

## TITRE EXPERTISE INFORMATIQUE ET SYSTEME D'INFORMATION

### BLOC n° 4 – Etudes et Développement

#### Cahier des Charges de la MSPR « Conception d'une solution applicative d'intelligence Artificielle respectant des contraintes temps-réel »

#### **Activité type A4 : Concevoir et développer des solutions applicatives « apprenantes » en temps-réel**

##### **COMPETENCES EVALUEES :**

Être capable de mettre en œuvre les compétences acquises sur les modules des UE "Mobilité & Systèmes embarqués" et "Intelligence Artificielle" :

- Systèmes embarqués
- Programmation en temps réel
- Web Services
- Apprentissage et réseaux neuronaux (base)
- Résolution de problèmes complexes

##### **Modalités et durée d'évaluation :**

Mise en œuvre d'un projet permettant d'utiliser de façon pratique et en groupe les concepts abordés au sein des UE de ce parcours métier - Ce projet est encadré par des formateurs intervenant dans ces modules.

##### **CONCEPTION ET EVALUATION DES SOLUTIONS D'APPRENTISSAGE**

**Durée de préparation :** 19.5 heures

**Mise en œuvre :** Travail d'équipe constituée de 6 apprenants-candidats (1 seul groupe impair toléré 5 ou 7) divisé en 2 sous-groupes qui confronteront leurs solutions

##### **Résultat attendu – Dossier à produire par équipe :**

- Réalisation d'une solution applicative fonctionnelle sous Windows ou Linux
- Un dossier de synthèse explicitant les choix retenus.

## **PHASE 2 : PRESENTATION ORALE COLLECTIVE + ENTRETIEN COLLECTIF**

**Durée totale par groupe** : 30 mn se décomposant comme suit :

- 10 mn de soutenance orale par l'équipe (présentation d'un défi et des résultats par l'équipe) ;
  - 20 mn d'entretien collectif avec le jury (questionnement complémentaire).
- Objectif : mettre en avant et démontrer que les compétences visées par ce bloc sont bien acquises

**Jury d'évaluation** : I. Chraibi Kaadoud, S.Labasse

## **PLAN DU DOCUMENT**

<b>I – PRESENTATION DE L'ENVIRONNEMENT MALMO .....</b>	<b>3</b>
<b>II – DESCRIPTION DU SYSTEME D'APPRENTISSAGE PAR RENFORCEMENT ....</b>	<b>3</b>
<b>III – DEFIS .....</b>	<b>4</b>
EATING .....	4
CLIFF WALKING .....	4
CATCH THE MOB .....	4
ATTIC / OBSTACLES / VERTICAL .....	4
DEFAULT WORLD.....	5
TRICKY ARENA .....	5
<b>IV – REALISATION ATTENDUE .....</b>	<b>5</b>
<b>V – LIVRABLES .....</b>	<b>6</b>
CODE SOURCE .....	6
DOSSIER DE SYNTHESE .....	6
PRESENTATION .....	6
<b>V – RESSOURCES NON EXHAUSTIVES .....</b>	<b>6</b>

## PRESENTATION DU CONTEXTE DU CAHIER DES CHARGES

### I – PRESENTATION DE L'ENVIRONNEMENT MALMO

La plate-forme Malmo (<https://www.microsoft.com/en-us/research/project/project-malmo/>) est une plate-forme d'expérimentation en intelligence artificielle construite par Microsoft Research sur Minecraft et conçue pour soutenir la recherche fondamentale en intelligence artificielle.

Elle comprend un mod pour la version Java et un code permettant à des agents de détecter et d'agir dans l'environnement Minecraft. Les composants peuvent fonctionner sous Windows, Linux ou Mac OS.

