

Aurélie Beynier  
Professeure des Universités  
LIP6, Sorbonne Université  
4 place Jussieu  
75005 Paris

aurelie.beynier@lip6.fr

Paris, le 10 novembre 2025

*Rapport sur le mémoire de thèse intitulé*  
**Interprétabilité, Arbres de Décision et Prise de Décision Séquentielle**  
présenté par **Hector Kohler**  
en vue de l'obtention du grade de docteur de  
l'Université de Lille  
Spécialité : Informatique et Applications

Les travaux de thèse présentés par Hector Kohler portent sur l'interprétabilité en prise de décision séquentielle. Ce travail de thèse s'inscrit dans les domaines de l'apprentissage par renforcement et de l'interprétabilité en Intelligence Artificielle. Hector Kohler s'intéresse plus particulièrement aux arbres de décision comme représentation interprétable des politiques dans des problèmes de classification supervisée et de prise de décision séquentielle. Hector Kohler étudie des approches de résolution permettant d'apprendre directement des représentations interprétables des stratégies de décision. Hector Kohler analyse en détail la difficulté de ces problèmes d'apprentissage et propose une nouvelle approche permettant d'apprendre directement des arbres de décision interprétables dans le cadre de la classification supervisée. Hector Kohler propose également une étude expérimentale de l'interprétabilité des modèles étudiés dans le domaine. Les contributions de la thèse sont ainsi à la fois théoriques, algorithmiques et expérimentales.

Cette thèse a été préparée au Centre de Recherche en Informatique, Signal et Automatique de Lille (CRISTAL), sous la direction de Philippe Preux et Riad Akrouf.

Le mémoire présenté par Hector Kohler comporte 166 pages et se compose de 13 chapitres organisés en 3 parties auxquels s'ajoutent un chapitre de conclusion et 3 annexes. L'ensemble du manuscrit est rédigé en anglais.

Dans ce rapport, je commencerai par décrire et commenter chaque partie du document puis, je donnerai un avis global sur le manuscrit.

Le premier chapitre introduit de manière synthétique le concept d'interprétabilité, le motive par différents cadres applicatifs de prise de décision, et replace la question de l'interprétabilité dans son contexte historique. Dans le cadre de l'apprentissage machine, ce chapitre motive le recours à des arbres de décision plutôt qu'à des réseaux de neurones afin de favoriser l'interprétabilité des décisions.

Le second chapitre du manuscrit présente les deux grandes familles d'approches visant l'interprétabilité en apprentissage par renforcement : l'apprentissage direct de modèles interprétables ou bien l'analyse post-hoc de modèles non interprétables afin d'apprendre une représentation interprétable. Dans ce chapitre, Hector Kohler met en évidence le compromis entre performance et interprétabilité sous-jacent à chaque classe de modèles. Hector Kohler termine ce chapitre en justifiant l'approche adoptée dans cette thèse consistant à développer des approches d'apprentissage direct d'arbres de décision interprétables.

Le troisième chapitre introduit les concepts théoriques et les problématiques techniques liés à la résolution de problèmes de décision séquentielle sous incertitude et à la construction d'arbres de décision interprétables. Après avoir présenté les Processus Décisionnels de Markov, Hector Kohler décrit les différentes approches de l'état de l'art permettant leur résolution : itération de la valeur, apprentissage par renforcement, réseaux de neurones. Hector Kohler traite ensuite des questions d'apprentissage d'arbres de décision interprétables et présente une méthode d'apprentissage indirect basée sur l'apprentissage par imitation d'une politique issue d'un réseau de neurones. Deux méthodes (Dagger et VIPER) permettant cet apprentissage sont présentées. En utilisant ces méthodes indirectes, Hector Kohler montre qu'il peut être, en pratique, très difficile de trouver des arbres de décision offrant un bon compromis entre performance et interprétabilité. Ces observations permettent à Hector Kohler de motiver le développement d'algorithmes d'apprentissage direct d'arbres de décision offrant un bon compromis entre performance et interprétabilité. Hector Kohler conclut ce chapitre en présentant l'organisation de la suite du manuscrit et les différentes contributions de la thèse.

