Influence Diagrams on R

J.A. Fernández del Pozo[1,*] , C. Bielza[1]

1. [1]Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Madrid, Spain

* Contact author: jafernandez@fi.upm.es

Keywords: Machine Learning, Probabilistic Graphical Models, Bayesian Networks, Classification, Decision Making.

We show, on the intersection of machine learning and decision analysis, a R package for development of probabilistic graphical models like influence diagrams and bayesian networks, two popular knowledge representation tools to deal with decision making and classification problems. *IdR* is a package we have developed to support some issues about tasks related to the influence diagrams building, inference,. . . .

We represent the decision model using a R scritp which is loaded into the environment as a variable-node list. We have implemented several evaluation algorithms from literature, descriptive tools and query procedures to run the decision model. Also we export the model to other environments. First, we can export to the *GeNie* (http://genie.sis.pitt.edu/, ©Decision Systems Laboratory, 2005-2007. All Rights Reserved) application the models to use its graphical interface among others features. Second, the results we will obtain can be retrieved on the *KBM2L* (http://www.dia.fi.upm.es/ jafernan/research/kbm32class.zip in java) system to perform explanations about the reasoning and sensitivity analysis. The idea is to concentrate on the core tasks (problem representation, model building, variable description and inference) and to develop in the future packages for graphical interface, explanation and other issues.

We suggest this package as educational resource, because it is easy to define the model and all its components and to perform analysis breakdown for the evaluation process. The evaluation output is the operations sequence performed over the model, and optionally the optimal decision tables and the posteriory conditional probability tables. Generally, we need to evaluate the model and to analyse the results (correctness, sensitivity, explanation,. . .). We point that the *IdR* package is useful for sensitivity analysis and simulation because the representation is clear, open and it is part of a powerful statistical environment.

An overview of the *IdR* package is summarizes on: we can define models with discrete domain variables like regular influence diagrams (with one utility node) and bayesian networks; *IdR* uses the *lattice* and *cluster* packages to analyse the conditional probability tables; we can summarize the graph, the strength of probability relationships and the conditional independences between nodes; the evaluation of influence diagrams and bayesian networks can be perform numerical (exact or rough) and qualitative way; also the evaluation can be full or partial, by mean of instances on any subset of variables; *IdR* uses the vectorization on its implementation of the Bayes rule and expected utility maximization, two complex operations involved on probabilistic inference.

Finally, we will try to implement more general decision networks (continuous variables, several utility nodes, non sequence decision nodes,. . .), alternatives to the conditional probability tables (linear models) and utility tables (multiattibute utility functions), evaluation algorithms, and learning algorithms from data for the bayesian networks. We are interesting on paralell evaluation of huge models, using packages like *snow*, i.e. very large decision sequences. We need also high computational power when try to perform global sensitivity analysis parametrized with tens of continuous variables.

References

Fernández del Pozo, J.A. (2008). Influence diagrams on R, `http://www.dia.upm.es/~jafernan/research/IdR.html`.

Fernández del Pozo, J.A. and Bielza, C. and Gómez, M. (2005). A List-Based Compact Representation for Large Decision Tables Management. European Journal of Operational Research, 160, Special Issue on Decision Making and AI, 638–662.

Bielza, C., Fernández del Pozo, J.A., Lucas, P. (2008) Explaining clinical decisions by extracting regularity patterns, Decision Support Systems, 44, 397408.