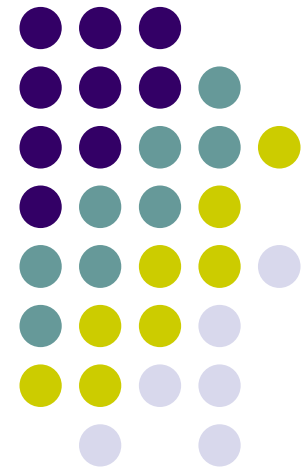


General Game Learning Using Knowledge Transfer

Article écrit par Bikramjit
Banerjee et Peter Stone

Présenté par Guillaume Bernard



Plan de l'exposé



1. Introduction
2. Apprentissage par renforcement
3. Extraction et utilisation des connaissances
4. Résultats expérimentaux
5. Conclusion



Introduction

- Domaine d'application : GGP
 - Écriture de jeux dans un langage commun (GDL)
- Contexte des jeux avec deux adversaires
- Objectif : créer un agent joueur capable d'utiliser des connaissances apprises précédemment dans d'autres jeux

Introduction



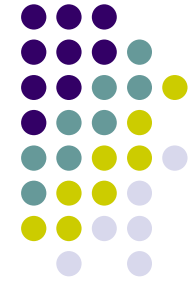
- Contributions des auteurs :
 - Création d'un agent joueur utilisant une méthode basé sur l'apprentissage par renforcement : $TD(\lambda)$
 - Modélisation des connaissances sous forme d'arbres
 - Traitement des connaissances avec un algorithme de recherche en profondeur
 - Réutilisation de ces connaissances dans des parties futures

Plan de l'exposé



1. Introduction
2. Apprentissage par renforcement
3. Extraction et utilisation des connaissances
4. Résultats expérimentaux
5. Conclusion

Apprentissage par renforcement



- Modélisation de l'environnement par un « *Markov Decision Process* »
- But de l'agent : apprendre la suite d'actions permettant de maximiser les gains
- Auteurs utilisent TD(λ) pour l'apprentissage
 - Méthode très rapide à exécuter : permet d'avoir un apprentissage performant

Plan de l'exposé



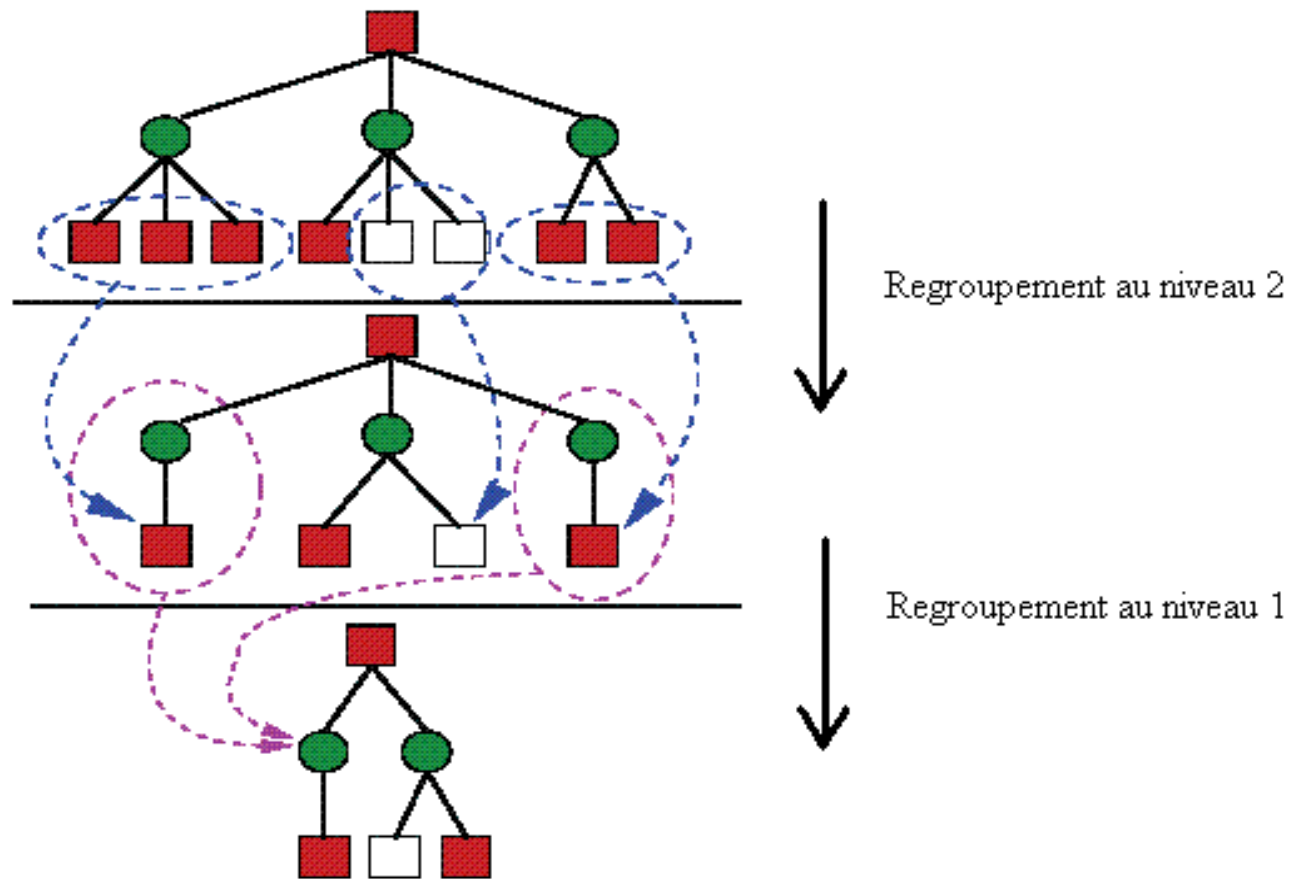
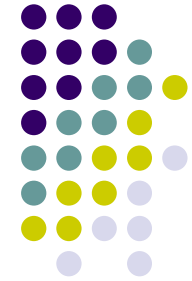
1. Introduction
2. Apprentissage par renforcement
3. **Extraction et utilisation des connaissances**
4. Résultats expérimentaux
5. Conclusion

