le tableau ci-dessous. Parcourez les points de données dans l'ordre, par ex. en utilisant le point de données #1 à l'étape 1.

Step	Weights	Score	Correct?
1	$[-1, 0, 0]$	$-1 \cdot 1 + 0 \cdot 1 + 0 \cdot 1 = -1$	yes
2	$[-1, 0, 0]$	$-1 \cdot 1 + 0 \cdot 3 + 0 \cdot 2 = -1$	no
3	$[0, 3, 2]$	$0 \cdot 1 + 3 \cdot 2 + 2 \cdot 4 = 14$	yes
4	$[0, 3, 2]$	$0 \cdot 1 + 3 \cdot 3 + 2 \cdot 4 = 17$	yes
5	$[0, 3, 2]$	$0 \cdot 1 + 3 \cdot 2 + 2 \cdot 3 = 12$	no

Les poids finaux sont $[-1, 1, -1]$.

3. A-t-on appris des poids qui séparent les données ? Avec les poids actuels, les points seront classés comme positif si $(-1) \cdot 1 + 1 \cdot A + (-1) \cdot B \geq 0$, ou $A - B \geq 1$. Nous aurons donc des prédictions incorrectes pour les points de données 3 :
$$(-1) \cdot 1 + 1 \cdot 2 + (-1) \cdot 4 = -3 < 0$$
et 4:
$$(-1) \cdot 1 + 1 \cdot 3 + (-1) \cdot 4 = -2 < 0$$
Notez que bien que le point 2 ait $w \cdot f = 0$, il sera classé comme positif (puisque nous classons comme positif si $w \cdot f \geq 0$).

4. Plus généralement, quelles que soient les données d'entraînement, vous souhaitez savoir si vos fonctionnalités sont suffisamment puissantes pour vous permettre de gérer une gamme de scénarios. Encerclez les scénarios pour lesquels un perceptron utilisant les fonctionnalités ci-dessus peut en effet parfaitement classer les films qui sont rentables selon les règles données :

   (a) Vos critiques sont géniaux : si le total de leurs scores est supérieur à 8, alors le film sera certainement rentable, sinon il ne le sera pas. Peut classifier (considérer les poids $[-8, 1, 1]$)

   (b) Vos évaluateurs sont des critiques d'art. Votre film sera rentable si et seulement si chaque critique donne une note de 2 ou une note de 3. Impossible de classifier.

   (c) Vos critiques ont des goûts bizarres mais différents. Votre film sera rentable si et seulement si les deux critiques sont d'accord. Impossible de classifier.

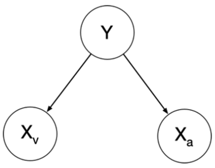
14.1 Examen 2022-2023

Vous voulez former une IA ornithologue, qui utilisera une caméra et un équipement audio pour déterminer si un oiseau repéré est un pinson ou un moineau. Vous obtenez ainsi deux observations : une visuelle et une auditive, désignées par les variables aléatoires X_v et X_a, respectivement. L'observation visuelle vous informe que l'oiseau est soit un pinson ($X_v = 1$) soit un moineau ($X_v = 0$). L'observation auditive X_a est définie de manière analogue. Vos observations sont une mesure bruyante du vrai type de l'individu, qui est noté Y. Après avoir personnellement examiné l'enregistrement visuel et auditif, vous découvrez ce qu'ils sont vraiment : soit un pinson ($Y = 1$), soit un moineau ($Y = 0$). Vous avez enregistré les observations et les vrais types des 20 premières rencontres d'oiseaux :

Oiseau i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Observation visuelle X_v	0	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	0
Observation auditive X_a	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
Type oiseau Y	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0

Maintenant, une observation pour un oiseau $i = 21$ est fait et vous voulez prédire le type Y de cet oiseau étant donné que vous avez observé $X_v = 1$ et$X_a = 1$.

Supposons que les types sont indépendants et que les observations sont indépendantes conditionnées par le type. Vous pouvez modéliser cela en utilisant un modèle Naïf Bayes, avec X_v et X_a, comme caractéristiques et Y comme label. Supposons que les distributions de probabilité sont les suivant :



X_v	Y	$P(X_v\|Y)$
0	0	p_v
0	1	$(1-p_v)$
1	0	$(1-p_v)$
1	1	p_v

X_a	Y	$P(X_a\|Y)$
0	0	p_a
0	1	$(1-p_a)$
1	0	$(1-p_a)$
1	1	p_a

Y	$P(Y)$
0	$(1-q)$
1	q

1. Quelle est l'estimation du maximum de vraisemblance (*Maximum likelihood estimation*) de p_v, p_a et q ? (2 points) $p_v = \frac{4}{5}$, $p_a = \frac{3}{5}$ et $q = \frac{1}{2}$. Pour l'estimation de q, on compte $10 \times Y = 1$ et $10 \times Y = 0$. Pour p_v, nous avons $\frac{8}{10}$ cas où $X_v = 1$ sachant que $Y = 1$ et pour $(1 - p_v)$ nous avons $\frac{2}{10}$ cas où $X_v = 1$ sachant que $Y = 0$. Pour p_a, nous avons $\frac{2}{10}$ cas où $X_a = 1$ sachant que $Y = 1$ et pour $(1 - p_a)$ nous avons $\frac{0}{10}$ cas où $X_a = 1$ sachant que $Y = 0$. La moyenne de $\frac{2}{10}$ et $\frac{10}{10}$ donne donc $p_a = \frac{3}{5}$.

