Jallais Bastien

Toussaint Jarod

Letondal Titouan

**Compte rendu SAE – Introduction à l’intelligence artificielle**

1. **Perceptrons multicouches**
2. **MLP - Discrimination**

**Introduction :**

Le perceptron Multicouche (MLP) est un type de neurones artificiels largement utilisé en apprentissage supervisé pour résoudre des problèmes complexes comme la classification et la régression. Il fait partie des modèles d’intelligence artificielle inspirés du fonctionnement du cerveau humain, ou les neurones sont organisés en couches et communiquent entre eux via des connexions pondérées.

**L’algorithme :**

1. **MLP – Un peu de programmation !**

**Un squelette de code :**

Le constructeur de la classe MLP initialise le réseau de neurones avec certaines spécifications comme le nombre de neurones par couche, le taux d’apprentissage, et la fonction de transfert. Il crée les couches et les neurones en définissant les poids, biais, et valeurs initiales. Concernant le constructeur, on a en entrée un tableau (layer) indiquant le nombre de neurones pour chaque couche, un taux d’apprentissage, ainsi qu’une fonction de transfert pour chaque neurone.

La méthode execute() est utilisé pour calculer la sortie du réseau pour une entrée donnée. Elle va mettre à jour les valeurs des neurones à chaque couche, en appliquant les poids, biais, et la fonction de transfert. Cette méthode prend en entrée un tableau représentant les données en entrée qui vont être injecté dans la première couche (par exemple les caractéristiques d’une image), et renvoie un tableau (output) représentant les prédictions/résultats du réseau générés par la dernière couche (par exemple classification).

La méthode backPropagate() permet d’ajuster les poids ainsi que les biais du réseau pour réduire l’erreur entre la sortie calculée et la sortie attendue. Il utilise la rétro propagation pour propager l’erreur des couches de sortie vers les couches cachées. Cette méthode prend en entrée deux tableau input et output représentant respectivement les données d’entrée pour le réseau ainsi que les sorties attendues de ces données d’entrée.

**MLP pseudo code :**

1. **MLP vs KNN ?**

**Un squelette de code :**

1. **Conclusion**
2. **Planification & jeux**
3. **Etudier les algorithmes de planification**

**Introduction :**

Dans ce projet nous allons comparer plusieurs algorithmes de recherche de chemin de la cadre de problème de planification générés à l’aide de la classe Dummy. Voici les algorithmes que nous allons étudier :

* **BFS :** Explore tous les nœuds à une profondeur donnée avant de passer à la suivante.
* **DFS :** Explore un chemin aussi profondément que possible avant de revenir en arrière.
* **UCS :** Etend le nœud ayant le coût cumulatif le plus faible.
* **GFS :** Oriente la recherche en fonction d’une heuristique, sans considérer les coûts passés.
* **A\* :** Combine les avantages de UCS et GFS en utilisant une fonction heuristique f(n) = g(n) + h(n), où g(n) est le coût pour atteindre un nœud et h(n) est l'estimation du coût restant.

Nous analyserons les performances de ces algorithmes sur des problèmes de différentes difficultés : **Très facile**, **Facile**, **Moyen**, et **Difficile**. Chaque niveau de difficulté se distingue par la taille du graphe **n** et le facteur de branchement **k :**

A green and black text on a green background

Description automatically generated

Commençons par la difficulté la plus simple, pour celle-ci on choisit une taille de graphe n très faible (100) et un facteur de branchement également très faible (2), voici ce qu’on obtient :

A screenshot of a computer

Description automatically generated

Pour ce niveau, tous les algorithmes affichent des performances similaires, avec des temps d'exécution très courts et des coûts quasi identiques.

Cela s'explique par la simplicité du problème, où la solution est rapidement accessible, quel que soit l'algorithme. Cependant, **DFS** se distingue légèrement avec un coût sensiblement plus élevé. Cela est dû à la nature de DFS qui explore les chemins en profondeur avant de revenir en arrière, générant ainsi plus de nœuds explorés inutiles.  
En revanche, BFS, UCS et A\* affichent exactement le même coût, car ils parcourent efficacement l'espace de recherche de manière optimale

Augmentons la difficulté, voyons ce qui se passe avec un problème facile (n = 1000 et k = 3) :

A screenshot of a computer

Description automatically generated

Pour un graphe plus grand et un facteur de branchement plus élevé, les résultats restent globalement similaires. Cependant, la difficulté accrue commence à révéler certaines distinctions :

* **DFS** trouve parfois la solution plus rapidement que les autres algorithmes, bien que cela dépende fortement de l'emplacement de la solution dans le graphe. Si la solution se situe profondément dans une branche à gauche, DFS peut y accéder directement sans explorer d'autres branches.
* Le coût de DFS reste néanmoins nettement plus élevé que celui des autres algorithmes, car il génère beaucoup de nœuds inutiles lorsqu'il explore un arbre de manière non optimale.
* BFS, UCS et A\* conservent des coûts similaires, car ils partagent des caractéristiques qui garantissent une exploration optimale.

