Uncertain Reasoning for Business Rules

Hamza Agli*, Philippe Bonnard,
Christophe Gonzales, and Pierre-Henri Wuillemin

IBM France Lab, 9 rue Verdun Gentilly, France
Sorbonne Universités, UPMC Univ Paris 6, CNRS, UMR 7606 LIP6
{hamza.agli,philippe.bonnard}@fr.ibm.com,
{christophe.gonzales,pierre-henri.wuillemin}@lip6.fr

Abstract. Business rules (BRs) have been widely adopted within decision making processes in the industrial fields (banking, assurance, transport ...). A BR is a high level description allowing non-computer scientists to author and/or make a decision by the use of vocabulary and concepts specific to the organization. It encompasses the business knowledge of experts and separates clearly the business logic from the application logic which implements it by defining and authoring it through a very structured and connected set of applications called a Business Rule Management System (BRMS).
In this paper we propose to investigate the possibility of integration of probabilistic reasoning in a business rules-based system. As a consequence, we can deal with incoherent and incomplete data. Our approach is to extend an object BR model with a probabilistic model. This will be done by coupling business rules and probabilistic engines . The result will allow to perform inferences in Bayesian networks and Probabilistic Relation Models (PRMs) in order to sophisticate the calculations performed in classical BR inference.

Keywords: Business rules, BRMS, Decision Making, Bayesian network, Probabilistic Relational Model (PRM), Bayesian inference

1 General description

URBS (Uncertain Reasoning for Business ruleS) is an applied research project of IBM France which has the ambition to promote the field of business rule-based decision making. It takes into account the uncertainty on data and the temporal aspect within the industrial context. This can be done through an hybridization between business rules and probabilistic graphical models.

* Corresponding PhD candidate in this research project of IBM and University of Pierre et Marie Curie, Paris.

On the one hand, the concept of Business Rule Management System (BRMS) has been introduced more than a decade ago in order to facilitate authoring, checking, deploying and executing the business policy of companies as conceived by their business (as opposed to technical) staff in the form of condition/action business rules. But it turns out that whereas BRMSs are well adapted to deal with structured and complete data by using classical Boolean inference, they face difficulties when they take into account incomplete or incoherent data.
On the other hand, Bayesian Networks (BNs) – which are very popular Probabilistic Graphical Models – were proposed in the late 80s for modeling uncertain knowledge and reasoning with uncertainty. The core of the Bayesian networks representation is a directed acyclic graph (DAG) whose nodes represent random variables and whose arcs represent probabilistic dependencies between these random variables. BNs encode compactly the joint probability over their nodes/random variables as the product of the conditional probabilities of the nodes given their parents in the DAG. Calculation in this graph are done through algorithms that implement probabilistic inference. They were initially used as tools for probabilistic reasoning in early expert systems. The BNs have been applied in several fields in which the decision aiding is a key element. For example: medicine, automatic diagnosis or the robotics.
Today's knowledge-based software systems have to inter-operate. Moreover, the emergence of the "Big Data" processing emphasizes the importance of analytics and probabilistic modeling of data. Hence, adapting the Business Rules to uncertain reasoning is essential. For this reason, this PhD thesis aims at building a bridge between two separate "worlds" by introducing the notion of Probabilistic Production Rules (PPR).
In effect, we propose to specify, implement and validate extensions of current BRMSs to enable them to take into account probabilistic information, which means notably, as mentioned above, dealing with incomplete and/or incoherent data. We will also investigate valid extensions of BNs dealing with rules, implementing these extensions and validating them on real-world decision-making applications. Once coupled with Probabilistic Relational Models (PRM), the BRs will benefit from the improvement of the inference quality as they will be able to apply both Bayesian and Boolean inferences. Since PRMs are an object-oriented extension of BNs, their object oriented nature seems to be very well suited for combining them with BRs.
The notion of PPR has been recently implemented as an intern prototype and it allows the generation of the probabilistic model, which is currently a BN, via annotations of the BR object model (java classes) and/or an externally located Bayesian network. Unfortunately, the current approach is rapidly limited whenever we tackle complex conditions because BNs show their limits when dealing with very large-scale models, which are difficult to create, maintain and utilize. These problems shall be dealt with by combining BRs with PRMs instead of with BNs. Actually, by essence, PRMs are designed to cope with very large systems.