*Ces trois chapitres introductifs posent très clairement le contexte de la thèse et motivent efficacement les problématiques étudiées par rapport aux approches existantes. Le choix des arbres de décision comme un modèle plus interprétable est assez rapidement établi mais la dernière partie de la thèse permettra de revenir sur cette question. Hector Kohler dresse un panorama complet et pédagogique des approches de l'état de l'art. De plus, Hector Kohler illustre judicieusement, par un exemple concret et des résultats expérimentaux, la problématique de la sous-optimalité des arbres de décision obtenus avec des approches indirectes d'apprentissage. Cet exemple contribue ainsi à motiver les orientations scientifiques suivies dans le reste de la thèse.*

La première partie des contributions de la thèse est présentée dans les trois chapitres qui suivent. Cette première partie démontre que l'apprentissage direct d'arbres de décisions interprétables pour des Processus de Décision Markoviens (MDP) est un problème difficile puisque qu'il revient à résoudre un Processus Décisionnel de Markov Partiellement Observable.

Le quatrième chapitre présente l'approche de l'état de l'art proposée par Topin et al.. Ce travail constitue, à l'heure actuelle, la seule approche permettant d'apprendre directement un arbre de décision pour un MDP en optimisant un compromis entre interprétabilité et performance du problème d'apprentissage par renforcement. Cette approche propose de reformuler le MDP

sous la forme d'un Iterative Bounded MDP (IBMDP) et montre qu'une politique déterministe partiellement observable pour ce IBMDP correspond à un arbre de décision pour le MDP original. Les algorithmes d'apprentissage par renforcement de l'état de l'art peuvent ainsi être utilisés afin d'apprendre directement un arbre de décision. Hector Kohler termine ce chapitre en introduisant un objectif d'interprétabilité pour l'apprentissage par renforcement permettant de réaliser le compromis entre performance et interprétabilité.

Le cinquième chapitre compare l'approche directe présentée dans le chapitre précédent à une approche indirecte d'apprentissage des arbres de décision basée sur l'apprentissage par imitation. Ce chapitre vise plus précisément à reproduire les résultats expérimentaux présentés par Topin et al. sur le problème de décision du "CartPole". Après avoir précisément présenté le contexte expérimental mis en œuvre, Hector Kohler décrit les résultats expérimentaux qu'il a obtenus. Ces résultats diffèrent légèrement de ceux décrits par Topin et al. En particulier, Hector Kohler observe que l'apprentissage par renforcement sur le IBMDP (apprentissage direct d'arbres de décision) n'est pas aussi efficace, en termes de performance, que l'apprentissage indirect par imitation. En effet, les résultats expérimentaux obtenus par Hector Kohler montrent que les méthodes d'apprentissage par renforcement résolvant les IBMDP retournent en moyenne des politiques offrant de faibles performances.

Dans le sixième chapitre, Hector Kohler établit des connexions entre les Processus Décisionnels de Markov Partiellement Observables (POMDP) et l'apprentissage direct d'arbres de décision pour des MDPs. Hector Kohler propose une formalisation de ce problème d'apprentissage comme un Partially Observable IBMDP (POIBMDP) dont la résolution consiste à apprendre une politique optimale déterministe partiellement observable. Sur un problème spécifique de déplacement d'un agent sur une grille  $2 \times 2$ , Hector Kohler montre comment définir un POIBMDP qui conduise à l'apprentissage d'arbres de décision de profondeur 1 offrant un bon compromis entre performance et interprétabilité. Des méthodes d'apprentissage par renforcement symétriques et asymétriques sont ensuite utilisées afin de tester l'efficacité des politiques issues de la résolution du POIBMDP. A travers une étude expérimentale comparant différents modèles et algorithmes d'apprentissage, Hector Kohler met en évidence que l'apprentissage direct d'arbres de décision à partir d'un POIBMDP est un problème difficile et trop coûteux pour mener à de bonnes performances en pratique. Les expérimentations menées dans ce chapitre donne un nouvel éclairage sur les sources de complexité liées à l'apprentissage direct d'arbres de décision. Les conclusions de ce chapitre mettent en évidence les limitations d'une approche d'apprentissage direct d'arbres de décision dans un cadre général de résolution de problèmes de décision séquentielle.

A partir des conclusions du précédent chapitre, dans le septième chapitre, Hector Kohler s'intéresse à une classe restreinte de POIBMDP formalisant des problèmes de classification supervisée. Hector Kohler présente tout d'abord une nouvelle formalisation des problèmes de classification par un Processus Décisionnel de Markov (MDP). En repartant de l'approche précédemment développée (et consistant à définir le POIBMDP associé pour l'apprentissage d'arbres de décision) Hector Kohler montre que pour ce type de problème le POIBMDP est en fait un MDP dont la résolution est par nature plus facile. Dans ce cadre totalement observable, les expérimentations développées par Hector Kohler mettent en évidence la possibilité d'apprendre directement des politiques interprétables optimales. Ceci conforte d'autant plus le lien entre observabilité partielle et complexité de l'apprentissage direct d'arbres de décision.

