

Sujet de Master 2 (Recherche)

# Optimisation d'apprentissage et coordination pour l'apprentissage distribué

(Mars - Août 2010)

---

*Lieu :* UMR MIA (Mathématiques et Informatique Appliquées) à AgroParisTech  
16, rue Claude Bernard, 75005 Paris

*Direction :* Antoine CORNUÉJOLS (AgroParisTech) ([antoine@lri.fr](mailto:antoine@lri.fr))

---

## 1. Sujet

De plus en plus d'applications soulèvent le problème d'un apprentissage incrémental à partir de données fournies en séquence. Les systèmes d'apprentissage sont alors susceptibles d'être sensible à l'ordre des données fournies. C'est le cas en particulier d'apprentissages dits hiérarchiques ou cumulatifs, par exemple par induction de programmes logiques (ILP) ou utilisant des Réseaux de neurones à réservoir (*Reservoir Neural Networks*). Se posent alors deux questions :

- Sous quelles conditions un système apprenant est-il indépendant de l'ordre ?
- Si il y a sensibilité à l'ordre, comment optimiser cet ordre ? (voir par exemple "Curriculum Learning" à ICML-09).

La première question est très fondamentale car on peut espérer associer à une invariance par l'ordre une quantité conservée comme en physique. La seconde question devient évidemment cruciale pour les apprentissages incrémentaux et par transfert.

Le travail envisagé concerne donc les deux volets :

- Un volet plutôt théorique sur l'étude des conditions d'invariance par rapport à l'ordre des données d'apprentissage.
- Un volet plus expérimental sur des systèmes d'apprentissage très simples, donc faciles à contrôler, et sur des systèmes plus complexes comme les réseaux de neurones à réservoir.

Par ailleurs, l'apprentissage incrémental peut être vu comme un apprentissage mettant en jeu plusieurs apprenants s'échangeant des messages le long de la ligne temporelle. On peut généraliser au cas d'un apprentissage par des agents locaux et limités (apprentissage distribué). On explorera également ce point de vue.

Ce stage s'inscrit dans un axe prioritaire des équipes d'accueil et bénéficiera d'un plein soutien. Il est rémunéré.

## 2. Environnement de la thèse

La thèse se déroulera entre le laboratoire d'informatique de l'UMR MIA (Mathématiques et Informatique Appliquées) d'AgroParisTech et l'équipe TAO du LRI à l'Université Paris-Sud.