

TaVAR : Une Table lumineuse pour Vulgariser l'Apprentissage par Renforcement¹

B. Girard, L. Gaillot, L. Laval, L. Le Peutit, E. Massi, F. Sangaré, I. Tuzun
ISIR, Sorbonne Université, CNRS, Paris, France.

Nous présentons ici la démarche de création d'une table lumineuse destinée à vulgariser le fonctionnement d'un apprentissage par renforcement par un robot dans une tâche de navigation. Ceci, en révélant les rouages internes de l'algorithme à la surface de la table, sous les roues du robot. Les caractéristiques visées étaient : visualisation des variables internes d'un algorithme de Q-Learning et mise en évidence de l'accélération de l'apprentissage par l'utilisation de réactivations d'épisodes mémorisés ("experience replay", Lin, 1992) ; affichage de la carte utilisée par le robot sur la surface de locomotion ; interactivité permettant au présentateur d'arrêter et relancer le robot à tout instant, de changer les données affichées et de contrôler le déclenchement des réactivations, le tout via des commandes au clavier ; faible coût, facilité de démontage et de transport (pour ne pas limiter son utilisation à une salle du laboratoire), accessibilité (visible par des enfants et utilisant des couleurs distinguables par des personnes souffrant des divers types de dyschromatopsie). La première version de TaVAR a été réalisée dans le cadre du projet industriel de la spécialité robotique de Polytech Sorbonne.

La création de cet objet est partie d'un constat fait par le passé, durant la préparation de démonstrations pour la Fête de la Science : il est difficile de vulgariser l'intelligence artificielle robotique, et l'apprentissage par renforcement en particulier. En effet, durant l'apprentissage, l'important de ce qui se passe dans la machine ne se voit pas. Les processus de calcul, les modifications des informations mémorisées par le système, si elles peuvent se traduire par des modifications progressives du comportement du robot, ne peuvent être perçues directement sans la conception d'une visualisation dédiée. Cette visualisation n'est pas nécessairement celle dont le scientifique a besoin lorsqu'il réalise une telle expérience, elle doit donc être réalisée spécifiquement pour le grand public.

Le dispositif TaVAR est composé d'une table en profilés d'aluminium et en planches de bois contreplaqué. Son plateau est une plaque de plexiglas (180cmx160cm) sur laquelle a été appliqué un film dépoli translucide. Un vidéo-projecteur à très courte focale est fixé sous la table de manière à projeter des images à travers le plateau translucide. Le robot est un turtlebot 3 burger utilisant ROS (Robotic Operating System), son contrôle est déporté sur un ordinateur avec lequel il communique via une borne wifi dédiée. Il se localise sur la table à l'aide de l'algorithme de SLAM (Simultaneous Localization And Mapping) classique GMapping (Grisetti *et al.*, 2007). Un ordinateur raspberry également sous ROS contrôle l'affichage du vidéo-projecteur. Il communique via le wifi avec

¹ This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101070381 (PILLAR project).

l'ordinateur principal : il reçoit des messages décrivant les valeurs des actions dans chaque état de la carte (les Q-valeurs) apprises par le robot, la trajectoire estimée du robot, et, lorsqu'une réactivation est demandée, ses états de départ et d'arrivée. Il peut envoyer des messages de commande pour arrêter le robot, le relancer, et jouer une



réactivation choisie au hasard parmi les transitions préalablement expérimentées.

Le dispositif a été utilisé pour la première fois pour la Fête de la Science 2022 dans les locaux de l'ISIR. Il a fonctionné 5 heures durant sans connaître de dysfonctionnement. Les retours d'expérience sont les suivants :

- Malgré la très courte focale du projecteur, la hauteur nécessaire du plateau est de 90 cm, auxquels s'ajoutent les 20cm de la paroi de l'arène, qui doit être perceptible par le capteur LIDAR fixé au sommet du robot. In fine, la table est un peu trop haute pour pouvoir être observée par tous les enfants de 6 ans et plus. On peut espérer résoudre ce problème en ajoutant une petite estrade.
- Une étude de la déformation mécanique du plateau nous a contraint à utiliser un plateau d'un seul tenant, plutôt qu'en deux morceaux : sous le poids du robot, un dénivelé de l'ordre du centimètre aurait pu apparaître entre les deux plaques lorsque le robot tenterait de passer de l'une à l'autre au milieu de la table, ce qui est trop élevé pour un turtlebot 3. L'objectif de transportabilité n'est pas pleinement atteint, le plateau pesant une vingtaine de kilogrammes.

Il sera intéressant, à l'avenir, de chercher à mesurer la compréhension que les spectateur auront eu de l'algorithme présenté, afin d'estimer la valeur pédagogique du dispositif.

Bibliographie

Grisetti, G., Stachniss, C., and Burgard, W. (2007). Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Trans. Rob.* 23, 34–46.

Lin, L. J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine learning*, 8(3), 293-321.