It is important to highlight the difference between the paradigms on which BNs and PRMs on the one hand, and BRs on the other hand are built on. Actually, while the first ones are "descriptive", the BRs paradigm is more procedural and keeping them first separated in a weak coupling approach, will facilitate the management of every component. Indeed, putting a procedural model inside a BN will downgrade its mathematical base. Decision making at that level are usually poor compared to BRMSs level. In addition, trying to manage BNs inference and update inside BRMS will be inefficient and difficult since our rule engine is not monotonic and is by essence a procedural engine on object data. Maybe in a tight coupling we will be oriented to change rules semantic and add more constraints in the BRs model.

2 The plan

In this PhD research project, we propose to couple PRMs and BRs in order to tackle the limitations mentioned above. We also plan to introduce, some advanced temporal concepts into probabilistic models. It is clear that we are looking for creating an effective hybridization of PRMs and BRs. In order to validate this approach, we are developing a new prototype based on this engine coupling called Bayesian Insight System (BIS).
We propose the following research directions:

1. In the first step, we are currently making a comparative study of the language expressivity of both PRMs and BRs. This will enable us to create a common language that holds the advantages of every paradigm. This comparison will point out the formal probabilistic models and inference algorithms that are the most useful for our study.
   Notably, it must be stressed that we deal with temporal probabilistic models and the processing of stochastic events.
   Moreover, our study should also include formal specification of the BNs, PRMs, OOBNs, BLOG models and a comparison with other logics like fuzzy logic, CRF and uncertain ontology.
   Finally, based on the results of our study, our approach will consist of extending current BRs by combining them with the best suited uncertainty models and their inference engines. To characterize the best models, we will exploit Use Cases to determine empirically the expressivity of our new language and to help designing it in a most efficient way. We will ensure that we keep the good properties of this expressivity with regards to the utilization and the efficiency of pattern matching algorithms which are an essential component of BRs.
   This will, precisely, lead us to perform two main extensions:
   (a) extend the object data model of the IBM BRMS, Operational Decision Manager (ODM), called XOM (for eXecution Object Model). The point is to practically use two models and keep their own logic: one linked to the application logic and the other is associated to the uncertain reasoning.

We demonstrate that O3PRM [15] (for Open Object Oriented PRM), which is a programming language that models and specifies PRMs using a strong object-oriented syntax, and the XOM have some good interaction properties and thus will be at the core of this extension.

(b) extend the O3PRM model to handle incremental operations, improve its query process and manage temporal evolution.

2. In the second part, we intend to implement and validate this hybrid language by designing a new prototype of probabilistic business rules engine, Bayesian Insight System prototype, which will use the open source C++ library, aGrUM [1]. This part contains:
   - experiments on how to integrate these prototypes in our BRMS environment and also on the coupling with existing Bayesian engines (SPSS, JSmile, Bayesia, ProBT..)
   - exploration of tight coupling techniques of algorithms developed for the new paradigm and defined in the first part.
   - introduction of the probability calculation in the BRs' temporal expressions, namely the Complex Event Processing.
   - one can also use the probabilistic data to implement the learning of probabilistic models and some parts of the rules within this framework.
   - another important aspect to explore is modeling helpers (aiding): IDE graphical interface, high level language etc.. are very interesting tools when dealing with modeling performed by business experts.
3. In the third part of this research, we will elaborate from the theoretical structure defined in the first stage and from its expressivity and its constraints, specific algorithms to deal with our hybridization to improve the time and size complexity of calculations. We should extend also the BR languages definition to probabilistic concepts. Which means that we will develop a natural and intuitive formulation of the stochastic notions for end-users (non-computer scientists) as well as for the experts of the domain.

The overall research work of this PhD will be validated and oriented through the utilization of some Use Cases from the industrial fields exploiting the collaborations of the laboratory and the company. For instance,

- Fraud detection: this use case proposes to introduce a probabilistic approach to model fraud-control issues. It aims to estimate the likelihood of a potential tax fraud, with the objective either to improve detecting and filtering "false positives", or to better identify "true positives".
- Gathering information about public data available on the Web: introducing probabilistic inference and evaluation in the rules is likely to reduce "false positive" detections in the web-site identification and improve the quality of the meta-data extraction
- Decision Support System diagnosis: rendering the DSS autonomic from costly human expert supervision.

