

# Sujet de thèse : Analyse des traces de résolution de problèmes pour une compréhension humaine

April 26, 2024

## 1 Encadrement

### 1.1 Contrat doctoral

Directeur de thèse (40%) :

- Éric Monfroy
- LERIA, université d'Angers
- Courriel : eric.monfroy@univ-angers.fr

Co-encadrant (30%) :

- Vincent Barichard
- LERIA, université d'Angers
- Courriel : vincent.barichard@univ-angers.fr

Co-encadrant (30%) :

- David Genest
- LERIA, université d'Angers
- Courriel : david.genest@univ-angers.fr

## 2 Description du sujet

L'avènement de l'intelligence artificielle a révolutionné notre manière d'appréhender et de résoudre une multitude de problèmes complexes dans divers domaines comme la finance, la logistique, la cybersécurité, ou plus généralement la prise de décisions. Les approches étudiées aujourd'hui, qu'elles soient stochastiques ou basées sur un langage formel, résolvent un problème donné en calculant une solution. Cette solution peut soit être mise en oeuvre directement, soit être évaluée par des utilisateurs qui décident de son application. Pour garantir une maîtrise humaine sur ces processus et éviter une autonomie totale des algorithmes, il est crucial de démystifier les solutions proposées ainsi que les processus ayant mené à leur élaboration.

L'un des domaines établis pour étudier et résoudre ces problèmes est celui de la résolution de problèmes sous contraintes. Ces problèmes sont caractérisés par des ensembles de variables et de relations, communément appelées contraintes, qui les interconnectent. Les problèmes sous contraintes, présents tant dans des contextes académiques qu'appliqués, traversent divers domaines comme l'ordonnancement, la planification, la conception d'emplois du temps ... Deux approches principales se distinguent pour leur résolution : les approches dédiées, développées spécifiquement pour un problème donné, et les approches génériques, applicables à une variété de problèmes. La trace d'une méthode, quelle que soit l'approche, constitue le reflet du chemin parcouru pour atteindre la solution. Mais celle-ci à une signification différente selon qu'elle provienne d'une méthode ad-hoc, le plus souvent stochastique, ou d'une méthode formelle.

Les méthodes dédiées englobent les algorithmes d'apprentissage (réseaux de neurones), ainsi que les méta-heuristiques (recuit simulé, recherche tabou, algorithmes génétiques) spécifiquement développées pour résoudre un problème donné. Ces méthodes encodent généralement les connaissances du problème directement dans le code source. L'algorithme implémenté est très efficace mais souffre d'un manque de généralité. À l'inverse, les approches génériques utilisent des langages de modélisation abstraits tels que CSP, SAT, ASP et CHR, offrant une description haut niveau du problème et de ses contraintes. Un solveur, agissant comme une boîte noire, est ensuite employé pour identifier une ou plusieurs solutions adéquates. Toutes ces méthodes sont capables de produire des traces de leur exécution. Dans le cas d'une méthode dédiée, la trace comporte des éléments provenant de décisions fondées sur des choix aléatoires ne permettant pas d'arguer d'un point de vue logique sur la pertinence d'un choix. La trace ne peut donc pas être utilisée pour expliquer formellement la solution. Dans le cas d'une méthode formelle, la trace reflète l'ordre des choix et de l'application des règles logiques qui ont été utilisées pour décrire le problème. Il est donc possible de mettre en relation les éléments de la trace avec les éléments de la description du problème. La trace constitue ici une explication de la solution, mais celle-ci est souvent très dense et trop complexe pour être compréhensible directement par un non-expert. Afin d'exploiter cette trace, il est essentiel de l'analyser, la reformuler et de la mettre en relation avec la description du problème avant de pouvoir la présenter sous une forme intelligible à un utilisateur.

L'objectif de cette thèse est d'étudier les traces de résolution d'un problème par une approche formelle et de proposer un cadre permettant l'analyse, la réécriture de cette trace afin d'en extraire une explication qui soit intelligible pour un utilisateur. La restitution auprès de l'utilisateur pourra se faire textuellement ou graphiquement. L'étude portera spécifiquement sur les traces générées par CHR++[1], un langage à base de règles et de contraintes développé au LERIA, en visant à élaborer des explications compréhensibles des solutions calculées. L'application de cette recherche à des cas de prise de décision permettra de générer des explications cohérentes ou de révéler les failles dans le processus décisionnel. Ce travail aspire à renforcer la confiance envers les solutions proposées par une méthode d'Intelligence Artificielle en clarifiant les fondements de ces dernières.

### 3 Plan de travail prévisionnel

Ce travail, qui s'inscrit sur une période de 3 ans, se déroulera comme suit :

1. La première année sera consacrée à produire l'état de l'art des approches développées dans la littérature. Cette étude servira de base aux travaux de l'année suivante et pourra être présentée dans une conférence nationale.
2. La seconde année sera consacrée à définition d'un cadre formel permettant de définir et calculer une explication pour l'exécution d'un programme logique ici en CHR (enrichissement des traces, définition d'une ontologie adaptée ...). Ce modèle fera l'objet d'une première implantation dans le système CHR++. Ces travaux pourront être publiés dans des congrès ou revues internationales.
3. La troisième année visera à mettre en application et valider le modèle sur des programmes logiques résolvant des problèmes académiques et réels de prise de décision. Parallèlement, le doctorant se consacrera à la rédaction de son manuscrit de thèse et la publication des résultats obtenus dans des congrès et revues internationales.

### References

- [1] Vincent Barichard. "CHR++: An efficient CHR system in C++ with don't know non-determinism". In: *Expert Systems with Applications* 238 (2024), p. 121810. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2023.121810>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417423023126>.