International Journal of Approximate Reasoning 68 (2016) 88–90

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Editorial

Special Issue on the Seventh Probabilistic Graphical Models Conference (PGM 2014)

This special issue of the International Journal of Approximate Reasoning is devoted to a selection of papers presented at the 7th Probabilistic Graphical Models Conference (PGM 2014), which took place in Utrecht, The Netherlands, from September 17th to September 19th, 2014.[1] The biennial PGM series continues to attract researchers from most of the European research groups on probabilistic graphical models as well as an increasing number of non-European researchers. The 7th edition welcomed authors and participants representing 22 different countries and 5 continents. First introduced in 2002 as a European workshop, PGM has now grown to become an international conference.

After a rigorous review process, 38 submissions were accepted both for publication in the Springer Lecture Notes in Artificial Intelligence [1] and for plenary presentation at the conference. The authors of ten of these papers were invited to submit a revised and extended version of their paper to the current special issue. Selection of papers was based upon the scientific quality and the potential for extension of the PGM contribution, judged both by PGM reviewers and (session) chairs. Submissions were received from all invited authors, and after a full reviewing process according to IJAR standards, nine papers were accepted for publication in this special issue. The nine accepted papers illustrate the wide scope of PGM, dealing with different types of probabilistic graphical model such as Bayesian networks, Influence diagrams and Chain graphs. The focus of the research described in the papers varies from studying foundations to introducing novel approaches to inference and learning.

The papers by Sonntag and Peña (On Expressiveness of the Chain Graph Interpretations), and by Peña and Gómez-Olmedo (Learning Marginal AMP Chain Graphs under Faithfulness Revisited) concern chain graphs, a probabilistic graphical model that builds upon graphs with both directed and undirected edges. Different approaches to encoding conditional independence relations in chain graphs exist, resulting in different chain graph interpretations. The paper by Sonntag and Peña compares three different chain graph interpretations in terms of the number of independence models that they can represent. One of their conclusions is that the AMP (Andersson–Madigan–Perlman) and MVR (multivariate regression) chain graph interpretations are more expressive than the LWF (Lauritzen–Wermuth–Frydenberg) interpretation. Marginal AMP (MAMP) chain graphs unify and generalise the AMP and MVR chain graph interpretations. The paper by Peña and Gómez-Olmedo focusses on learning such MAMP chain graphs. More specifically, the paper presents an algorithm for learning MAMP chain graphs from probability distributions whose independence relation can be perfectly captured in such a graph.

A popular probabilistic graphical model, which can in fact be seen as a special case of a MAMP chain graph, is the Bayesian network. The paper by Butz, Oliveira and Madsen (Bayesian Network Inference using Marginal Trees) addresses repeated inference in Bayesian networks to answer a series of queries. It introduces so-called marginal trees as a secondary structure that represents past computations. The marginal trees enable the prevention of unnecessary computations and the efficient reuse of repeatedly required computations. The inference algorithm that exploits this secondary structure can be viewed as modified one-way propagation in a join tree; experiments comparing marginal tree inference with both lazy propagation and variable elimination demonstrate that the new approach is promising.

Bayesian networks of restricted topology are often used for classification tasks where the class of, for example, an object is to be predicted from a description of its features. Different types of Bayesian network classifier exist, varying in the number of class variables and the restrictions on the structure allowed among the class variables, among the feature variables, and between the class and feature structures. The paper by De Campos, Corani, Scanagatta, Cuccu and Zaffalon (Learning Extended Tree Augmented Naive Structures) proposes ETAN, a one-dimensional classifier (i.e. with single class variable) which extends the tree-augmented naive Bayes (TAN) classifier by relaxing some of its structural constraints. The

[1] http://www.projects.science.uu.nl/PGM/.

http://dx.doi.org/10.1016/j.ijar.2015.09.002

paper in addition presents and evaluates an efficient structure learning algorithm for ETANs that incorporates a search strategy and new score function, and results in improved prediction accuracy compared to TAN and Naive Bayes. The ideas behind the presented algorithm may also prove useful for structure learning of Bayesian networks of general topology.

In multi-label or multi-dimensional classification, the classification task consists of assigning a value to more than one class variable. The paper by Varando, Bielza and Larrañaga (Decision Functions for Chain Classifiers Based on Bayesian Networks for Multi-Label Classification) investigates the expressiveness of Bayesian network-augmented naive Bayes (BAN) classifiers with multiple binary-valued class variables by studying the number of distinct decision functions induced by two types of such multi-label classifiers. One, built using the binary relevance method, consists of a set of one-dimensional classifiers that each solves the problem independently, upon which the individual results are combined to obtain a multi-label prediction. The other, a chain classifier, captures dependencies between class variables by linking them into a chain, ensuring that each classifier includes as additional features the class predicted by the previous classifiers. The authors prove upper bounds on the expressive power of the two methods and show that chain classifiers are more expressive than the binary relevance-based models.

