Baptiste Chemier & Aisi Nguyen

POLYTECH 4ème Année APPRENTISSAGE

Compte rendu de TP

TP1. Planification stochastique & MDP

# Partie IV - Influence des paramètres

## Question 1

Si nous mettons le bruit à 0, nous sommes alors sur que la politique optimale permette à l’agent de traverser le pont. En effet, un bruit égal = 0 correspond à un environnement déterministe. Ainsi, avec un bruit de 0, l’agent n’aura aucune chance de ne pas aller dans la direction souhaitée.

## Question 2

### L’agent suit un chemin risqué pour atteindre l’état absorbant de récompense +1

Il faut mettre une valeur de récompense aux autres négatives afin que l’agent trouve l’état +1 en premier, l’agent ira donc plus rapidement vers la première récompense qu’il trouvera.

### L’agent suit un chemin risqué pour atteindre l’état absorbant de récompense +10

Afin d’atteindre la meilleure des récompenses, par le plus court chemin, il suffit de mettre un bruit à 0 comme cela, il n’aura aucune chance de ne pas suivre la direction voulue. La récompense +10 étant la récompense maximale, l’agent se dirigera vers cette récompense.

### L’agent suit un chemin sûr pour atteindre l’état absorbant de récompense +1

### L’agent évite les états absorbants

Pour que l’agent ne se dirige pas vers les états absorbants, il faut donner une récompense aux autres cases de +25 afin qu’il trouve plus avantageux d’aller vers celles-ci plutôt que sur les états absorbants et où la récompense maximale est de +10.

# Partie V – Modélisation d’un problème sous forme de MDP