



# Apprentissages dans les CBNs non paramétriques

## Contexte

Les réseaux bayésiens permettent la représentation compacte d'une distribution de probabilité multivariée par le biais d'un graphe dirigé acyclique (DAG) dont les nœuds sont associés aux variables aléatoires ([1], [2]). La structure du graphe, quant à elle, encode les indépendances conditionnelles entre ces variables et induit une factorisation de la loi jointe en un produit de distributions de probabilités conditionnelles (CPD). L'inférence, qui consiste à connaître la probabilité d'un ensemble de variables aléatoires d'intérêt étant donné un ensemble d'observations, est facilitée par le formalisme des réseaux bayésiens [3], bien que restant NP-difficile [4]. La structure du graphe est la plupart du temps apprise à partir de données via plusieurs types de méthodes d'apprentissage ([4], [5], [6]). Cependant, ces méthodes s'étendent difficilement à des variables continues et les extensions se restreignent souvent au cas gaussien ([7], [8]) pour pouvoir mener les calculs.

Le modèle des copules ([9], [10], [11], [12]), d'un autre côté, permet d'encoder les dépendances entre variables continues tout en s'abstrayant des comportements marginaux de celles-ci. Les copules ont cependant le défaut de n'être maniable qu'en faible dimension ( $n < 10$ ). Il existe donc d'un côté les réseaux bayésiens qui permettent de réduire la complexité (grande dimension) en utilisant la localité mais qui s'étendent difficilement aux variables continues et d'un autre côté les copules qui permettent de traiter les dépendances entre variables continues mais seulement en faible dimension.

Pour réconcilier les deux mondes, il existe de nombreux modèles graphiques à base de copules ([13], [14], [15]). Parmi ceux-ci, ce sont les *copula bayesian networks* (CBN) ([16], [17]) qui retiendront notre attention, car ils présentent les mêmes propriétés de localité que les réseaux bayésiens.

## Objectifs

Le but de cette thèse est de contribuer au domaine de l'apprentissage des CBNs. Il s'agira dans un premier temps d'étudier les possibilités d'utiliser les méthodes d'apprentissage du cadre classique (discret). La voie choisie actuellement consiste à proposer des scores ou des tests d'indépendance dans le cadre de lois continues, non paramétriques. En particulier un algorithme de type PC [6] utilisant un test d'indépendance non-paramétrique [18] sera envisagé. D'autres méthodes d'apprentissages non-paramétriques seront par la suite étudiées et en particulier des méthodes utilisant des notions de théorie de l'information (entropie, information mutuelle) continue.

Ces nouvelles techniques d'apprentissage des CBNs seront implémentées dans le module C++ wrappé en python: *otagrum*. Ce module permet la communication entre deux bibliothèques C++ que sont *aGrUM*, spécialisé dans la gestion de modèles graphiques discrets et *openTURNS*, spécialisé dans la gestion de distributions continues et en particulier de copules.



## Bibliographie

- [1] Koller, Daphne, and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
- [2] Darwiche, Adnan. *Modeling and reasoning with Bayesian networks*. Cambridge university press, 2009.
- [3] Shenoy, Prakash P., and Glenn Shafer. "Axioms for probability and belief-function propagation." *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Springer, Berlin, Heidelberg, 2008. 499-528.
- [4] Cooper, Gregory F. "The computational complexity of probabilistic inference using Bayesian belief networks." *Artificial intelligence* 42.2-3 (1990): 393-405.
- [5] Heckerman, David. "A tutorial on learning with Bayesian networks." *Innovations in Bayesian networks*. Springer, Berlin, Heidelberg, 2008. 33-82.
- [6] Sprites, P., C. Glymour, and R. Scheines. "Causation, prediction and search, Cambridge." (2000)
- [7] Cooper, Gregory F., and Edward Herskovits. "A Bayesian method for the induction of probabilistic networks from data." *Machine learning* 9.4 (1992): 309-347.
- [8] Lauritzen, Steffen Liholt, and Nanny Wermuth. "Graphical models for associations between variables, some of which are qualitative and some quantitative." *The annals of Statistics* (1989): 31-57.
- [9] Geiger, Dan, and David Heckerman. "Learning gaussian networks." *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1994.
- [10] Nelsen, Roger B. *An introduction to copulas*. Springer Science & Business Media, 2007.
- [11] Joe, Harry. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.
- [12] Genest, Christian, and Anne-Catherine Favre. "Everything you always wanted to know about copula modeling but were afraid to ask." *Journal of hydrologic engineering* 12.4 (2007): 347-368.
- [13] Sklar, M. "Fonctions de repartition an dimensions et leurs marges." *Publ. inst. statist. univ. Paris* 8 (1959): 229-231.
- [14] Huang, Jim C. *Cumulative distribution networks: Inference, estimation and applications of graphical models for cumulative distribution functions*. University of Toronto, 2009.



- [15] Bedford, Tim, and Roger M. Cooke. "Probability density decomposition for conditionally dependent random variables modeled by vines." *Annals of Mathematics and Artificial intelligence* 32.1-4 (2001): 245-268.
- [16] Aas, Kjersti, et al. "Pair-copula constructions of multiple dependence." *Insurance: Mathematics and economics* 44.2 (2009): 182-198.
- [17] Elidan, Gal. "Copula bayesian networks." *Advances in neural information processing systems*. 2010.
- [18] Elidan, Gal. "Lightning-speed structure learning of nonlinear continuous networks." *Artificial Intelligence and Statistics*. 2012.
- [19] Bouezmarni, Taoufik, Jeroen VK Rombouts, and Abderrahim Taamouti. "Nonparametric copula-based test for conditional independence with applications to granger causality." *Journal of Business & Economic Statistics* 30.2 (2012): 275-287.
- [20] Sancetta, Alessio, and Stephen Satchell. "The Bernstein copula and its applications to modeling and approximations of multivariate distributions." *Econometric theory* 20.3 (2004): 535-562.