If the labels in a multi-label classification problem are ordered in a predefined structure such as a tree or an acyclic directed graph (DAG), e.g. based upon "is-a" relations, then the classification task is called Hierarchical Multi-label Classification (HMC). In HMC, consistent predictions are predictions where the associated labels form a path in the hierarchy, i.e. if an instance belongs to a certain class, it should also belong to all its super-classes. Such a path may be a complete path from the root to a leaf (mandatory leaf node prediction), or may end in an intermediate node in the hierarchy. The paper by Ramírez-Corona, Sucar and Morales (Hierarchical Multi-label Classification Based on Path Evaluation) concerns the latter type of prediction. The authors propose a novel HMC approach called Chained Path Evaluation (CPE) that can cope with large datasets by training a local classifier for each non-leaf node in the hierarchy. As in chain classifiers, CPE incorporates information from the hierarchy by including the parent label; CPE then selects the best path(s) in the hierarchy based on a newly designed score that avoids the bias towards conservative predictions found in many existing evaluation measures. Experiments demonstrate that CPE is at least competitive compared to state-of-the-art methods, outperforming them on DAG hierarchies and on deep hierarchies with many leaf nodes.

To specifically capture pairwise interactions between class variables in the context of multi-dimensional classification, the paper by Arias, Gamez, Nielsen and Puerta (A Scalable Pairwise Interaction Framework for Multidimensional Classification) proposes a two-step approach. In the first phase, for each pair of class variables, a one-dimensional classifier is learnt for the combined state space of the two original class variables. Subsequently, in phase 2 the individual class predictions from the one-dimensional classifiers are combined in a Markov random field, which upon inference returns the multi-dimensional prediction. In addition to the main framework and its properties, the authors describe strategies for supporting the scalability of the framework. Experiments demonstrate that the proposed method significantly outperforms or is comparable to state-of-the-art methods.

Decision making can be explicitly supported by extending a Bayesian network to an Influence diagram, which incorporates preferences and actions. In a Bayesian network, the computation of the maximum a-posteriori probability assignment for a subset of network variables (MAP) is an important task. Likewise, in an Influence diagram the computation of the maximum expected utility of a strategy (MEU) is important. Mauá (Equivalences between Maximum A Posteriori Inference in Bayesian Networks and Maximum Expected Utility Computation in Influence Diagrams) constructively shows that the two problems are quite similar: one can be reduced to the other, and vice versa, with only small overhead and limited increase in tree-width. Experiments show that MAP solvers are often better at solving the corresponding MEU problem than MEU solvers. The author also exploits the reductions to determine the complexity of MEU problems with imprecise probabilities from results concerning the complexity of MAP with imprecise parameters.

The Limited Memory Influence Diagram (LIMID) is an Influence diagram variant that allows for representing problems that violate the *no forgetting* assumption that variables known at the time of a decision should be known at all later decision moments. Solving LIMIDs is often done by local search. The paper by Mauá and Cozman (Fast Local Search Methods for Solving Limited Memory Influence Diagrams) proposes (exact and approximate) dominance pruning strategies to speed up k-local search. Approximate pruning is fully polynomial additive in bounded-tree-width bounded-variable cardinality LIMIDs with k action variables. Experiments show very promising results for $k > 5$ and only minor loss in accuracy due to approximate pruning.

This special issue was made possible with the help of many others. First of all the authors must be complemented for their hard work that has resulted in high quality papers with a nice balance between theory and experimental work. Valuable input for all papers was provided by the reviewers; I am therefore grateful to the PC members of PGM 2014 and all additional reviewers for their efforts and competence. For aiding in the initial selection of papers for this special issue, I would in addition like to thank Linda van der Gaag and Ad Feelders — PGM 2014 general chairs and my program co-chairs — and all PGM session chairs. Finally, I would like to thank IJAR's Editor-in-Chief Thierry Denoeux for the chance to publish this special issue and for guiding me through the process of managing and editing it. Thank you all!

References

[1] L.C. van der Gaag, A.J. Feelders (Eds.), Proceedings of the 7th European Workshop on Probabilistic Graphical Models (PGM 2014), Lecture Notes in Artificial Intelligence, vol. 8754, Springer-Verlag, 2014.

Silja Renooij
Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands
E-mail address: S.Renooij@uu.nl

22 September 2015
Available online 28 September 2015