Augmentons encore la difficulté du problème :

A screenshot of a computer

Description automatically generated

**BFS** commence à afficher un coût légèrement différent de UCS et A\*, ce qui s'explique par sa stratégie d'exploration exhaustive en largeur, qui peut générer plus de nœuds que nécessaire dans certains cas.

**A**\* se montre performant, en équilibrant efficacement le coût cumulé et l'heuristique, ce qui le rend souvent plus rapide que UCS seul.

**GFS**, bien qu'orienté par une heuristique, peut parfois manquer de pertinence dans les problèmes aléatoires générés ici, ce qui explique des performances parfois inférieures.

**DFS** a trouvé une solution en seulement **0,112 secondes**, ce qui est beaucoup plus rapide que les autres algorithmes. Cela s’explique par le fonctionnement de DFS, qui explore un chemin en profondeur avant de revenir en arrière.  
Dans ce cas précis, **la solution était située en profondeur à gauche de l’arbre**, ce qui a permis à DFS de la rencontrer rapidement, par **pure chance**.

Cependant, il est important de noter que cette rapidité est aléatoire et dépend fortement de la position de la solution dans l’arbre. Dans d'autres configurations, DFS pourrait prendre beaucoup plus de temps.

Passons maintenant au problème difficile (n = 50 000 et k = 5) qui permettra de comparer efficacement les différents algorithmes :

A green and black text

Description automatically generated

Ce niveau met véritablement les algorithmes à l'épreuve. Voici les observations principales :

* **UCS et A**\* nécessitent un temps d'exécution environ deux fois plus long que certains autres algorithmes. Cela s'explique par leur recherche exhaustive et leur garantie d'optimalité. UCS, en particulier, explore souvent plus de nœuds, car il ne tient pas compte de l'heuristique.
* **DFS** devient inefficace pour ce niveau, générant un coût très élevé et nécessitant un temps parfois exponentiel en fonction de l'emplacement de la solution.
* **GFS** se montre rapide, mais au détriment de l'optimalité, car il ne prend pas en compte les coûts cumulés. Cela peut expliquer des résultats très variables selon les instances du problème.

Un point clé à noter ici est que, dans un environnement aléatoire comme celui généré par la classe Dummy, les heuristiques utilisées par A\* et GFS ne sont pas forcément très efficaces. En effet, ces heuristiques ne reflètent pas toujours fidèlement la structure sous-jacente du graphe. Cela limite la capacité de ces algorithmes à surpasser BFS ou UCS.

**Etude des résultats**

Ce graphe **montre l’évolution du temps de résolution** des algorithmes en fonction de la difficulté du problème.

**A graph with lines and dots

Description automatically generated**

**C**e graphe montre le **coût utilisé par chaque algorithme en fonction de la difficulté** du problème. On constate ici que UCS et A\* ont exactement le même coût peu importe la difficulté, BFS a quant à lui des coûts très similaire. On voit que GFS est légèrement plus coûteux que ces 3 algorithmes. Le plus important est de constater que DFS nécessite un coût largement supérieur et semble avoir un coût exponentiel en fonction du niveau de difficulté, ce qui le rend inutilisable pour des problèmes trop complexes.

**Que peut-on comprendre ?**

**UCS et A**\* : Ces algorithmes affichent des coûts identiques, car tous deux garantissent une solution optimale. Cependant, A\* tend à être plus rapide grâce à l'utilisation de l'heuristique.

**DFS** : Bien qu'il puisse parfois trouver une solution rapidement, son inefficacité en termes de coût le rend inadapté aux problèmes de grande taille ou aux environnements où l'optimalité est importante.

**GFS** : Rapide dans certains cas, il peut échouer à garantir une solution optimale, surtout lorsque l'heuristique manque de pertinence.

**BFS** : Efficace pour des graphes de petite taille, mais devient rapidement coûteux en termes de mémoire et de calcul pour des graphes plus grands.

**Conclusion**

En conclusion, le choix de l'algorithme dépend fortement du problème à résoudre :

* **DFS** peut être performant lorsque la solution est située en profondeur dans une branche bien définie, mais son coût élevé limite son application.
* **UCS** est adapté aux problèmes nécessitant une solution optimale, mais son temps d'exécution peut être prohibitif pour des graphes très grands.
* **A\*** se distingue par sa capacité à équilibrer optimalité et rapidité, surtout lorsque l'heuristique est bien définie.
* **GFS** est utile lorsque la rapidité prime sur l'optimalité, à condition de disposer d'une heuristique efficace.
* **BFS** reste une approche simple et efficace pour les graphes de petite taille, mais souffre de limitations dès que les dimensions augmentent.

Enfin, dans le cadre de ce projet basé sur des problèmes générés aléatoirement, l'efficacité des algorithmes est limitée par l'absence d'une structure prédéfinie ou d'une heuristique adaptée. Cela explique pourquoi les différences entre A\*, UCS et BFS restent faibles, et pourquoi DFS et GFS ne se démarquent pas significativement.