

L'ARÈNE

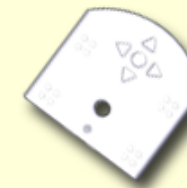
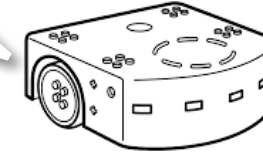


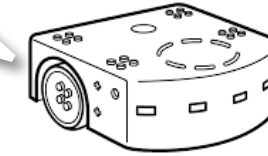
MISSION DE THYMIO ET LE PRINCIPE DE SON APPRENTISSAGE

- Thymio se trouve à l'intérieur d'une arène. Son but est d'explorer cette arène sans toucher les murs. Comme d'habitude, au début, Thymio ne sait rien faire.
- Nous décidons de lui apprendre à effectuer sa mission avec une nouvelle méthode appelée apprentissage par renforcement. C'est une méthode qui fait aussi partie du domaine de l'intelligence artificielle. Nous utiliserons à nouveau un réseau de neurones.
- Dans cette méthode, on ne dit pas à Thymio quelle action il doit faire, parmi les diverses actions possibles. C'est à lui d'en choisir une. Mais comment le robot peut-il savoir quelle action choisir ? Grâce à des récompenses qu'on lui donnera, ou non[1].

[1] Cela évoque le jeu de devinette, "T'es froid, t'es chaud, tu brûle ! ..."

FICHE ÉLÈVE





L'ARÈNE

L'APPRENTISSAGE PAR RENFORCEMENT

- Notre nouvelle manière d'entraîner Thymio consistera à lui indiquer si l'action qu'il a choisi, pour une situation des capteurs donnée, est un bon ou un mauvais choix en lui donnant pour cette action une récompense ou pénalité.
- Le réseau de neurones prend donc connaissance de ces récompenses pour chaque action effectuée. Comment le réseau va-t-il réagir ? Comme le ferait un être humain ou un animal, il va chercher à obtenir le plus de récompenses possible et à fuir les pénalités.
- Le réseau utilise les mathématiques pour progresser. Si on doutait de l'utilité des mathématiques, en voilà un bon exemple d'application. Petit à petit, Thymio va effectuer de plus en plus souvent les "bonnes" actions (celles qui rapportent des récompenses), et de moins en moins souvent les "mauvaises".
- Pour découvrir les bonnes ou mauvaises actions, le robot doit toutes les explorer.
- Donc, de temps en temps, le robot tentera des actions nouvelles, sans tenir compte des récompenses.

Youpi !



Action : avancer
Capteur : rien devant

*Récompense
Accordée !*

Ouille !

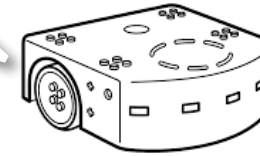


Action : avancer
Capteur : mur devant

*Pénalité
infligée !*

L'ARÈNE

FICHE ÉLÈVE



RÉGLAGES ALPHAI

RAPPEL POUR CONNECTER THYMIO

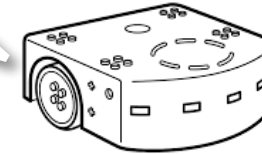
- Connecter Thymio avec une des 2 méthodes suivantes :
 - soit relier le robot à l'ordinateur à l'aide du câble USB
 - soit brancher la clé USB et allumer le robot
 - Lancer le logiciel Thymio Suite et attendre quelques instants.
 - Lancer le logiciel Alphai.
 - Dans Alphai, sélectionner votre robot qui apparaît sur l'écran bleu pour établir la connexion.

RÉGLAGES

- Les réglages sont présents dans un fichier que nous allons charger.
- Dans le menu **Paramètres**, sélectionner l'option **Charger des paramètres d'exemple...**
- Dans le cadre qui apparaît, double cliquer sur l'option **Évitement d'obstacles** (apprentissage par renforcement).
- Au dessous du réseau on remarque deux barres noires de progression.
 - La barre de gauche affiche les récompenses ou pénalités sous forme de nombre.
Une récompense est représentée par un nombre positif, une valeur négative correspond à une pénalité. Les valeurs possibles pour les récompenses sont déjà fixées par Alphai.
 - La barre droite indique le niveau, c'est à dire la moyenne des récompenses précédentes. Il correspond à l'état de l'apprentissage de Thymio.
- Par ailleurs, on laissera les boutons **apprentissage** et **exploration** activés.

L'ARÈNE

FICHE ÉLÈVE



UN PREMIER APPRENTISSAGE

Observer Thymio dans l'arène

Lancer le mode autonome. Thymio se met en mouvement. Son premier déplacement est choisi au hasard.

Observer ses mouvements et ses changements de couleur. Cliquer à nouveau sur le bouton autonome au bout de 20 secondes environ pour arrêter l'apprentissage.

Question 1

Établir des liens entre ce que fait le robot et les couleurs qu'il prend. Noter la réponse :
Poursuivre l'apprentissage pendant 5 à 10 minutes.

Question 2

Constatez-vous une évolution dans le comportement de Thymio ? Si oui, la décrire.
Noter la réponse.



LE RÔLE DE L'EXPLORATION

Cliquer sur le bouton **exploration** pour le désactiver.

Question 3

Constatez-vous une évolution dans le comportement de Thymio ? Si oui, la décrire.

Noter la réponse :

Réinitialiser l'IA avec le bouton **réinitialiser l'IA**.

Cela fait « oublier » à Thymio tout ce qu'il a appris, il recommence un apprentissage de zéro (garder autonome activé, mais exploration désactivé).

