TME 2. Programmation Dynamique

Apprentissage par Renforcement

Ce TME a pour objectif d’expérimenter les modèles d’algorithme de programmation dynamique sur un MDP classique de type GridWord.

Dans notre cas, la tâche est une tâche de RL où un agent doit récolter des éléments jaunes dans un labyrinthe 2D et terminer sur une case verte. Nous avions à notre disposition un dizaine de plan de labyrinthe afin de tester nos algorithmes.

Afin de tester les différents algorithmes nous avons effectué 1000 tests sur chaque plan.

# Baseline : Algorithme aléatoire

Afin de pouvoir interpréter les résultats des algorithmes de d’itération de politique ou de valeur nous avons d’abord réalisé les tests avec un algorithme de déplacement aléatoire.

Dans un premier temps on peut s’intéresser aux résultats de l’algorithme sur le plan 0.

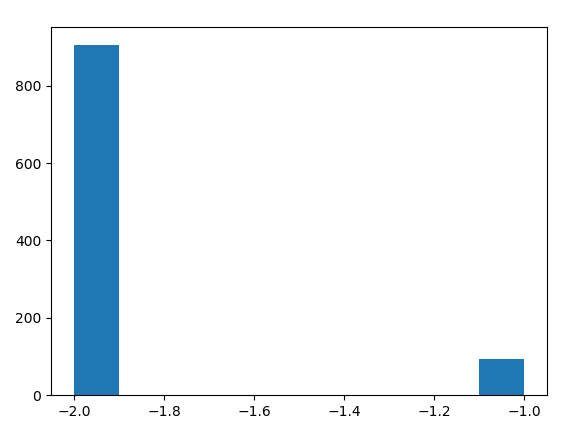
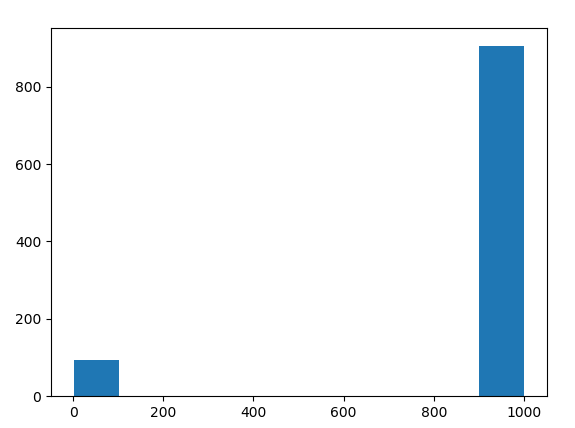


Tableau 1: répartition des scores et mouvements

INTERPRETATION

Dans un second temps on peut observer les résultats sur chaque plan.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **10** |
| **movement** | 907 | 901 | 1001 | 24 | 1001 | 1001 | 1001 | 1001 | 1001 | 1 |
| **score** | -1.91 | -1.80 | -2.00 | 0.98 | -2.00 | -2.00 | -2.00 | -2.00 | -2.00 | -1.00 |

Tableau 2: Résultats des tests pour l'algorithme aléatoire

INTERPRETATION

# Algorithme itératif de politique

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **10** |
| **movement** | 15 | 4 | 10 | 7 | 1001 | 11 | 10 | 523 | 1001 | 1001 |
| **score** | 0.76 | 1.41 | 0.99 | 0.99 | -2.00 | 0.99 | 0.99 | 0.89 | -2.00 | -2.00 |

Tableau 3: Résultats selon les plans pour epsilon 0.001 et gamma 0.5

INTERPRETATION

**Influence du paramètre gamma :**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** |
| **movement** | 28 | 15 | 15 | 16 | 15 | 15 | 5 | 5 | 5 |
| **score** | 0.97 | 0.73 | 0.74 | 0.73 | 0.73 | 0.77 | 0.54 | 0.52 | 0.50 |

Tableau 4: Variation du paramètre gamma pour epsilon 0.001

INTERPRETATION

**Influence du paramètre epsilon :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **0.1** | **0.01** | **0.001** | **0.0001** |
| **movement** | 15.1 | 15.3 | 14.4 | 14.6 |
| **score** | 0.718 | 0.736 | 0.737 | 0.732 |

