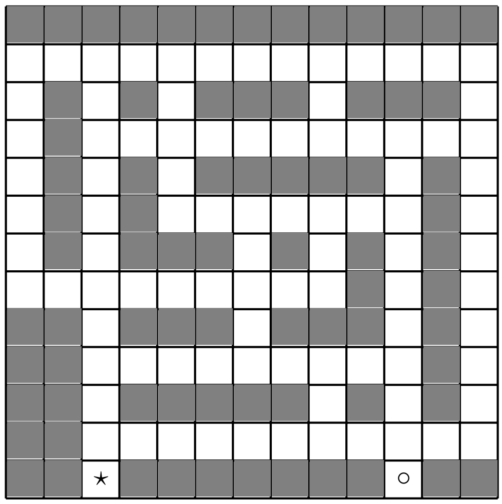
Série 1 : Recherche Heuristique

Exercice 1: Modélisation de graphes d'états

On s'intéresse pour commencer à un labyrinthe proposé par Robert Abbott. Comme vous l'aurez deviné, il s'agit d'entrer par la case représentée par $\circ$ et de sortir par la case représentée par $\star$. Facile... mais l'histoire va un peu se compliquer.
Je vous invite à consulter son site (`http://www.logicmazes.com/`) pour de nombreux autres exemples.



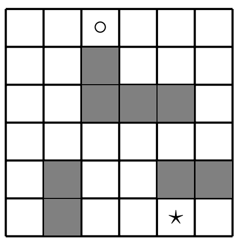
1. Quel modélisation de ce problème pouvez-vous proposer ? Quel est le facteur de branchement de votre problème ?

2. Résoudre le problème en utilisant une stratégie de recherche *greedy best-first* (expansion du noeud de la frontière avec la plus petite valeur de h) en prenant pour h la distance de Manhattan.

3. A présent, changeons légèrement les règles : il est désormais interdit de tourner à gauche (et de faire demi-tour). Quelle modification devez vous apporter à votre modélisation ? Quelle est la conséquence sur le nombre d'état ? L'heuristique employée reste-t-elle admissible ? (Au fait, avez-vous trouvé la solution ?)

Exercice 2: Recherche de chemins

Abordons à présent un environnement moins contraint. L'objectif est encore une de vous déplacer de $\circ$ jusqu'à $\star$. Les déplacements autorisés sont verticaux et horizontaux, pour commencer.



1. Vous souhaitez utiliser l'algorithme A^* pour résoudre ce problème : vaut-il mieux utiliser la distance de Manhattan ou la distance euclidienne ?

2. Déroulez l'algorithme A^*, en indiquant à chaque étape l'état de la *frontière* courante et de la *réserve*. Lorsque plusieurs choix s'offrent à vous pour le noeud à étendre, vous être libres de faire le choix le plus "malin".

3. Pour aller plus loin : imaginons maintenant que notre agent ne se déplace pas de manière discrète sur une grille, mais comme il le souhaite dans le plan. Quelle technique pourriez-vous imaginer pour trouver le plus court chemin dans ce cas ?

Exercice 3: Autour des heuristiques

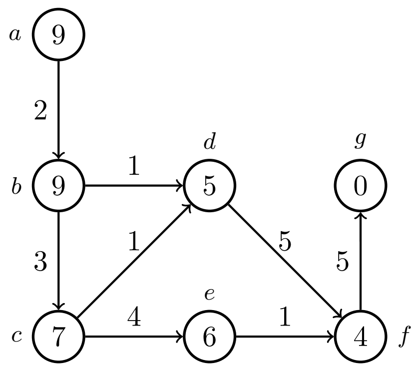
1. Montrer que, comme nous l'avons affirmé en cours, sous l'hypothèse de consistance de l'heuristique, les valeurs de $f(n)$ sont non-décroissantes sur les noeuds n composant n'importe quel chemin vers un noeud but.

2. Montrer que consistance implique l'admissibilité d'une heuristique.

3. Que se passe-t-il si l'heuristique est *parfaite*, c'est-à-dire qu'elle évalue toujours exactement la longueur du chemin jusqu'à l'état but ?

4. Dans de nombreuses situations, la garantie d'optimalité n'est pas cruciale et on peut alors utiliser des heuristiques non-admissibles qui permettent d'accélérer le traitement. Pourriez-vous en tester une sur un des exemples précédents ?

5. Sur l'exemple suivant, la valeur indiquée dans le noeud est la valeur de l'heuristique. La valeur sur les arcs est la coût du déplacement d'un noeud à l'autre. L'heuristique donnée est-elle admissible ? Est-elle consistante ?



6. Une technique connue permettant de rendre une heuristique consistante est de simplement mettre à jour les heuristiques en appliquant l'ajustement de valeur suivant (avec p fils de n) :

$$h'(n) = \max(h(p), h(n) - cost(n,p))$$

Cette méthode est communément appelée le *MaxPath trick*. Appliquez là sur l'exemple précédent.

