

PROGRAMME DE DOCTORAT EN INFORMATIQUE
FACULTÉ DES SCIENCES ET DE GÉNIE

SOUTENANCE DE THÈSE

de Sami Zhioua

Le jeudi 31 janvier 2008 à 10h

à la salle 3632 du pavillon Louis-Jacques Casault

Stochastic Systems Divergence through Reinforcement Learning

PRÉSIDENT

Professeur Mario Marchand, Ph.D.

Directeur des programmes de deuxième cycle
Département d'informatique et de génie logiciel

SECRÉTAIRE DE LA SOUTENANCE

Professeur Mario Marchand, Ph.D.

Directeur des programmes de deuxième cycle
Département d'informatique et de génie logiciel

EXAMINATEURS

Professeure Josée Desharnais, Ph.D. (directrice de thèse)
Département d'informatique et de génie logiciel, Université Laval

Professeur François Laviolette, Ph.D. (co-directeur)
Département d'informatique et de génie logiciel, Université Laval

Professeure Marta Kwiatkowska, Ph.D. (examinatrice)
Computing Laboratory, Oxford University

Professeure Doina Precup, Ph.D. (examinatrice)
School of Computer Science, McGill University

Professeur Danny Dubé, Ph.D. (examineur)
Département d'informatique et de génie logiciel, Université Laval

Résumé

Les mathématiques offrent un cadre convenable pour raisonner rigoureusement sur les systèmes et phénomènes réels. Par exemple, en génie logiciel, les méthodes formelles sont parmi les outils les plus efficaces pour détecter les anomalies dans les logiciels. Plusieurs systèmes réels sont stochastiques par nature dans le sens où leur comportement est sujet à un aspect d'incertitude. La représentation de ce genre de systèmes requiert des modèles stochastiques comme les processus de Markov étiquetés (LMP), les processus de Markov décisionnels (MDP), etc. Cette thèse porte sur la quantification de la différence entre les systèmes stochastiques. Le résultat principal de la thèse est que l'apprentissage par renforcement, qui est une branche de l'intelligence artificielle particulièrement efficace en présence d'incertitude, peut être utilisé pour quantifier efficacement cette divergence. L'idée clé est de définir un MDP à partir des systèmes à comparer de telle sorte que la valeur optimale de cet MDP corresponde à la divergence entre eux. La caractéristique la plus attrayante de l'approche proposée est qu'elle est complètement indépendante des structures internes des systèmes à comparer. Pour cette raison, l'approche peut être appliquée à différents types de systèmes stochastiques. La deuxième contribution est une nouvelle famille de notions d'équivalence, que nous appelons K-moment, qui est plus forte que l'équivalence par trace mais plus faible que la bisimulation. Cette famille se définit naturellement à travers la coïncidence de moments de variable aléatoires (d'où son nom) et possède une caractérisation simple en terme de tests. Nous montrons que K-moment fait partie d'un cadre plus grand, appelé test-observation-équivalence (TOE), qui constitue la troisième contribution de cette thèse. Il s'agit d'un cadre plus flexible pour la définition des notions d'équivalence basé sur les tests.

Abstract

Modelling real-life systems and phenomena using mathematical based formalisms is ubiquitous in science and engineering. The reason is that mathematics offer a suitable framework to carry out formal and rigorous analysis of these systems. For instance, in software engineering, formal methods are among the most efficient tools to identify flaws in software. The behavior of many real-life systems is inherently stochastic which require stochastic models such as labelled Markov processes (LMPs), Markov decision processes (MDPs), predictive state representations (PSRs), etc. This thesis is about quantifying the difference between stochastic systems. The important point of the thesis is that reinforcement learning (RL), a branch of artificial intelligence particularly efficient in presence of uncertainty, can be used to quantify efficiently the divergence between stochastic systems. The key idea is to define an MDP out of the systems to be compared and then to interpret the optimal value of the MDP as the divergence between them. The most appealing feature of the proposed approach is that it does not rely on the knowledge of the internal structure of the systems. Only a possibility of interacting with them is required. Because of this, the approach can be extended to different types of stochastic systems. The second contribution is a new family of equivalence notions, K-moment, that constitute a good compromise between trace equivalence (too weak) and bisimulation (too strong). This family has a natural definition using coincidence of moments of random variables but more importantly, it has a simple testing characterization. K-moment turns out to be part of a bigger framework called test-observation-equivalence (TOE), which we propose as a third contribution of this thesis. It is a refined testing framework to define equivalence notions with more flexibility.