

# Recent Advances in Probabilistic Graphical Models

Concha Bielza,<sup>1,\*</sup> Serafín Moral,<sup>2,†</sup> Antonio Salmerón<sup>3,‡</sup>

<sup>1</sup>*Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, 28660 Boadilla del Monte, Madrid, Spain*

<sup>2</sup>*Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain*

<sup>3</sup>*Department of Mathematics, University of Almería, 04120 Almería, Spain*

Probabilistic graphical models constitute a fundamental tool for the development of intelligent systems. They provide a sound and well-founded approach for performing inference and belief updating in complex domains endowed with uncertainty. A probabilistic graphical model is the result of the combination of a qualitative component (a graph) encoding conditional independence relationships among the variables in the system and a quantitative component consisting of a collection of local probability distributions matching the independence properties specified by the graph. The union of the two components provides a compact representation of the joint probability distribution over the domain being modeled. Bayesian networks are the most prominent type of probabilistic graphical models and have experienced a remarkable methodological development during the past two decades. This has come along with a wide variety of successful applications in different domains. Regardless of the increasing interest in the area, probabilistic graphical models are still facing a number of challenges, covering modeling, inference and learning.

This special issue contains seven papers that contribute to the methodological development beyond the current state-of-the-art knowledge. Five of the papers were selected among the contributions presented at the 15th Conference of the Spanish Association of Artificial Intelligence (CAEPIA'2013, Madrid, Spain, September 17–20, 2013). These papers were substantially extended and went through a new and strict review process. The other two papers were invited contributions not presented at the conference and went through the same review process.

The first three papers deal with hybrid models, where both discrete and continuous variables coexist. Lucas and Hommersom propose a framework for handling causal independence covering discrete and continuous variables simultaneously. The methodology is based on the convolution concept from probability theory, and the

\*Author to whom all correspondence should be addressed; e-mail: [mcbielza@fi.upm.es](mailto:mcbielza@fi.upm.es).

†e-mail: [smc@decsai.ugr.es](mailto:smc@decsai.ugr.es)

‡e-mail: [antonio.salmeron@ual.es](mailto:antonio.salmeron@ual.es)

authors show its potential impact as a modeling tool as well as a means of speeding up probabilistic inference.

Also with discrete and continuous variables, Varando *et al.* present two methods for learning mixture of polynomial approximations of conditional densities. One is based on conditional sampling and the other on interpolation. Both methods involve estimating the joint and the marginal densities and do not require discretizing the conditioning variables.

Shenoy *et al.* analyze the practical implications of the presence of deterministic conditionals in hybrid Bayesian networks modeled as mixtures of exponentials and mixtures of polynomials. They identify the key problems that arise when carrying out inference in such situations and propose different approximation schemes aimed at making inference tractable.

The next three papers concern probabilistic graphical model learning. Arias *et al.* improve the efficient and scalable local-search algorithm called fast constrained hill climbing (FastCHC) to learn Bayesian networks. Three modifications try to relax FastCHC constraints by adding diversification to the search process in the structure learning. The computational complexity is not altered, and the resulting networks are highly scored.

Hernández-González *et al.* learn the parameters of a multidimensional Bayesian network classifier when the data are (subjectively) labeled by a crowd of annotators. Annotator expertise, estimated by the expectation-maximization algorithm, is used to calibrate the annotators' contribution to the learning process, promoting the labels provided by outstanding annotators.

Cano *et al.* consider a generalization of Bayesian networks, recursive probability trees, able to represent joint probabilities in a compact way when there are product factorizations or/and contextual conditional independence relationships. They propose a greedy algorithm to learn the structure of a recursive probability tree based on computing a dependence degree between each couple of variables. An experimental study shows that the proposed approach is competitive with classical learning algorithms as PC or K2.

Finally, the paper by Sonntag *et al.* focuses on the graphical part of the model. It tries to answer some basic questions that have important implications in the tasks of modeling and learning. The first one is the ratio between directed acyclic graphs and essential graphs for a fixed number of variables. The number of chain graphs with respect to directed acyclic graphs is also computed. A Markov chain Monte Carlo approach is used to obtain an approximation of these quantities. An important conclusion is that the average Markov equivalent class size is small, and an important gain in efficiency should not be expected when learning using essential graphs instead of directed acyclic graphs.

We would like to thank Prof. R. R. Yager, Editor-in-Chief of the *International Journal of Intelligent Systems*, for giving us the opportunity to edit this special issue. We also appreciate the valuable help of the anonymous reviewers for their professional, accurate, and timely comments.