

---

# Apprentissage séquentiel avec graphes de feedback

Frédéric Koriche\*<sup>1</sup>

<sup>1</sup>Centre de Recherche en Informatique de Lens (CRIL) – Université d’Artois, CNRS : UMR8188 – Rue Jean Souvraz - SP 18 62307 LENS CEDEX, France

## Résumé

Parmi les nombreux modèles de l’apprentissage automatique, l’apprentissage séquentiel est probablement l’un des plus généraux, avec des applications diverses s’étant depuis la classification à partir de flux de données, jusqu’à la prédiction de préférences dans les systèmes de recommandation. L’apprentissage séquentiel est profondément relié à la théorie des jeux ; il s’agit essentiellement d’un jeu répétitif à somme nulle entre deux joueurs: le prédicteur et l’apprenant. Durant chaque tour de jeu, le prédicteur choisit une action et, simultanément, l’environnement choisit une fonction de perte. Le prédicteur subit alors une perte sur son action jouée et, selon le feedback qui lui a été transmis, révisé sa stratégie de choix d’actions. Dans cet exposé, nous étudierons le modèle d’apprentissage séquentiel (stochastique ou non-stochastique) dans lequel le feedback est un graphe (potentiellement orienté, et potentiellement pondéré) ; à partir de ce graphe, le prédicteur peut faire des inférences sur les actions qu’il aurait pu jouer à la fin de chaque tour. Ce modèle est suffisamment flexible pour englober d’un côté l’apprentissage supervisé (comme par exemple la classification ou la régression) et de l’autre l’apprentissage par renforcement (plus particulièrement, les bandits multi-bras). Dans ce contexte, la convergence des algorithmes mesurée par la notion de regret, fait appel à des outils mathématiques provenant à la fois de l’optimisation séquentielle et de la théorie des graphes. Sur ce point, nous concluons l’exposé par diverses questions ouvertes, susceptibles d’intéresser à la fois les chercheurs en optimisation et les théoriciens des graphes.

---

\*Intervenant