

# Sujet 1

## Auto-organisation pour l'apprentissage du support dans le cadre de l'approximation de fonction et application à l'apprentissage par renforcement

Proposé par : Jérémy Fix, Alain Dutech

Equipe : Biscuit

### Informations générales

Encadrants Jérémy Fix, Alain Dutech  
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy  
Téléphone 03 54 95 86 36  
Email jeremy.fix@loria.fr, alain.dutech@loria.fr  
Bureau C 044

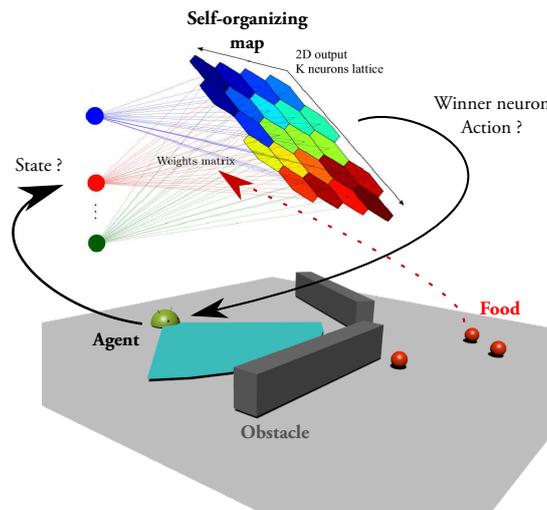


FIGURE 1 – Dans cette application virtuelle de l'apprentissage par renforcement, une carte de Kohonen est utilisée pour estimer la Fonction Valeur. Comment apprendre cette carte? Comment décider de l'action? (L'image de la SOM est tirée de [1])

### Motivations

Si certains développements récents en Intelligence Artificielle ont permis des réalisations importantes en Apprentissage par Renforcement (battre un maître au jeu de GO par exemple), ces travaux s'appuient largement sur de l'apprentissage *supervisé*. C'est cependant sur d'autres types d'apprentissages qu'il conviendrait de se pencher pour progresser vers des machines artificielles vraiment autonomes. Ces thématiques, notamment l'apprentissage *non-supervisé* et l'*auto-organisation*, sont au centre des préoccupations de l'équipe Biscuit. Et nous voulons tout particulièrement guider ces méthodes d'adaptation en utilisant le cadre de l'Apprentissage par Renforcement, par exemple pour concevoir l'architecture décisionnelle d'agents intelligents.

Les approches supervisées d'apprentissage par renforcement nécessitent qu'un signal de supervision non nul soit présent pour que l'apprentissage ait lieu. Or, dans bien des cas, par nature de l'apprentissage par renforcement, le signal de supervision est très rare puisqu'il dépend essentiellement des instants où une récompense ou punition est obtenue. Pour bien des exemples où l'agent doit construire une politique pour atteindre un état

final (e.g. gagner à un jeu), le signal vraiment informatif n'est disponible qu'à l'état final du jeu. Il en vient que beaucoup d'information sur la nature des états rencontrés pendant l'exploration est perdue. Au contraire, les approches par apprentissage non supervisé vise à représenter la distribution des états rencontrés; cet apprentissage est donc agnostique du but mais capture les propriétés des états visités.

## Sujet

On se propose dans ce sujet d'explorer l'utilisation des cartes auto-organisatrices [5] pour l'apprentissage de fonction de valeur utilisée dans les algorithmes *model-free* en apprentissage par renforcement [4]. Les cartes auto-organisatrices de Kohonen sont un algorithme de quantification vectorielle qui, à la différence des K-means, dispose d'une topologie : des unités proches dans la carte ont des prototypes proches. On peut alors espérer disposer de bonnes capacités de généralisation lorsqu'il est question d'approximer une fonction à partir du support quantifié. Par ailleurs, des extensions récentes [3], moins sensibles à la densité des échantillons de la distribution à quantifier, permettent d'envisager de biaiser l'apprentissage du support et donc la distribution des prototypes, vers les parties du support où il est difficile d'apprendre la fonction à approximer. Le travail débutera par une étude de ces idées sur des fonctions "jouet" et on appliquera ensuite ces idées sur des problèmes d'apprentissage par renforcement avec des approches *model-free* nécessitant d'approximer une fonction de valeur.

## Cadre du travail

Le stage s'effectuera au Loria, un poste de travail sera fourni. Des bonnes notions d'apprentissage automatique sont attendues, notamment en ce qui concerne la quantification vectorielle et l'auto-organisation. Les développements se feront sous Linux en C/C++ ou python. Le travail pourra être réalisé en tout ou partie sur le site de Nancy ou sur le site de Metz (campus de CentraleSupélec).

## Références

- [1] D. Jankowski and M. Amanowicz. *Intrusion detection in software defined networks with self-organized maps* in Journal of Telecommunications and Information Technology, 2015.
- [2] T. Kohonen *Self-organized formation of topologically correct feature maps* in Biological Cybernetics, vol. 43(1), pp. 59–69, 1982.
- [3] N. Rougier and Y. Boniface *Dynamic self-organizing maps* in Neurocomputing, vol. 74(11), pp. 1840–1847, 2011.
- [4] R. Sutton and A. Barto, *Reinforcement Learning : An introduction*, 2nd edition, MIT Press, Cambridge, MA, 2018.