

Résumé

Le cortex cérébral apparaît dans de nombreux travaux comme une architecture de modules autonomes connectés rétroactivement, interagissant autour des informations sensorielles ou de plus haut niveau traitées par les différentes aires, implémentant des tâches d'apprentissage extrêmement sophistiquées. Cette notion bio-inspirée d'architecture modulaire présente un intérêt computationnel dans la recherche de nouveaux paradigmes d'apprentissage. Il s'agit en effet de systèmes complexes, propices à faire émerger des mécanismes d'apprentissage dus à l'interaction entre les modules. Dans cette démarche exploratoire, cette thèse propose d'étudier la mise en relation de cartes auto-organisatrices en architectures modulaires non-hiérarchiques, c'est-à-dire présentant des rétroactions entre les modules. Les cartes auto-organisatrices sont un algorithme d'apprentissage non-supervisé permettant de représenter de façon ordonnée un espace d'entrée en faible dimension, et qui s'inspire de l'organisation présente dans les aires corticales. Par la simplicité de leurs règles de mise à jour et leur capacité de générer une représentation positionnelle d'une entrée, elles nous apparaissent comme des candidates naturelles à la conception d'une architecture modulaire. Nous développons et étudions dans ces travaux un modèle modifié de cartes auto-organisatrices permettant de les associer au sein d'une architecture non-hiérarchique. Nous appelons ce modèle CxSOM, pour *Consensus Driven Multi-SOM*. La thèse constitue ensuite une analyse expérimentale des mécanismes d'organisation et d'apprentissage émergeant de l'association des modules. Nous nous concentrons sur la mise en évidence de mécanismes de mémoire associative entre modalités ; l'objectif est de pouvoir apprendre une représentation de plusieurs espaces d'entrées au sein de l'architecture et d'extraire des relations existant entre ces entrées. Pour analyser ces mécanismes, nous mettons l'accent sur une méthode de représentation des réponses de l'architecture, et proposons des outils de visualisation et de mesure de l'apprentissage. Grâce à ce cadre expérimental, nous avons pu mettre en lumière des comportements d'apprentissage associatifs spécifiques au modèle d'architecture et des perspectives d'étude possibles. En particulier, le modèle présente un comportement de prédiction d'entrée rendu possible par les rétroactions et l'apprentissage associatif des entrées au sein de l'architecture. La proposition du modèle CxSOM et l'analyse des comportements sur des architectures simples nous permet d'élaborer une base de travail, vers la conception d'architectures non-hiérarchiques comportant de nombreuses cartes.

Mots-clés: Cartes auto-organisatrices, architecture modulaire, mécanismes d'apprentissage

Abstract

The cerebral cortex appears in many models as an architecture of autonomous modules connected retroactively, interacting around sensory and higher-level information and implementing extremely sophisticated learning tasks. This bio-inspired concept of a modular architecture has also a computational interest in the search for new learning paradigms, especially for the design of autonomous learning networks that evolve over time. They are indeed complex systems, which can conduct to the emergence of learning mechanisms due to the interaction between the modules. This thesis proposes to explore how to connect self-organizing maps in non-hierarchical modular architectures, that involves retroactions between the modules. Self-organizing maps are an unsupervised learning algorithm that creates an ordered representation of a low-dimensional input space and is inspired by the organization present in cortical areas. Due to the simplicity of their update rules and their ability to map an input to a position, they appear to us as natural candidates for the design of a modular architecture. In this work, we introduce a modified self-organizing map model and an interface method for associating them within a non-hierarchical architecture. We call this model CxSOM, standing for Consensus-Driven Multi-SOM. The thesis constitutes an experimental analysis of the organization and learning mechanisms emerging from the association of those modules. We focus on highlighting associative memory mechanisms between several modalities ; the goal for the model is to learn a representation of multiple input spaces within the architecture, as well as the relationships existing between these inputs. To analyze these mechanisms, we focus on method to represent the architecture's responses, and propose visualization and learning measurement tools. Through this analysis framework, we were able to highlight specific learning behaviors of the architecture model and possible study prospects. Particular, we show an input prediction behavior that emerges from the interactions between the maps of the architecture. The construction of the CxSOM model and the analysis of its behavior on simple architectures stand as groundwork towards the design of a non-hierarchical architecture model with numerous maps.

Keywords: Self-organizing maps, modular architecture, learning mechanisms