2. Quelle est la probabilité que le prochain oiseau soit un pinson compte tenu de vos observations pour $i = 21$? Exprimez votre réponse en fonction des paramètres p_v, p_a et q (dont vous n'aurez peut-être pas tous besoin). (1 point)

$$P(Y = 1|X_v = 1, X_a = 1) = \frac{p_a p_v q}{p_a p_v q + (1 - p_a)(1 - p_v)(1 - q)}$$

Le numérateur étant la distribution conjointe des trois variables et le dénominateur la somme de la distributions conjointe pour $Y = 1$ et $Y = 0$.

3. À l'aide de le vôtre réponse à la partie (b) ainsi que les valeurs obtenues à la partie (a), calculez la probabilité réelle que le prochain oiseau soit un pinson. (1 point)

$$\begin{aligned} P(Y = 1|X_v = 1, X_a = 1) &= \frac{p_a p_v q}{p_a p_v q + (1 - p_a)(1 - p_v)(1 - q)} \\ &= \frac{\frac{4}{5}\frac{3}{5}\frac{1}{2}}{\frac{4}{5}\frac{3}{5}\frac{1}{2} + \frac{1}{5}\frac{2}{5}\frac{1}{2}} \\ &= \frac{0.24}{0.28} \approx 0.86 \end{aligned}$$

Est-il possible de créer un prédicteur en utilisant un perceptron ? Supposons que nous utilisions les observations répertoriées en haut et que les caractéristiques suivantes soient utilisées : $f_0 = 1$, f_1 est l'observation visuelle, et f_2 est l'observation auditive. Prenez les éléments

$T = 1, 2, 3, 5, 7, 8, 10, 11, 12, 13, 16, 17, 18, 19, 20$

comme ensemble d'entrainement et le reste comme ensemble d'évaluation ($E = 4, 6, 9, 14, 15$).

Oiseau i	1	2	3	5	7	8	10	11	12	13	16	17	18	19	20	4	6	9	14	15
Observation visuelle X_v	0	0	1	1	0	1	1	0	1	1	1	1	0	0	0	0	0	1	0	1
Observation auditive X_a	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Type oiseau Y	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1

1. Exécutez un passage à travers les données d'entrainement avec l'algorithme perceptron. Vous pouvez sauter des étapes si elles n'entraînent pas de changements dans le vecteur de poids. Parcourez les points de données dans l'ordre en supposant que le vecteur de poids initial est $[-1, 0, 0]$. (2 points)

Etape	Poids	Score Remarque : '.' est la multiplication	Class	Correct?
1	$[-1, 0, 0]$	$-1.1 + 0.0 + 0.0 = -1$	$Y = 0$	Oui
2	$[-1, 0, 0]$	$-1.1 + 0.0 + 0.0 = -1$	$Y = 0$	Oui
3	$[-1, 0, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	Oui
		Aucun changement jusqu'à étape 8		
8	$[-1, 0, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	Non
10	$[0, 1, 0]$	$0.1 + 1.1 + 0.0 = 1$	$Y = 1$	Oui
11	$[0, 1, 0]$	$0.1 + 0.1 + 0.0 = 0 (\geq 0)$	$Y = 1$	Oui
		Aucun changement jusqu'à étape 18		
18	$[0, 1, 0]$	$1.0 + 0.1 + 0.1 = 0$	$Y = 1$	Non
19	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	Oui
20	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	Oui

2. Donnez la probabilité que l'instance appartienne à la classe $Y = 1$ et $Y = 0$ et expliquez votre réponse. (2 points) On a donc comme vecteur $[-1, 1, 0]$. Pour le cas $i = 21$ on a comme score $-1.1 + 1.1 + 0.1 = 0$, donc $Y = 1$. La probabilité est calculée par la fonction sigmoïde $1/(1 + e^{-z}) = \frac{1}{2}$. Notre perceptron est donc moins sûr de ses conclusions.

3. Parmi toutes les observations dans l'ensemble d'entrainement, combien sont prédites correctement et combien sont fausses ? Faites le même calcul pour l'ensemble de validation. Est-ce un bon prédicteur ? Motivez votre réponse. (2 points) Résultats sur les données d'entraînement:

Instance	Poids	Score	Class	montant
$[1, 0, 0]$	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	6/7
$[1, 0, 1]$	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.1 = -1$	$Y = 0$	0/0
$[1, 1, 0]$	$[-1, 1, 0]$	$-1.1 + 1.1 + 0.0 = 0$	$Y = 1$	4/6
$[1, 1, 1]$	$[-1, 1, 0]$	$-1.1 + 1.1 + 1.0 = 0$	$Y = 1$	2/2

Résultats sur les données d'évaluation:

Instance	Poids	Score	Class	montant
$[1, 0, 0]$	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.0 = -1$	$Y = 0$	$2/3$
$[1, 0, 1]$	$[-1, 1, 0]$	$-1.1 + 1.0 + 0.1 = -1$	$Y = 0$	$0/0$
$[1, 1, 0]$	$[-1, 1, 0]$	$-1.1 + 1.1 + 0.0 = 0$	$Y = 1$	$2/2$
$[1, 1, 1]$	$[-1, 1, 0]$	$-1.1 + 1.1 + 1.0 = 0$	$Y = 1$	$0/0$

Ce n'est pas un mauvais prédicteur car il a une précision de 80% sur toutes les instances d'entraînement et aussi 80% avec l'ensemble d'évaluation. Comme vous pouvez le voir, l'observation auditive ne semble pas être une caractéristique pertinente pour la prédiction, l'observation visuelle est apparemment suffisante. Notez cependant qu'il n'a jamais vu de cas avec $f_1 = 1$ et $f_2 = 0$ pendant l'entraînement.

References

[1] Russell, S., Norvig, P. (2021) Artificial Intelligence, A modern approach. 4^{th} edition, global edition. Pearson Education.

[2] Homeworks and exercices available for the course CS188, ai.berkeley.edu.