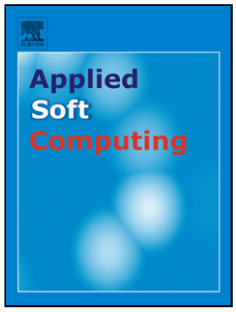


Title: Automated Extraction of Fragments of Bayesian Networks from Textual Sources

Author: Marcello Trovati Jer Hayes Francesco Palmieri Nik Bessis

PII: S1568-4946(17)30411-8
DOI: http://dx.doi.org/doi:10.1016/j.asoc.2017.07.009
Reference: ASOC 4337

To appear in: *Applied Soft Computing*

Received date: 1-12-2016
Revised date: 1-6-2017
Accepted date: 5-7-2017

Please cite this article as: Marcello Trovati, Jer Hayes, Francesco Palmieri, Nik Bessis, Automated Extraction of Fragments of Bayesian Networks from Textual Sources, *<![CDATA[Applied Soft Computing Journal]]>* (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.009

# Automated Extraction of Fragments of Bayesian Networks from Textual Sources

Marcello Trovati

*Department of Computer Science, Edge Hill University, UK*

Jer Hayes

*IBM Research, Dublin Lab, Ireland*

Francesco Palmieri

*Department of Computer Science, University of Salerno, Italy*

Nik Bessis

*Department of Computer Science, Edge Hill University, UK*



**Abstract**

Mining large amounts of unstructured data for extracting meaningful, accurate, and actionable information, is at the core of a variety of research disciplines including computer science, mathematical and statistical modelling, as well as knowledge engineering. In particular, the ability to model complex scenarios based on unstructured datasets is an important step towards an integrated and accurate knowledge extraction approach. This would provide a significant insight in any decision making process driven by big data analysis activities. However, there are multiple challenges that need to be fully addressed in order to achieve this, especially when large and unstructured data sets are considered.

In this article we propose and analyse a novel method to extract and build fragments of Bayesian Networks (BNs) from unstructured large data sources. The results of our analysis show the potential of our approach, and highlight its accuracy and efficiency. More specifically, when compared with existing approaches, our method addresses specific challenges posed by the automated extraction of BNs with extensive applications to unstructured and highly dynamic data sources.

The aim of this work is to advance the current state-of-the-art approaches