*Les travaux présentés dans cette partie du manuscrit apportent un éclairage nouveau sur les sources de complexité liées à l'apprentissage direct d'arbres de décision. A partir de problèmes de décision bien choisis et grâce à des expérimentations exhaustives, Hector Kohler met en évidence l'impact de l'ob-*

*servabilité partielle sur la difficulté de la tâche d'apprentissage. Bien que les résultats présentés dans cette partie modèrent fortement la pertinence de l'apprentissage direct d'arbres de décision en décision séquentielle, Hector Kohler identifie une classe de problèmes significative, les problèmes de classification, pour lesquels cet apprentissage s'avère une tâche plus facile. On peut alors se demander s'il est possible d'identifier d'autres classes de problèmes qui s'avèreraient elles aussi plus faciles. Ces résultats originaux ouvrent ainsi de nouvelles orientations de recherche pour le domaine, au delà de la thèse d'Hector Kohler.*

Dans la seconde partie du manuscrit, Hector Kohler se concentre sur le problème de l'apprentissage d'arbres de décision en apprentissage supervisé, tâche pour laquelle il a été illustré précédemment que l'apprentissage direct d'arbres de décision était envisageable.

Le huitième chapitre fait office d'introduction à cette deuxième partie et motive la pertinence du développement de nouveaux algorithmes d'induction d'arbres de décision pour la résolution de problèmes de classification supervisée. Hector Kohler introduit les méthodes existantes pour l'induction d'arbres de décision en apprentissage supervisé. Afin d'offrir un compromis entre la précision des approches optimales et l'efficacité en temps des approches gloutonnes, Hector Kohler propose une approche nouvelle basée sur la formulation du problème sous forme d'un MDP et sa résolution par programmation dynamique. L'approche DPDT (Dynamic Programming Decision Trees) ainsi proposée a pour objectif de minimiser la perte d'entraînement tout en limitant le nombre de découpages possibles.

Le neuvième chapitre détaille l'approche DPDT ébauchée dans le précédent chapitre. Hector Kohler détaille tout d'abord la nouvelle formulation du problème sous forme de MDP. Hector Kohler présente ensuite l'algorithme DPDT d'induction d'arbres. Afin de limiter le nombre de découpages possibles de l'ensemble d'exemples, Hector Kohler propose plusieurs heuristiques permettant de limiter le coût de construction du MDP. L'algorithme DPDT exploite ensuite la programmation dynamique afin de résoudre le MDP. Des garanties de performances pour cet algorithme sont ensuite démontrées. Le chapitre se conclut par une discussion sur la complexité des différentes étapes de l'approche.

Le dixième chapitre présente une étude expérimentale précise de l'approche DPDT en la comparant à plusieurs approches de l'état de l'art. Les performances de l'approche sont tout d'abord évaluées et comparées à une approche gloutonne, une approche optimale et l'approche Top-B. Hector Kohler montre ainsi que DPDT permet d'atteindre une perte d'entraînement quasi optimale et une meilleure précision que les autres approches non-optimales. DPDT permet également de réduire significativement le nombre d'opérations nécessaires à l'induction des arbres par rapport aux approches optimales. Ces résultats démontrent la pertinence de l'approche DPDT proposée par Hector Kohler. DPDT offre ainsi un juste milieu entre les solveurs exacts et les approches gloutonnes et complète de façon pertinente le panorama des approches existantes. Dans la deuxième partie du chapitre, Hector Kohler étudie les capacités de généralisation et de "boosting" de l'approche DPDT. Les résultats expérimentaux renforcent la pertinence de l'approche en montrant que DPDT offre de meilleures capacités de généralisation que les autres approches étudiées.

*Cette deuxième partie du manuscrit présente des contributions à la fois théoriques, algorithmiques et expérimentales pour l'induction d'arbres de décision dans le cadre d'un problème de classification. Hector Kohler propose une formulation originale et efficace de ce problème sous forme d'un MDP. L'algorithme DPDT ensuite proposé associe de façon judicieuse la programmation dynamique et des heuristiques efficaces. L'approche ainsi développée offre une alternative convaincante pour l'induction d'arbres de décisions dans un problème de classification. La rigueur des expérimentations menées par Hector Kohler est*

*à souligner et renforce d'autant plus la pertinence des résultats obtenus.*

La troisième partie du manuscrit revient sur la question de départ de l'interprétabilité en apprentissage par renforcement et questionne le compromis interprétabilité-performance de différents modèles de représentation des politiques, en particulier celui des arbres de décision adopté dans les précédentes parties.

