Apprentissage relationnel et sélection de l'action pour les infrastructures critiques

Fabien Flacher **Pierre Gérard** **Céline Rouveirol**

Thalès, Palaiseau *Université de Paris 13, Villetaneuse*

fabien.flacher@thalesgroup.com prenom.nom@lipn.univ-paris13.fr

Cette thèse CIFRE se déroulera en partenariat entre le laboratoire d'Innovation de THALES (ThereSIS) et l'équipe A3 du Laboratoire d'Informatique de l'Université de Paris Nord (LIPN). Elle permettra d'appliquer aux infrastructures critiques des techniques récentes d'acquisition de modèles d'action relationnels, ainsi que leur exploitation dans le cadre de la recherche automatique de failles dans les ensembles de règles métier.

Introduction

Le problème général abordé dans cette thèse est celui de l'apprentissage automatique au sein des systèmes de supervision. Ces systèmes sont amenés à être déployés dans des structures de plus en plus complexes (ambassade, gare, aéroport, etc.) et soulèvent de nombreuses difficultés liées notamment au grand nombre de processus métiers qu'ils doivent implémenter et faire cohabiter.

Ces processus --- ou règles métiers --- sont au cœur de la valeur ajoutée d'un système de supervision car ils permettent de capturer le fonctionnement d'une infrastructure critique et de définir les réactions d'une application donnée face à l'évolution du système.

Le développement de ces règles métiers constitue aujourd'hui l'un des enjeux majeurs de la conception des systèmes de supervision. Deux difficultés doivent êtres abordées en particulier :

- **La création et de la configuration de nouvelles règles** capturant le fonctionnement d'une infrastructure existante. Lors de la mise au point d'un système de supervision pour une nouvelle infrastructure, le processus de capture et d'intégration de la connaissance du fonctionnement de l'infrastructure sous forme de règles métiers est généralement très coûteuse ;
- **La validation d'un ensemble complet de règles**. Par ailleurs, la quantité de règles modélisées ainsi que la pluralité de leur contexte de mise en œuvre rend leurs interactions potentielles très difficile à appréhender et à valider (ou invalider).

Apprentissage de modèles

L'apprentissage automatique permet justement de construire automatiquement des représentations intentionnelles à partir d'exemples. Typiquement, il peut s'agir de découvrir par apprentissage quelques règles générales permettant de synthétiser un grand nombre d'exemples fournis au système d'apprentissage.

Dans le cadre applicatif défini plus haut, ce paradigme peut être employé à plusieurs niveaux :

- Pour construire automatiquement une représentation synthétique du fonctionnement d'une infrastructure observée ;
- Pour exploiter le modèle obtenu afin de chercher automatiquement comment le fonctionnement du système de supervision pourrait être mis en défaut. Ce second problème revient à chercher, pour chaque état possible de l'infrastructure, la meilleure action à entreprendre pour parvenir à un objectif défini. Si on veut mettre en défaut un système, on peut par exemple donner comme objectif les situations indésirables, et laisser le système apprendre un moyen de les atteindre, si c'est possible.

En apprentissage, ce double problème peut être abordé de deux manières :

- Acquisition d'un modèle puis **planification** : le modèle appris est utilisé pour raisonner et construire une séquence d'actions permettant d'atteindre l'objectif assigné. Si l'objectif est la mise en défaut, on peut ainsi obtenir une séquence d'actions qui montrent comment invalider les règles métier envisagées ;
- Acquisition d'un modèle conjointement à **apprentissage par renforcement** : c'est par expérimentations successives que l'on apprend à atteindre les objectifs fixés. Le modèle est employé pour "simuler" des

expériences et accélérer l'apprentissage des bonnes/mauvaises actions dans chaque situation possible.

Dans tous les cas, la combinaison du nombre d'objets d'une infrastructure et de la présence de nombreux acteurs humains engendre un fonctionnement hautement dynamique des infrastructures critiques supervisées. Ce facteur multiplie les situations possibles et rend complexe la modélisation ainsi que son exploitation. Il convient donc de se tourner vers des méthodes permettant un passage à l'échelle important.

Représentations relationnelles

Afin de permettre un apprentissage dans des espaces d'états et d'action de grande taille, on peut envisager plusieurs pistes :

- Raffiner encore des techniques déjà très sophistiquées pour leur permettre de mieux passer à l'échelle. Par exemple, les derniers développement en apprentissage par renforcement dit "factorisé" sont extrêmement efficaces mais restent limitées à des problèmes de quelques millions d'états ;
- Utiliser des représentations plus riches pour les espaces d'états et d'action.

Avec des langages de représentation utilisant des restrictions de la logique d'ordre un plutôt que des langages propositionnels, les états et les actions sont représentées par des prédicats logiques exprimant des relations entre objets plutôt que par des vecteurs de valeurs numériques. Ce changement de paradigme offre de nouvelles possibilités d'apprentissage, notamment des possibilités de généralisation et de passage à l'échelle des solutions apprises qui sont hors de portée systèmes opérant dans des langages propositionnels.

Tous les problèmes d'apprentissage ne s'expriment pas d'emblée de manière relationnelle mais dans de nombreuses applications (jeux vidéo, simulations d'infrastructures par exemple) il peut être plus facile de produire une représentation relationnelle qu'une représentation propositionnelle. Les objets et leurs relations étant immédiatement disponibles, il peut même sembler peu naturel de produire indirectement des représentations par attributs/valeurs.

Conclusion

La proposition de cette thèse est donc de partir des dernières avancées en matière d'apprentissage automatique, notamment les méthodes par renforcement relationnel, et de les appliquer aux problématiques de la conception des ensembles de règles métiers des systèmes de supervision d'infrastructures critiques. Ces travaux seront intégrés chez THALES au sein du développement des nouveaux outils de conception et de validation de ces systèmes de supervision. Ils permettront de répondre aux besoins d'outils d'assistance à la conception de ces ensembles de règles métiers. En particulier, en s'appuyant sur l'apprentissage automatique d'un modèle de l'infrastructure étudiée, ces travaux devraient permettre de fournir automatiquement une cartographie de la dynamique de fonctionnement l'infrastructure critique étudiée. Cette cartographie servira alors de base à un outil de construction de nouvelles règles métiers capturant automatiquement cette dynamique.

De plus, en exploitant la capacité de ces algorithmes à apprendre une politique optimale, il sera également possible de construire un outil apprenant automatiquement à identifier les cas potentiellement défectueux, instanciant une véritable méthode d'invalidation automatique de ces ensembles de règles. Cet outil pourrait constituer le premier pas vers un procédé industriel permettant d'évaluer automatiquement ces systèmes.

Ces travaux s'inscriront dans la continuité des travaux actuellement menés au sein du laboratoire d'Innovation de THALES (ThereSIS) sur l'apprentissage, ainsi que dans celle des travaux de recherche menés actuellement dans le cadre du projet ANR HARRI, au sein de l'équipe d'apprentissage du Laboratoire d'Informatique de Paris Nord.

Informations diverses

Financement : these CIFRE entre THALES et de LIPN, participation aux frais de transports et aux déjeuners
Lieux :THALES sur le campus de l'Ecole Polytechnique (Palaiseau 91120), LIPN (Villetanneuse 93430)

Compétences requises

Algorithmique, bon niveau en C++ ainsi qu'en modélisation objet, prolog et système d'apprentissage automatique.

Pour candidater

Envoyer par mail aux trois encadrants : un CV, une lettre de motivation et si possible, des relevés de notes récents.