Extraction et utilisation des connaissances



- Modélisation des connaissances sous forme d'arbre
- Traitement de chaque post-état
- Exploration en appliquant l'algorithme de recherche en profondeur (max = 2)
 - On étiquette chaque nœud (victoire, défaite ...)
 - Regroupement des nœuds identiques
 - Si il y a un terminal dans l'arbre, on mémorise

Extraction et utilisation des connaissances



Extraction et utilisation des connaissances



- Application de l'algorithme tout le long de l'apprentissage
- Calcul de la valeur de chaque connaissance F_i
 - On compare chaque connaissance F_i avec les post-états σ
 - $\text{val}(F_i) = \text{avg}_w \{Q(\sigma) \mid \sigma \text{ équivalent à } F_i\}$ (w est le nombre de fois ou a été visité le post – état)
- Réutilisation des connaissances :
 - Initialisation de chaque $Q(\sigma)$ équivalent à un F_i avec $\text{val}(F_i)$

Plan de l'exposé



1. Introduction
2. Apprentissage par renforcement
3. Extraction et utilisation des connaissances
4. Résultats expérimentaux
5. Conclusion



Résultats expérimentaux

- Comparaison avec deux autres méthodes
 - Agent n'utilisant pas le transfert de connaissances
 - Agent avec une recherche en profondeur et minmax
- Trois joueurs adverses
 - Faible, aléatoire, ϵ -greedy
- Apprentissage sur le jeu du Tic-Tac-Toe contre les trois joueurs

Résultats expérimentaux



- Résultats partagés :
 - Contre agent ϵ -greedy, méthode des auteurs moins efficace
 - Contre les autres, méthode gagnante
- Causes des résultats :
 - Heuristique de l'agent avec minmax faite pour traiter les agents de type ϵ -greedy
 - Méthode des auteurs plus généraliste

Plan de l'exposé



1. Introduction
2. Apprentissage par renforcement
3. Extraction et utilisation des connaissances
4. Résultats expérimentaux
5. Conclusion

État de l'art



- Extraction des connaissances : Fawcett 1993
- Combinaison de la recherche en profondeur et de l'apprentissage par renforcement : Tesauro 1994
- Méthode d'apprentissage par renforcement avec transfert de connaissances : Asgharbeygi 2006



Conclusion

- Gestion du transfert de connaissance
 - Permet d'améliorer les résultats à un niveau plus général
 - Moins efficace qu'une heuristique pour les cas particuliers
- Commentaires personnels
 - Méthode intéressante mais résultats pas assez positifs
 - Trop générale
 - Trouver un juste milieu entre cette méthode et celles plus classiques avec des heuristiques