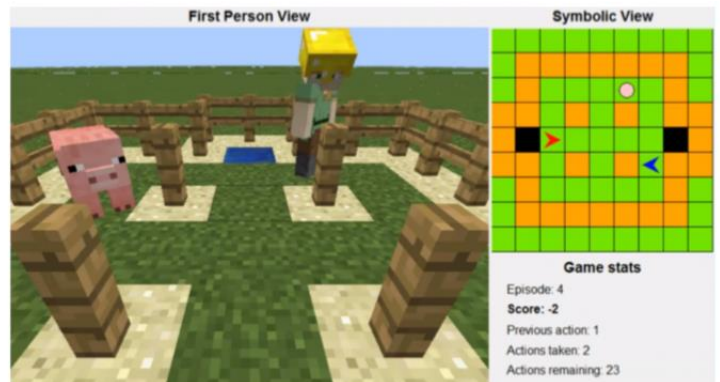
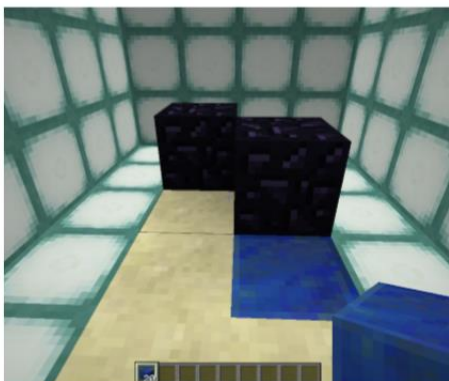


L'API de Malmo permet d'accéder aux actions, aux observations (emplacement, environnements, images vidéo, statistiques de jeu) et à d'autres données générales fournies par Minecraft.

### II – DESCRIPTION DU SYSTEME D'APPRENTISSAGE PAR RENFORCEMENT

MarLÖ (abréviation de Multi-Agent Reinforcement Learning dans MalmÖ) est une API de haut niveau construite sur le projet MalmÖ pour faciliter les expériences d'apprentissage par renforcement pour des agents uniques ou multi-agents.

L'API de Marlo est un wrapper pour Malmo qui fournit un API de plus haut niveau et un environnement plus standardisé et compatible RL pour les études scientifiques.



Le Framework est une extension de Gym d'OpenAI, une boîte à outils permettant de développer et de comparer des algorithmes d'apprentissage par renforcement, fournissant ainsi une plate-forme standard et familière aux scientifiques, développeurs et étudiants.

### III – DEFIS

La description des défis est tirée des sources, un aperçu est disponible sur la page du projet Marlö : <https://github.com/crowdAI/marLo>

#### Eating

This environment is a flat map littered with lots of food items which the agent can pick up. The goal is to pick up only healthy foods - the agent is thus prompted to prefer certain items over others.

Actions : Move forward,backward / Jump / Turn / Strafe / Crouch / Use

Rewards :

2 points for picking up: fish, porkchop, beef, chicken, rabbit, mutton

1 point for picking up: potato, egg, carrot

-1 points for picking up: apple, melon

-2 points for picking up: sugar, cake, cookie, pumpkin pie

#### Cliff Walking

The goal of this mission is for the agent to reach the end of the cliff maze and to pick up the diamond item laying at the end of it. The walking cliff is surrounded by lava and the walking terrain itself has holes to fall through.

Actions : Move / Jumpmove / Strafe / Turn / Movenorth, Moveeast, Movesouth, Movewest / Jumpnorth, Jumpeast, Jumpsouth, Jumpwest / Jump / Look / Use / Jumpuse

Rewards : 100 points for reaching the end goal, -1 points for every step taken, -100 points for falling in lava.

#### Catch The Mob

The goal of this mission is for one or two agents (depending on the playmode) to catch the given target monster by cornering it such that it can no longer escape the block that it is currently in. This can be done via cornering in various angles depending on the map's layout.

Actions : Move / Jumpmove / Strafe / Turn / Movenorth, Moveeast, Movesouth, Movewest / Jumpnorth, Jumpeast, Jumpsouth, Jumpwest / Jump / Look / Use / Jumpuse

Rewards : 1 points for catching the mob, 0.2 points for exiting the current map, -0.02 for each step taken (maximum 50 steps per episode)

#### Attic / Obstacles / Vertical

Attic : The goal of this mission is for one agent to find the goal which is typically hidden within the attic of a mansion, which represents the full extent of the map. There are various obstacles such as stairs, bookshelves and indeed holes filled with lava that the agent must learn to navigate on their path to finding the goal, which is represented by a gold block, a diamond block or a redstone block.

