

Dynamique des systèmes complexes : modèles probabilistes, relationnels et dynamiques de grande taille

Mots clés :

- **Directeur de thèse** : Christophe Gonzales
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire d'informatique de Paris 6
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

Les réseaux bayésiens (BNs) représentent aujourd'hui l'un des modèles les plus utilisés pour la représentation des connaissances probabilisées. Un réseau bayésien représente graphiquement la probabilité jointe d'un grand nombre de variables, de manière assez compacte pour pouvoir être manipulée. Des algorithmes d'inférence et d'apprentissage ont été et sont encore développés afin de maîtriser au mieux l'explosion combinatoire inhérente à la représentation exhaustive d'un système complexe. Dès que les algorithmes de base des BNs se sont révélés assez efficaces pour permettre de tester des modélisations de systèmes complexes et de grandes tailles, l'idée d'améliorer le modèle par une représentation structurée des BNs a naturellement émergé ; les lacunes de la représentation en BN qu'il s'agit ainsi de combler sont, par exemple, la complexité de l'inférence, la rigidité du langage de représentation, l'effort important que demande la spécification d'un modèle de grande taille, etc. Dans cette optique, plusieurs extensions des BNs ont été proposées avec, en particulier, les réseaux bayésiens dynamiques (dBNs) et les modèles probabilistes relationnels (PRM Probabilistic Relational Model). Les réseaux bayésiens dynamiques (Dean, T. and Kanazawa, 1990) permettent de modéliser des processus stochastiques et étendent les modèles tels que les chaînes de Markov, les chaînes de Markov cachés, hiérarchiques, les filtres de Kalman, etc. La particularité des dBNs étant d'exploiter les hypothèses de Markov d'ordre 1 et d'homogénéité du processus afin de pouvoir représenter l'état du système, ainsi que sa dynamique, de manière factorisée à l'aide d'un BN. Bien que proposés au début des années 1990, les dBNs constituent encore un objet de recherche (Singliar, T. and Dash, D., 2008) de par la complexité des modèles obtenus qui imposent des algorithmes d'apprentissage tout à fait spécifiques (Murphy, K., 2002). Les modèles structurés englobent tous les langages de description d'objets liés entre eux par des relations probabilistes d'arité quelconque. Depuis une dizaine d'années, un ensemble de propositions ont donc vu le jour : - Graph-grammaire (Edgar and Musen, 1985), - Multi-Sectionned BN (Xiang, 95), - Relational Bayesian Networks (Jaeger 97), - Probabilistic Relational Models (Friedman et al, 99), - OOBN (Koller and Pfeffer 97), - Probabilistic Frame-Based Systems (Koller and Pfeffer 98), - OOBN (Bangso et Wuillemin, 2000), - Multi-Entity BN, (Laskey and Pfeffer, 05). Nous considérons principalement les propositions type OOBN et PRM dans cette thèse (Torti et al. 2010). Cette structuration de haut niveau des BNs apporte des avantages supplémentaires telles que des algorithmes ad-hoc et plus efficaces, parallélisables. Les algorithmes efficaces d'apprentissage dans les réseaux bayésiens orientés objets (implantant complètement un langage objet) sont encore rares. Si certains algorithmes existent (Flores et al (2003), Langseth, and Bangsø (2001)), il manque encore des outils de modélisation capables de fournir efficacement les fonctionnalités qui permettraient d'exploiter pleinement cette méthode de modélisation, notamment pour des systèmes évoluant dans le temps. Le but de cette thèse est d'étudier des modèles hybrides (PRM dynamique) (Friedman, N., Koller, D. and Pfeffer, A., 1998) et de proposer des algorithmes d'apprentissage de PRM dynamiques, en particulier des PRMs dont la structure et les paramètres peuvent évoluer dans le temps. Dans un second temps, il s'agira de développer des méthodes d'élicitation de savoir d'experts pouvant être conjuguées avec les algorithmes d'apprentissage afin d'étendre le champs d'application des PRMs. Pour se faire, cette thèse s'appuiera sur les travaux déjà effectués par l'équipe dans le cadre de l'ANR SKOOB, qui ont permis de développer un langage PRM complètement orienté objet qu'il s'agira d'étendre au cadre dynamique.