Série 2 : Jeux

Exercice 1: Un premier jeu simple...

Dans la table qui suit, les vecteurs de gains donnent le gain de l'agent 1, puis celui de l'agent 2. L'agent 1 dispose de 2 stratégies possibles : a et b, tandis que l'agent 2 dispose de 3 stratégies : r, s, t. Ainsi, si 1 joue a et 2 joue t, le gain de 1 sera 0 et celui de 2 sera 5 (case (a, t) dans la matrice).

	r	s	t
a	(3,3)	(1,2)	(0,5)
b	(1,4)	(0,0)	(2,1)

1. Existe-t-il une stratégie dominante pour l'un des joueurs ? Existe-t-il un ou des équilibre(s) de Nash ?
2. En supposant à présent que les agents jouent de manière séquentielle, indiquez pour chacun des agents s'il est prérérable de jouer en premier ou en second.

Exercice 2: Colonel Blotto

Le jeu dit du Colonel Blotto oppose deux joueurs, qui ont chacun n troupes à allouer sur des champs de batailles. Ces champs de batailles sont au nombre de k. On suppose que, sur un champ de bataille donné, le joueur qui a disposé le plus de troupes gagne la bataille (la bataille peut donner un match nul en cas de même nombre de troupes). Au final, le joueur qui remporte le plus de batailles gagne la partie. Il s'agit donc d'un jeu à somme nulle. Une stratégie pour un joueur consiste donc à décider de l'allocation de ses troupes sur les champs, et peut s'écrire comme un vecteur

$$(x_1, x_2, \dots x_k)$$

où la coordonnée x_i est le nombre de troupes allouées sur le champ i.

Commençons par considérer le cas $k = 3$ et $n = 6$, avec la contrainte suivante : le nombre de troupes doit nécessairement être croissante avec les indices des champs.

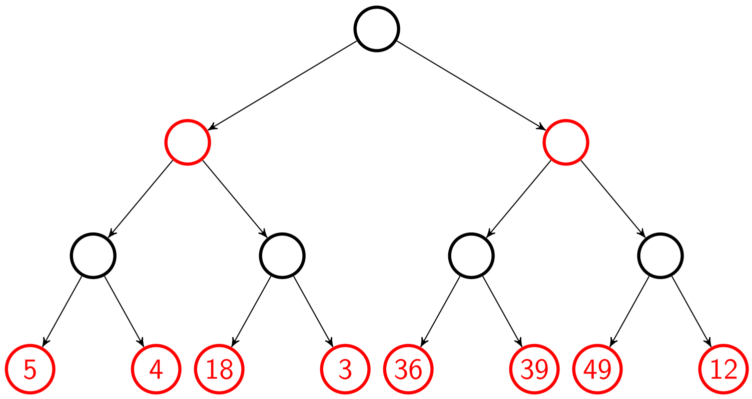
1. Lister les stratégies des agents. Représentez le jeu sous forme matricielle, en considérant que le gain est : 1 pour un joueur qui gagne la partie, 0 pour un match nul, et -1 pour une défaite.

2. Existe-t-il une stratégie dominante ? Donnez les équilibres de Nash de ce jeu.

3. Montrer, de manière générale, qu'il n'existe pas de stratégie dominante à ce jeu. Existe-t-il des stratégies dominées ?

Exercice 3: Minimax et élagage alpha-beta

On considère l'arbre de recherche suivant, dans lequel les nœuds noirs sont des nœuds MAX, tandis que les nœuds gris sont des nœuds MIN. Les valeurs dans les feuilles sont les valeurs obtenues dans chacun de ces états finaux.

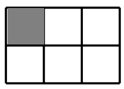


1. Indiquez par utilisation de l'algorithme minimax la valeur du nœud racine

2. Utilisez ensuite l'algorithme avec les coupes α-β. Que se passerait-il si vous aviez évalué les nœuds dans l'ordre inverse ?

Exercice 4: Le jeu de la tablette de chocolat

Dans le jeu de la tablette de chocolat (*chomp game*), deux joueurs doivent manger à tour de rôle une tablette de $m \times n$ carrés, mais malheureusement le carré en haut à gauche est empoisonné : le joueur qui le mange a perdu. La règle est la suivante : lorsqu'un joueur mange un carré, tous les carrés situés dans la zone sud-est du carré mangé disparaissent.



1. Sur un jeu de tablette 2×3, analysez l'arbre de recherche obtenu pour ce jeu. Que pouvez-vous en conclure ?

2. Sur une tablette $2 \times n$, existe-t-il selon vous une stratégie dominante ? Si non, justifiez votre réponse. Si oui, quelle est-elle ?

3. Sur une tablette $n \times n$, existe-t-il selon vous une stratégie dominante ? Si non, justifiez votre réponse. Si oui, quelle est-elle ?

$\triangleright$ Petite note historique : la version “barre chocolatée” du jeu a été proposée par David Gale (le même Gale que celui de l'algorithme des mariages stables).