Le chapitre 11 traite de l'évaluation de l'interprétabilité d'un modèle. Hector Kohler soulève la question de la mesure de l'interprétabilité sans expérience utilisateur. A partir de travaux de la littérature, Hector Kohler propose d'évaluer l'interprétabilité suivant deux mesures : le temps d'inférence de la politique et la taille de la politique. Afin de s'abstraire de différents détails techniques, Hector Kohler introduit une conversion standardisée des différentes formes de politiques ("unfold policies") qui sera ensuite utilisée dans les expérimentations. La présentation du langage utilisé est assez succincte et aurait mérité d'être plus détaillée. La description du choix de la conversion n'est de plus pas donnée. On peut se demander dans quelle mesure le langage choisi influence les mesures d'interprétabilité qui sont ensuite réalisées.

Le chapitre 12 présente la méthodologie mise en place afin de réaliser l'étude expérimentale sur l'interprétabilité de différentes représentations des politiques. Quatre classes de politiques sont considérées (politiques linéaires, arbres de décision, arbres de décision obliques, réseaux de neurones Relu), différents paramètres sont testés pour chaque classe. Les environnements de tests, issus de bibliothèque du domaine, sont également listés. Hector Kohler propose d'utiliser une approche d'apprentissage par imitation afin d'apprendre les différentes formes de politiques. Hector Kohler illustre enfin la nécessité d'utiliser la conversion standardisée des politiques ("unfold policies") afin d'obtenir des mesures significatives.

Le chapitre 13 présente les résultats expérimentaux obtenus par Hector Kohler sur l'interprétabilité des différentes représentations des politiques. Les résultats obtenus corroborent plusieurs résultats de la littérature comme la difficulté de calculer des politiques interprétables pour des MDP de très grandes tailles ou le fait que le nombre de caractéristiques des états de l'environnement soit prépondérant dans l'interprétabilité d'une politique. Hector Kohler étudie également les variations du compromis entre interprétabilité et performance des différentes représentations. Il s'avère que les résultats sont dépendants des problèmes considérés et que les arbres de décision peuvent s'avérer moins interprétables que des réseaux de neurones sur certains problèmes.

*La troisième partie du manuscrit tente d'apporter des réponses à la difficile question de l'évaluation de l'interprétabilité des politiques de prise de décision. L'absence d'évaluation utilisateur renforce la complexité de cette question. A partir de travaux de la littérature, Hector Kohler donne des réponses convaincantes à cette question en proposant des mesures fondées théoriquement. Hector Kohler mène ainsi avec succès une évaluation sur 4 modèles de politiques dont celui des arbres de décision étudié tout au long du manuscrit. Le protocole expérimental mis en place par Hector Kohler est argumenté et permet de conforter les résultats de la littérature. Les résultats expérimentaux mettent en évidence des niveaux d'interprétabilité mitigés suivant les modèles et les problèmes considérés. De nouveau, la rigueur et l'importance du travail expérimental mené par Hector Kohler est à souligner. Ce travail de qualité permet à Hector Kohler de mener à bien la difficile tâche de l'évaluation de l'interprétabilité. Les résultats mis en évidence par Hector Kohler ouvrent de nombreuses perspectives pour le domaine.*

Un dernier chapitre de conclusion résume les contributions de la thèse et ouvre plusieurs perspectives à ce travail.

*Hector Kohler propose un manuscrit bien rédigé et agréable à lire. Le manuscrit présente des contributions théoriques et expérimentales au domaine de l'interprétabilité en apprentissage par renforcement. Les résultats présentés dans le manuscrit ont en particulier nécessité un travail expérimental très important mené avec une très grande rigueur. Ce travail remarquable a permis à Hector Kohler d'identifier les difficultés et challenges liés à l'apprentissage de politiques interprétables dans une grande variété de problèmes de décision. Hector Kohler a ainsi pu proposer de nouvelles contributions théoriques à l'apprentissage d'arbres de décision interprétables, en particulier pour les problèmes de classification. La pertinence de ces travaux est attestée par une publication à la conférence internationale Knowledge Discovery and Data Mining (KDD 25) et par deux publications dans des workshops internationaux (Seventeenth European Workshop on Reinforcement Learning et Workshop on Programmatic Reinforcement Learning).*

Au vu de tous ces éléments, je donne donc **un avis très favorable** à la soutenance de thèse de Hector Kohler en vue de l'obtention du grade de docteur en Informatique.

Aurélie BEYNIER

A handwritten signature in dark ink, reading "A. Beynier", with a long horizontal flourish extending to the right.