

Projet de Statistiques, 2021

Deep checker

Arthur Correnson, Igor Martayan, Manon Sourisseau

Introduction

Dans ce projet, nous nous intéressons au jeu de dame. L'objectif est d'utiliser des méthodes d'apprentissage statistiques pour construire une intelligence artificielle capable de jouer au jeu. L'idée maîtresse étant d'apprendre une heuristique évaluant la qualité des coups. Pour se faire, nous utiliserons des données récoltées lors de la simulation d'un grand nombre de parties de jeu de dame. Notons que ce projet est un projet de bout en bout, c'est à dire qu'il couvre [1] la génération du jeu de donnée, [2] le choix des modèles et enfin [3] l'entraînement des modèles et leur mise en pratique.

1 - Génération des données & création d'un simulateur de jeu de dames

Les techniques d'apprentissage statistiques utilisent, comme leur nom l'indique, des données comme matière première. Pour avoir les données nécessaires à l'apprentissage d'une bonne heuristique, nous avons commencé par créer notre propre simulateur de jeu de dames afin de générer un grand nombre de parties. Pour un bon apprentissage, il est primordial de disposer d'un grand nombre de données, sous peine d'obtenir une heuristique qui serait sur-spécialisée pour les quelques scénarios présentés par le jeu de données mais qui se généraliserait très mal aux autres scénarios de partie possibles. Il était donc nécessaire d'avoir un simulateur qui offre des grandes performances, mais également une représentation très compacte des données pour éviter de faire exploser la mémoire nécessaire à la simulation des parties et aux stockages des informations collectées. Nous avons donc fait le choix d'écrire le simulateur en langage C et de définir par la même une représentation binaire astucieuse des données de jeu.

% Décrire comment marche le simulateur

% Décrire les optimisations de code faites

% Décrire la forme des données finales.

Modèles et heuristiques

Présentation de KNN Notre heuristique suit l'algorithme de KNN (K-Nearest Neighbors). Cet algorithme de machine learning est une algorithme d'apprentissage supervisé, qui fonctionne de la manière suivante :

- On calcule la distance entre l'entrée et chaque éléments de la base de donnée
- On sélectionne les K plus proches voisins selon la distance calculés.

- On attribut à notre entrée la moyenne des valeurs des K plus proches voisins.

Adaptation de KNN à notre heuristique Dans une situation de plateau donnée, on attribue un score à chaque coups possible avec l'aide de KNN :

- choix de la distance : Nombre de bytes dans un Xor entre deux coups.
- choix de $K = 4$
- La valeur associée aux K plus proches voisins correspond à la moyenne des scores.

Enfin, on sélectionne le coups possible ayant le plus grand score.

Conclusion