Question 4

Constatez-vous des différences entre ce nouvel apprentissage et le précédent ? Si oui, les décrire. Si vous ne constatez pas de différence, recommencer encore un nouvel apprentissage en appuyant sur **réinitialiser l'IA**.

Noter la réponse :

L'ARÈNE



LE COMPORTEMENT DU RÉSEAU DE NEURONES

Observer le comportement du réseau de neurones

Nous allons observer et noter le comportement en détail sur quelques étapes, au début de l'apprentissage.

1. Réinitialiser l'IA avec le bouton **réinitialiser l'IA**. Placer Thymio au milieu de l'arène.

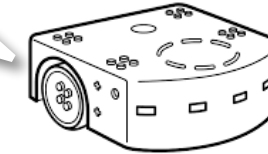
On rappelle que son premier déplacement est choisi au hasard. Pour vous en convaincre vous pouvez cliquer plusieurs fois de suite sur **réinitialiser l'IA** et **autonome**.

2. Compléter la première ligne du tableau. Le petit tiret signifie que les capteurs avant de Thymio ne captent rien car il n'y a rien.

3. Bien regarder Thymio et cliquer sur le bouton **Pas à pas**.

Etape Pas à pas	capteurs avant	reculer à gauche	avancer à gauche	avancer	reculer à droite	avancer à droite	Valeur récompense	Valeur niveau
0	-							
1	-							
2	-							

FICHE ÉLÈVE



Question 5

Quel mouvement Thymio a-t-il fait ? Dans la ligne du tableau que vous venez de remplir, relevez la valeur correspondante à ce mouvement. Comparer ce nombre aux valeurs des autres actions. Que constate-t-on ?

Noter la réponse :

Question 6

On observe aussi que le robot a reçu une récompense. Cette récompense semble-t-elle en accord avec le but de la mission ?

Expliquer :

Une fois la première récompense décernée, les valeurs de sortie sont recalculées par le réseau de neurones.

Compléter la 2ème ligne du tableau et devinez alors quel sera le prochain mouvement de Thymio.

Faire quelques clics supplémentaires sur le bouton **pas à pas** en regardant évoluer les récompenses et le niveau.

Question 7

Comment évolue le niveau quand le robot reçoit une récompense ou au contraire une pénalité ? Que représente le niveau ?

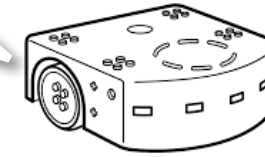
Noter la réponse :

A présent, appuyer sur **autonome** pour laisser Thymio continuer son apprentissage.

Question 8

Comment évolue le niveau au cours de l'apprentissage ?

Pourquoi ? Expliquer :



L'ARÈNE



INFLUENCE DE LA PÉNALITÉ SUR LE COMPORTEMENT APPRIS

Si on résume les différentes valeurs apparues dans la barre de progression, on a relevé :

- 100 : Quand Thymio avance tout droit sans obstacle devant, c'est la valeur la plus élevée
- 55 : Quand Thymio tourne sans obstacle devant.
- -50 : Thymio effectue une des multiples "mauvaises" actions comme avancer contre un mur, reculer quand il n'y a rien devant ...

Nous pouvons modifier la valeur de la pénalité.

- ouvrir l'onglet récompense et mettre la pénalité à une petite valeur, par exemple 0 : réinitialiser l'IA et relancer l'apprentissage pendant quelques minutes.

Observer le comportement de Thymio, est-il plus audacieux, plus prudent ? tape-t-il plus ou moins souvent les murs ?

- mettre une plus grande pénalité, 1,5 par exemple. Réinitialiser à nouveau l'IA et relancer l'apprentissage pendant quelques minutes.

Même question que précédemment.

Question

Résumer comment le comportement de Thymio évolue quand on modifie la valeur de la pénalité.

Notez la réponse :

RÉPONSES



THYMIO LE CRAINTIF

Nom:.....

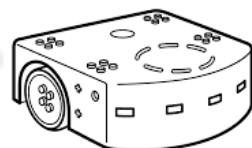
Classe.....

Question 1 : Établir des liens entre ce que fait le robot et les couleurs qu'il prend.

Question 2 : Constatez-vous une évolution dans le comportement de Thymio ?

Question 3 : Constatez-vous des différences entre ce nouvel apprentissage et le précédent ? Si oui, les décrire.

ÉLÈVE



RÉPONSES



THYMIO LE CRAINTIF

Nom:.....

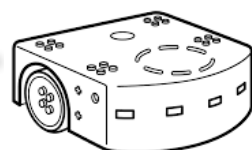
Classe:.....

Question 4 : Constatez-vous des différences entre ce nouvel apprentissage et le précédent ?

Question 5 : Quel mouvement Thymio a-t-il fait ? Dans la ligne du tableau que vous venez de remplir, relevez la valeur correspondante à ce mouvement. Comparer ce nombre aux valeurs des autres actions. Que constate-t-on ?

Question 6 : On observe aussi que le robot a reçu une récompense. Cette récompense semble-t-elle en accord avec le but de la mission ?
Compléter la 2ème ligne du tableau et devinez alors quel sera le prochain mouvement de Thymio.

ÉLÈVE



RÉPONSES



THYMIO LE CRAINTIF

Nom:.....

Classe:.....

Question 7 : Comment évolue le niveau quand le robot reçoit une récompense ou au contraire une pénalité ? Que représente le niveau ?

Question 8 : Comment évolue le niveau au cours de l'apprentissage ? Pourquoi ?

Question 9 : résumer comment le comportement de Thymio évolue quand on modifie la valeur de la pénalité.

ÉLÈVE

