

## TD5 – Apprentissage par renforcement

L'objectif de ce TD est d'implémenter en *Python* le modèle de Q-learning pour l'apprentissage par renforcement étudié en cours. Des expérimentations seront ensuite menées pour évaluer la performance du modèle sur des jeux d'essai.

### 1. Présentation des éléments fournis

Un fichier Python (**q\_learning.py**) est fourni, contenant des fonctions à compléter pour rendre le programme opérationnel.

Afin de tester les fonctions développées, le programme principal reprend l'exemple vu en cours sur la recherche de la sortie à partir d'une pièce d'une maison.

### 2. Ecriture du programme

Compléter les fonctions dans le fichier **q\_learning.py** pour rendre le programme fonctionnel. Il est conseillé de suivre l'ordre des fonctions apparaissant dans le fichier, et de les tester à chaque étape.

### 3. Expérimentations

Une fois le programme fonctionnel, réaliser des expérimentations afin d'évaluer l'influence des paramètres de l'algorithme (alpha et gamma). Une complexification du problème sera mise en œuvre en considérant des maison contenant plus de pièces, et de configurations différentes.

### 4. Pour aller plus loin

Adapter l'algorithme précédent pour résoudre le problème des tours de Hanoï (3 tours, 3 anneaux).