Tableau 5: Variation du paramètre epsilon pour gamma 0.5

INTERPRETATION

**Histogramme de répartition des résultats sur le plan 0 :**

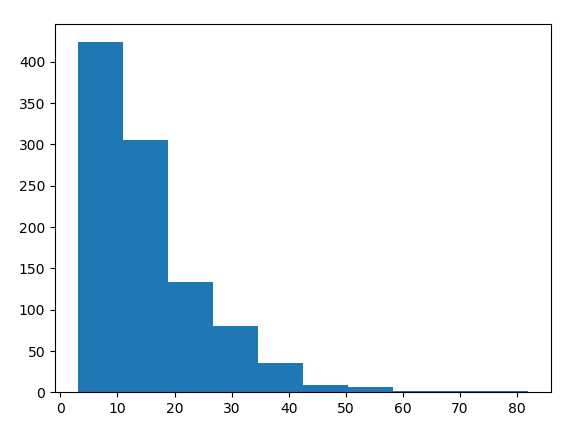
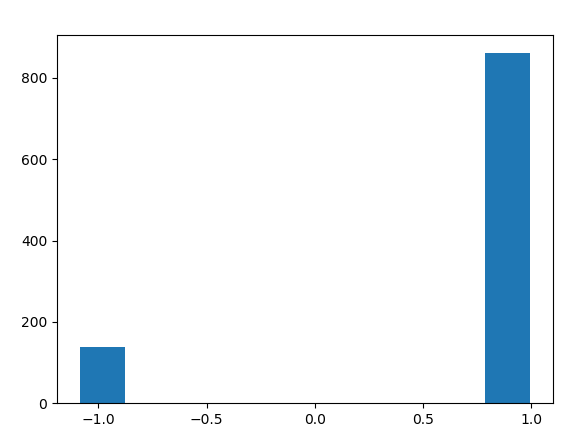


Tableau 1: répartition des scores et mouvements

INTERPRETATION

# Algorithme itératif de valeur

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **10** |
| **movement** | 15 | 5 | 10 | 7 | 1001 | 10 | 10 | 566 | 1001 | 1001 |
| **score** | 0.76 | 1.42 | 0.99 | 0.99 | -2.00 | 0.99 | 0.99 | 0.89 | -2.00 | -2.00 |

Tableau 6: Résultats selon les plans pour epsilon 0.001 et gamma 0.5

INTERPRETATION

**Influence du paramètre gamma :**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** |
| **movement** | 28 | 15 | 15 | 14 | 15 | 15 | 15 | 5 | 5 |
| **score** | 0.97 | 0.75 | 0.73 | 0.77 | 0.72 | 0.76 | 0.75 | 0.53 | 0.61 |

Tableau 7: Variation du paramètre gamma pour epsilon 0.001

INTERPRETATION

**Influence du paramètre epsilon :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **0.1** | **0.01** | **0.001** | **0.0001** |
| **movement** | 15.88 | 14.91 | 15.07 | 15.01 |
| **score** | 0.7551 | 0.7161 | 0.7539 | 0.7300 |

Tableau 8: Variation du paramètre epsilon pour gamma 0.5

INTERPRETATION

**Histogramme de répartition des résultats sur le plan 0 :**

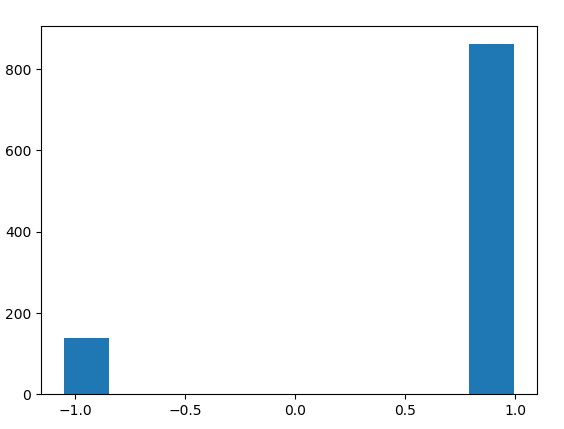
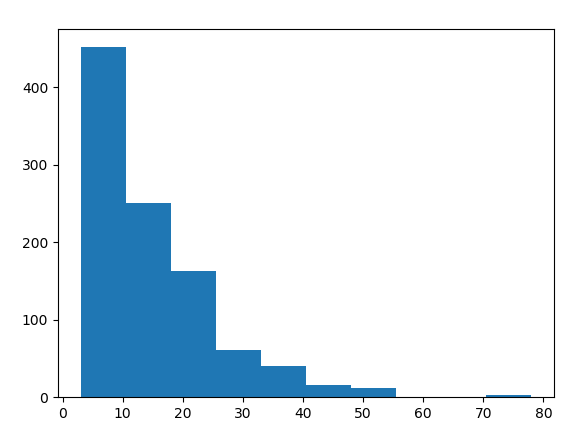


Tableau 1: répartition des scores et mouvements

INTERPRETATION

# Conclusion