Obstacles : The layout of this mission is that of a series of interconnected rooms made out of stone bricks. Various obstacles are placed in and between these rooms, and doors actioned by levers separate them. The agent's goal is to find the gold/diamond/redstone block in one of the rooms.

**Vertical :** This is a map with many vertical emplacements such as stairs or ladders. The agent's task is to find the gold/diamond/redstone block at the top of the tower by climbing the stairs/ladders suitably.

Actions : Jump / Move / Pitch / Strafe/ Turn / Crouch / Use

Rewards : 1000 points for finding the goal (8000 for Vertical), 20 points for finding a gold, diamond or redstone ore, -1000 points for running out of time

#### **Default World**

This environment thrusts the agent into a normal, random Minecraft world. The seed is generated randomly, therefore the location is not always the same. The goal is to find a gold/diamond/redstone block in time.

Actions : Forward / Backward / Turning

Rewards : 1000 points for finding the goal, -1000 points for running out of time, -10000 points for death

#### **Tricky Arena**

The layout of this mission is that of a flat map littered with redstone, obsidian and ice blocks, as well as water and lava holes. The goal of the agent is to step on as many obsidian blocks as possible before heading towards a redstone block, at which point the mission ends.

Actions : Jump / Move / Pitch / Strafe / Turn / Crouch / Use

Rewards : 100 points with a 1 second delay for each obsidian block touched / -1000 points upon death / -900 points for running out of time / 100 points for leaving the arena (touching stained glass) / -800 points for falling in a water hole / 400 points for finding a redstone block

## **IV – REALISATION ATTENDUE**

La réalisation attendue des apprenants consiste à résoudre le plus efficacement possible au moins un des défis cités en expérimentant plusieurs méthodes sur chaque défi et en les confrontant entre elles.

Pour proposer des solutions réellement distinctes, les apprenants se divisent en deux sous-groupes. Chacun met au point une solution d'apprentissage au sein d'un programme commun. Le programme voit alors s'affronter deux propositions dégageant éventuellement une IA gagnante et une IA perdante.

Les méthodes d'apprentissage à expérimenter et confronter :

- Seront issus du domaine de l'apprentissage par renforcement (Voir la section « RL Theory » en bas de page <https://skymind.ai/wiki/deep-reinforcement-learning>)
- Pourront utiliser des frameworks populaires tel que Keras très répandu en ingénierie, pyTorch très populaire en recherche ou encore Tensorflow, qui est plus bas niveau (il est possible d'utiliser plusieurs frameworks ensemble selon la stratégie choisie)

Dans le benchmark des solutions proposées, les apprenants doivent attacher autant d'importance au qualitatif de la solution qu'au quantitatif (temps et délais de traitements). Ainsi, chaque étape de traitement doit être mesurée.

En plus des sources, le document à remettre s'attardera sur les points suivants :

- Détail des solutions d'apprentissage
- Résultats obtenus avec une analyse critique
- Mesures de performance et méthodes d'optimisation et/ou ordonnancement mises en œuvre

## **V – LIVRABLES**

### **Code source**

Le code source des réalisations est remis sous la forme d'un lien GitLab/GitHub accessible publiquement et spécifié dans le dossier de synthèse.

### **Dossier de synthèse**

Le dossier de synthèse suit le plan suivant :

- Contexte : Rappel du contexte du projet et des objectifs. Défi choisi.
- Sous-Groupe 1 : Solutions proposées, évolutions et améliorations consécutives, difficultés rencontrées/solutions proposées, mesures de performances et optimisations
- Sous-Groupe 2 : idem
- Résultats des « affrontements »
- Discussion : Au regard du ou des défi-s traité-s, quelles conclusions tirer quand aux solutions choisies et leur efficacité

### **Présentation**

Une présentation de type PowerPoint est à remettre au format pdf juste avant la présentation orale le 26 juin.

## **V – RESSOURCES non exhaustives**

<https://marlo.readthedocs.io/en/latest/installation.html>

<https://keras.io/>

Vidéos en français sur l'Apprentissage par renforcement :

<https://www.actuia.com/contribution/thibault-neveu/lapprentissage-par-renforcement/>