Résumé du projet de recherche (Langue 2)

La modélisation probabiliste des systèmes complexes dynamiques est un domaine où les verrous technologiques et scientifique sont nombreux. De nombreux travaux et équipes travaillent sur des sujets adjacents ou connexes. Toutefois, il est à noter que l'emphase est souvent mise sur l'apprentissage de domaine structuré (Getoor et Taskar, 2007 ou Bakir et al, 2007), nous adopterons donc ici un point de vue plus original consistant à s'intéresser non seulement à l'apprentissage de données dans une structuration orientée objet mais également à la modélisation via des savoirs d'experts de tels modèles. En terme applicatif, les champs visés tout particulièrement sont la maintenance, la gestion du risque, les applications militaires et marketing, les modélisations de processus complexes comme des processus biologiques, etc.

Informations complémentaires (Langue 2)

Bakir, G., Hofmann, T., Scölkopf, B., Smola, A.J., Taskar, B. And Vishwanathan, S.V.N. Predicting Structured Data, MIT Press, 2007

Bangsø O. and P.-H. Wuillemin (2000a). Top-down construction and repetitive structures representation in Bayesian Networks. Thirteenth International Florida Artificial Intelligence Research Symposium Conference, Florida, USA.

Bangsø, O. and Wuillemin, P.H., 2000. Object Oriented Bayesian Networks: A Framework for Topdown Specification of Large Bayesian Networks and Repetitive Structures. Technical Report CIT-87.2-00-obphw1. Aalborg: Department of Computer Science, Aalborg University.

Dean, T. and Kanazawa, K. A model for reasoning about persistence and causation. *Comput. Intell.*, 5(3) :142–150, 1990

Egar, J. and Musen, M. Graph-grammar assistance for automated generation of influence diagrams, in Proceedings of the 1st Annual Conference on Uncertainty in Artificial Intelligence (UAI-85). Elsevier Science Publishing Company, Inc., 1985.

Friedman, N., Koller, D. and Pfeffer, A. (1998). Structured representation of complex stochastic systems. Proceedings of the 15th National Conference on Artificial Intelligence (AAAI) (pp. 157-164).

Friedman, N. and Getoor, L. and Koller, D. and Pfeffer, A. Learning Probabilistic Relational Models (1999) Getoor, L. and Taskar, B. Eds. Introduction to Statistical Relational Learning, MIT Press, 2007

Jaeger, M. Relational Bayesian networks, in Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence, M. Kaufmann, Ed., pp. 266--273, 1997.

Koller, D. and Pfeffer, A. Object-oriented bayesian networks, in Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97), pp. 302--313, 1997.

Koller, D. and Pfeffer, A., Probabilistic frame-based systems, in AAAI/IAAI, pp. 580--587, 1998.

Laskey, K.B., MEBN: A Logic for Open-World Probabilistic Reasoning George Mason University C4I Center Technical Report C4I06-01, 2006.

Murphy, K. Dynamic Bayesian Networks : Representation, Inference and Learning. PhD thesis, 2002

Pearl, J. Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference. Morgan Kaufmann, 1988

Singliar, T. and Dash, D. Efficient inference in persistent dynamic bayesian networks. Uncertainty in Artificial Intelligence, 2008

Lionel Torti, Pierre-Henri Wuillemin and Christophe Gonzales "Reinforcing the Object-Oriented Aspect of Probabilistic Relational Models". Proceedings of PGM, pp. 273-280, 2010

Xiang, Y., Jensen, F. Inference in Multiply Sectioned Bayesian networks with extended Shafer-Shenoy and lazy propagation. In Proceedings on the Fifteenth Conference on Uncertainty in Artificial Intelligence. Pages 680-687. Morgan Kaufmann Publishers, San Francisco, 1999.