Série 3 : Dynamiques multi-agents

Exercice 1: Dynamique de meilleures réponses

On considère une famille composée d'un père (p), d'une mère (m), d'une fille (f) et d'un garçon (g). Chacun doit prendre une décision binaire, aller au cinéma (oui ou non). Leur décision dépend de celles des autres. Les stratégies des agents sont les suivantes :

— le père préfère toujours prendre la même décision que la mère ;

— le fils s'oppose toujours au père, quelque soit sa décision ;

— la mère préfère choisir comme son fils et sa fille lorsque ceux-ci sont d'accord ; et sinon elle choisit comme le père.

— la fille est complètement indépendante.

1. Indiquez pourquoi les dynamiques de réponses améliorantes et de meilleures réponses sont équivalentes dans ce problème

2. Donnez les tables de meilleures réponses des différents agents.

3. Vous considèrerez les deux cas : selon que la fille choisit d'aller ou cinéma ou non. Dans chacun des cas, indiquez si il existe une situation d'équilibre. Indiquez également si une dynamique de meilleure réponse permet, dans chacun des cas, de converger vers une situation d'équilibre.

Exercice 2: Jeu de congestion

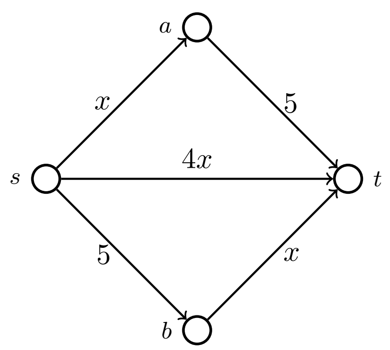
On considère un jeu dit de "congestion", dans lequel 6 agents doivent se déplacer sur un réseau pour aller d'un nœud s à un nœud t. Dans ce problème :

— les stratégies pour les agents sont les chemins possibles pour aller de s à t. Une situation de jeu est la collection $S = \langle s_1, \dots s_6 \rangle$ des stratégies des agents.

— le coût à passer sur un arc est, pour un agent, donné par une fonction linéaire $cout(e, x) = ax+b$ où x est le nombre d'agents empruntant cet arc

— le coût pour un agent i d'emprunter un chemin $\pi = [e_1, \dots, e_k]$ dans une situation S, est donné par $\sum_{e_j \in \pi} cout(e_j, S(e_j))$, où $S(e_j)$ est le nombre d'agents empruntant l'arc dans S.

— une réponse améliorante consiste à changer son chemin pour un chemin de moindre coût



1. En partant de la situation initiale où tous les agents empruntent le chemin $\pi = [(s, a), (a, t)]$, indiquez la séquence d'actions observées dans une dynamique de meilleures réponses.

2. Exhibez à présent une situation d'équilibre qui peut être atteinte par une dynamique de réponses améliorantes, mais pas pas une dynamique de meilleure réponse.

3. Pour une situation S, considérez la fonction $\phi(s) = \sum_{e \in G} \sum_{i=1}^{S(e)} cout(e, i)$, où rappelons-le $S(e)$ est le nombre d'agents empruntant l'arc e. Calculez sa valeur pour la séquence de meilleures réponses de la question 1. Montrez que la valeur de cette fonction décroît nécessairement après une réponse améliorante. Que pouvez-vous en déduire sur les séquences de réponses améliorantes dans ce problème ?

4. Un concepteur du réseau vous suggère de remplacer l'arc (s, t) par un arc (a, b) de coût fixe nul. Que pouvez-vous dire de la situation où tous les agents empruntent le chemin $[(s, a), (a, b), (b, t)]$?

Exercice 3: Fictitious Play

On considère la situation de jeu donnée par la matrice suivante :

$R \backslash B$	a	b
a	$(3, 2)$	$(0, 0)$
b	$(0, 0)$	$(2, 3)$

Il s'agit d'un jeu de coordination : le joueur B a une préférence pour la coordination sur l'action a, le joueur R a une préférence sur l'action b, mais les deux joueurs préfèrent être coordonnés sur la même action que non-coordonnés.

1. Calculer sur 4 itérations les coups joués si les deux agents suivent une stratégie de fictitious play, en partant de l'information a priori suivante : pour R : $(a : 3, b : 1)$ et pour B : $(a : 1, b : 2)$. Autrement dit, R dispose d'un échantillon initial où il a observé que B avait joué 3 fois a et 1 fois b, tandis que R dispose d'un échantillon où B a joué 1 fois a et 2 fois b. Si deux actions sont également préférées, on pose par convention que les agents choisissent l'action par ordre alphabétique.

2. Donnez un autre exemple d'information a priori pour lequel la convergence s'effectuera sur un équilibre de Nash différent.

3. Si un agent joue constamment la même action et que l'autre joue en fictitious play, est-il vrai que ce joueur "borné" aura à la limite (c'est-à-dire, pour un nombre d'itérations suffisamment grand) un gain cumulé plus grand que celui de l'autre ?