

Apprentissage par renforcement à base de modèles pour le contrôle de processus de décision semi-markoviens déterministes par morceaux, partiellement observables

Orlane Le Quellenec

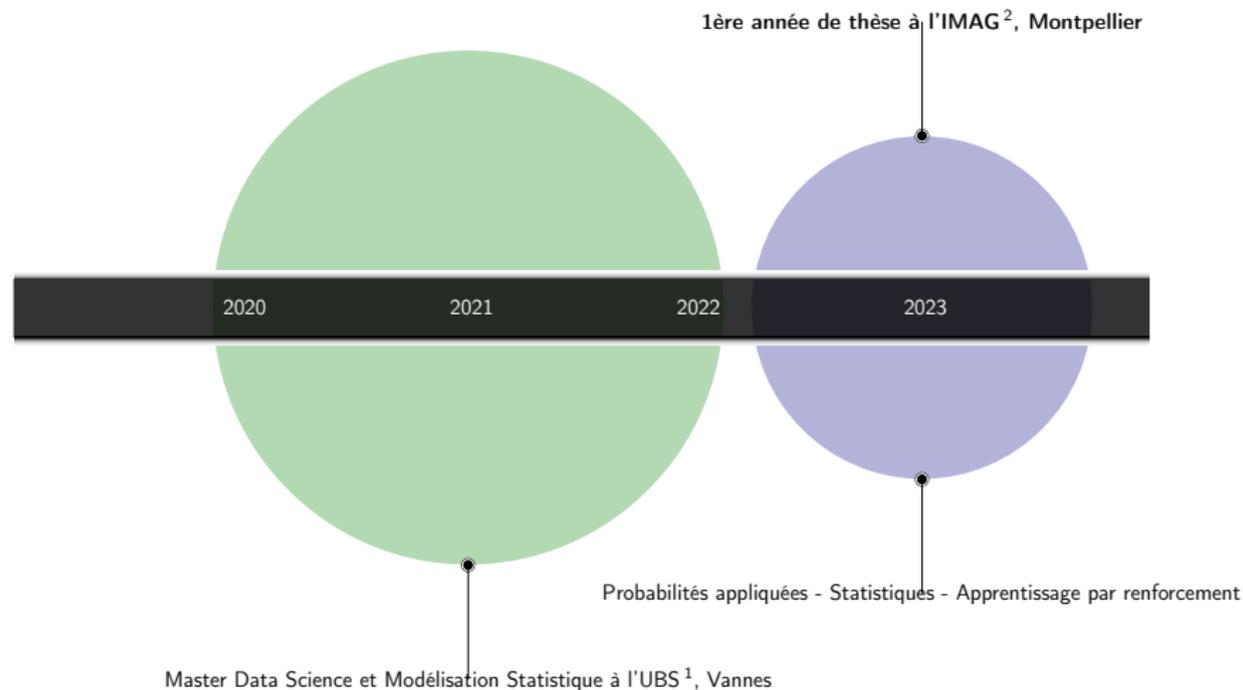
Sous la direction de :

Benoîte de Saporta ¹, Alice Cleynen ² et Régis Sabbadin ³

¹IMAG, Université de Montpellier

²CNRS Montpellier

³INRAE-MIAT, Université de Toulouse



1. Université Bretagne-Sud
2. Institut Montpelliérain Alexander Grothendieck

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de processus de décision semi-markoviens déterministes par morceaux, partiellement observés.

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de processus de décision semi-markoviens déterministes par morceaux, partiellement observés.

Definition

L'apprentissage par renforcement est un procédé d'apprentissage automatique consistant, pour un système autonome, à apprendre les actions à réaliser, à partir d'expériences, de façon à optimiser une récompense quantitative au cours du temps.

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de **processus de décision semi-markoviens**³ déterministes par morceaux, **partiellement observés** (POsMDP).

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de **processus de décision semi-markoviens**³ déterministes par morceaux, **partiellement observés** (POsMDP).

Definition

Un POMDP^a est une généralisation d'un MDP^b dont l'agent n'a qu'une information partielle de l'état courant.

3. ne respecte pas la propriété de Markov

Le sujet

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de **processus de décision semi-markoviens**³ déterministes par morceaux, **partiellement observés** (POsMDP).

Definition

Un POMDP^a est une généralisation d'un MDP^b dont l'agent n'a qu'une information partielle de l'état courant.

- a. processus de décision markovien partiellement observé
- b. processus de décision markovien

Definition

Un MDP^a est un modèle mathématique de décisions séquentielles qui satisfait la propriété de Markov^b.

- a. processus de décision markovien
- b. l'état actuel ne dépend pas du passé

3. ne respecte pas la propriété de Markov

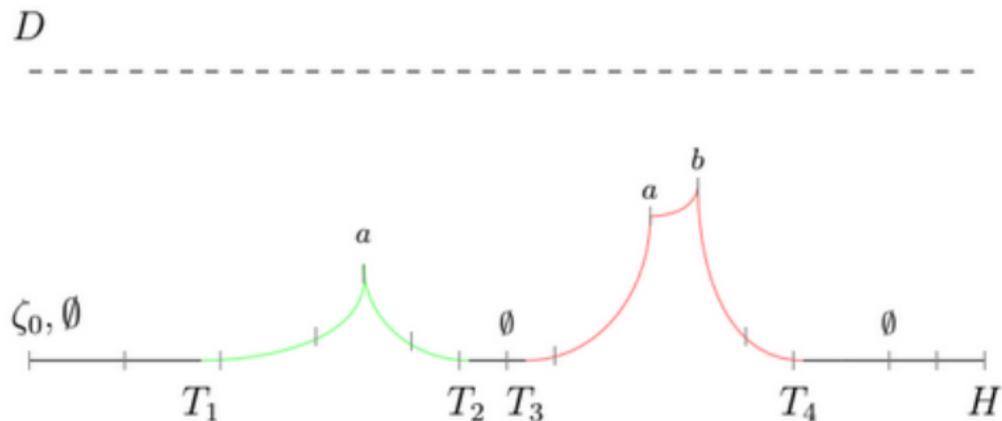


Figure – Exemple de trajectoire contrôlée

Sujet : Apprentissage par renforcement à base de modèles pour le contrôle de processus de décision semi-markoviens déterministes par morceaux, partiellement observés (POsMDP).

Objectif : Proposer un cadre d'inférence et d'optimisation conjointe pour un processus semi-markovien déterministe par morceaux, partiellement observé et à modèle mal connu.

Contexte : Application au domaine médical sur des problématiques de suivi et traitement de cancers.

- 1ère année de thèse à Montpellier
- Domaines de recherche :
 - Probabilités appliquées
 - Statistiques
 - Apprentissage par renforcement
- Processus de décision markovien partiellement observables
- Problématiques de suivi et de traitement de cancers

orlane.lequellenec@umontpellier.fr

INRAE anr®

