	<b>Cycle ingénieur 2<sup>ème</sup> année</b> <b>Examen de rattrapage</b> <i>K. El Ganaoui - J.-P. Forest - Y. Le Nir - P. Loubière - H. Senoussi</i>	
	<i>Matière : IA Applications GSI-SIE</i>	<i>Date : Mai 2021</i>
		<i>Durée : 3h</i>
		<i>Nombre de pages : 2</i>

- Durée : 3h,
- Type : Machine,
- Rendu : Vous rendrez une archive contenant vos codes sources incluant les réponses aux questions sous forme de commentaires. **Rappel** : En python, # donne un commentaire monoligne et les triples guillemets """ donnent des commentaires multi-lignes.  
**L'archive sera nommée : NOM-Prenom-rendu-IA.**
- **Merci d'adresser vos rendus par mail au responsable du cours ET à votre chargé de TD.**
- Tous documents autorisés,
- Aucune question ne peut être posée aux enseignants, posez des hypothèses en cas de doute.

## Apprentissage par renforcement (10 points)

On considère un agent qui se déplace sur le *gridworld* de la figure 1.

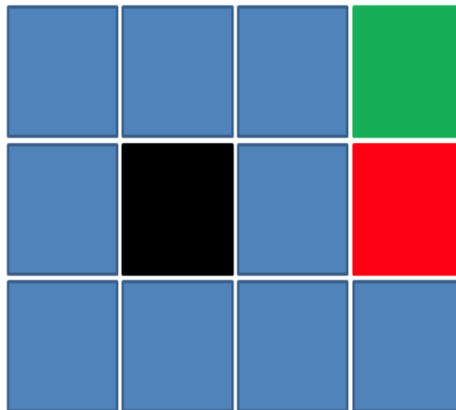


Figure 1 – Gridworld

La case rouge est une case interdite et l'atteindre provoque une punition de  $-1$ .

La case verte est l'objectif et l'atteindre permet à l'agent d'obtenir une récompense de  $+1$ .

La case noire est inaccessible.

L'agent peut se déplacer dans les 4 directions. S'il tente de sortir du *gridworld* ou accéder à la case noire, il ne bouge pas.

Chaque action coûte  $-0.04$ .

Lorsque l'agent prend une direction  $d$ , il réussit avec une probabilité 0.8 et prend la direction située à sa droite ou celle située à gauche avec une probabilité de 0.1.

1. Donnez le processus de décision markovien décrivant ce problème.
2. Écrire un programme Python permettant d'appliquer le Q-learning pour trouver une stratégie optimale.  
On prendra  $\gamma = 0.9$ ,  $\epsilon = 0.2$  et on essaiera plusieurs valeurs de  $\epsilon$  et de  $\gamma$ .
3. Étudiez les conséquences des changements de valeur de  $\gamma$  et de  $\epsilon$ .

## Métaheuristiques (10 points)

Dans cet exercice, nous souhaitons minimiser la fonction suivante :

$$f(x) = 418.9829 * d - \sigma_{i=1}^d x_i * \sin(\sqrt{|x_i|})$$

dans le domaine de définition  $[-500; 500]^d$ ,  $d = 2$ , dimension du problème

1. Afin de réaliser cette minimisation, on discrétise l'espace de recherche à tous les points de coordonnées entières. On évalue la fonction en chaque point.
  - (a) Écrivez un code qui réalise cette opération qui implémente cet algorithme et retourne l'optimum et son score.
  - (b) Combien d'évaluations de la fonction  $f$  sont réalisées (on notera  $n_1$  ce nombre d'évaluations) ?
2. Écrivez un code qui tire uniformément aléatoirement  $n_2$  points dans le domaine de définition et retourne l'optimum et son score.
  - (a) Quel ordre de grandeur de  $n_2$  est judicieux ? Justifiez.
  - (b) Que pouvez-vous conclure sur la qualité de la solution ?
3. Écrivez le code permettant de résoudre le problème avec l'algorithme PSO. On justifiera :
  - le critère d'arrêt choisi ;
  - le nombre de particules ;
  - la méthode choisie pour gérer les problèmes aux limites de l'espace de recherche.
4. Expliquez les effets du choix du voisinage (anneau versus graphe complet) sur le comportement de l'algorithme.
5. Quel critère de comparaison équitable pouvez-vous mettre en œuvre pour comparer les résultats obtenus lors des 3 premières méthodes ? L'appliquer et comparer les résultats, concluez.
6. Nous allons étudier l'impact de passer de  $d = 2$  à  $d = 5$ . Reprenez les questions 1 à 3 avec  $d = 5$ , analysez les résultats, que pouvez-vous conclure ?