[1] aGrUM, for a Graphical Universal Model, is an open source software that implements O3PRM

References

1. L.Torti, P.Wuillemin, "O3PRM Language specification", Technical report UPMC, 2013
2. B.Berstel–Da Silva, "Verification of Business Rules Programs", PhD thesis, Uni Freiburg, Germany 2012
3. L. Torti, C. Gonzales, and P. Wuillemin, "Speeding-up structured probabilistic inference using pattern mining" International Journal of Approximate Reasoning, 900–918, 2013
4. L. Torti, "Structured Probabilistic Inference in Object-Oriented Bayesian Networks," PhD Thesis ,2012, UPMC, France.
5. P. Weber, G. Medina-Oliva, C. Simon, and B. Iung, "Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas," Engineering Applications of Artificial Intelligence, vol. 25, iss. 4, pp. 671-682, 2012
6. B. Iung, G. Medina-Oliva, P. Weber, and E. Levrat, "Using probabilistic relational models for knowledge representation of production systems: A new approach to assessing maintenance strategies," CIRP Annals – Manufacturing Technology, vol. 61, iss. 1, pp. 419-422, 2012.
7. L. Torti, C. Gonzales, and P. Wuillemin, "Patterns Discovery for Efficient Structured Probabilistic Inference," in Proceedings of the Fifth International Conference on Scalable Uncertainty Management, 2011.
8. C. Gonzales and P. Wuillemin, "PRM inference using Jaffray Fays Local Conditioning," Theory and Decision, vol. 71, iss. 1, pp. 33-62, 2011.
9. I. Siskeviciute, "Introduction of temporal probabilistic reasoning in business rules management system" . 2012
10. J. Pearl.Probabilistic Reasoning in Intelligent Systems: "Networks of Plausible Inference" . San Mateo, CA., 1988
11. K.R. McNaught and A. Zagorecki. Using dynamic bayesian networks for prognostic modelling to inform maintenance decision making. In Industrial Engineering and Engineering Management . IEEE IEEM, 2009
12. H.Ait Kaci, P.Bonnard, "Probabilistic Production Rules", Technical report
13. Cohen T, S Chvaneveldt R.& W Iddows D. "Reflective random indexing and indirect inference : A scalable method for the discovery of implicit connections. Biomed Inform 43 iss. 2, 240–256"
14. A.ElGhali, P.Bonnard, K.ELGhali, D.Hromada, H.Ait Kaci, "Règles de production et réseaux bayésiens pour l'extraction de mots clés, Technical report.
15. Web site for the Open Object Orientd PRM http://o3prm.lip6.fr/
16. Web site for aGrUM : https://forge.lip6.fr/projects/agrum/wiki

Hamza AGLI

9 Rue Verdun
94250 Gentilly, France
Phone: +33(0)149 08 36 87
Email: hamza.agli@fr.ibm.com

Date of birth: 07 June 1988
Nationality: Moroccan
Marital status: Married

Software engineer & PhD Candidate

Educational Background

2012 – 2013 Université Paris Dauphine- École des mines de Paris, Paris, France.
M.Sc (Research Master) in Operations Research and Decision Aiding .
▷ *Modeling, Optimization, Decision & Organization.*

2009 – 2012 Toulouse Institute of Technology (Enseeiht), Toulouse, France.
Engineering Diploma (M.Sc Degree equivalent).
▷ *Applied Mathematics & Computer Science.*

2011 – 2012 Toulouse Institute of Technology (INP), Toulouse, France.
M.Sc (Master Research) in Computer Science and Telecommunication.
▷ *Distributed systems and Critical Software Program.*

2007 – 2009 Lycée Ibn Taimia , Marrakesh, Morocco.
Two-year intensive undergraduate course to prepare for competitive entrance to National Engineering Schools.
▷ *Classes Préparatoires aux Grandes École (MPSI/MP*).*

Experiences

2014-..

Feb-.. IBM France Center for Advanced Studies.
LIP6, Computer science lab of UPMC and CNRS, Paris area, France.
PhD researcher & software engineer in Business Rules and Probabilistic Graphical Models.

2013

May-Oct (5 months) ⋆ Siemens - Mines ParisTech, Paris, France.
Logistic decision aiding.
– Forecasting the European Silicon Valley (Plateau de Saclay) logistic demand within the Grand Paris project framework.
– Introducing a new mean of transportation : Modeling & analysis of the impact on the supply chain of an urban freight system.
– This solution will integrate the freight in the infrastructure of the fully automated Great Paris Métro.

2012

Apr-Sept (6 months) Inria, Sophia Antipolis, France.
Orbital transfer & the low thrust 3 or 2-body problem : Optimization, model, simulation, optimal control.

Jan-Mar Enseeiht, Toulouse, France.
Numerical resolution for a medical imaging contrast problem : Optimal control, study and evaluation of three solvers performances.

2011

June-Oct (3 months) Total SA {Jean Féger Scientifc and Technical Center}, Pau, France.
Advanced optimization algorithms and development techniques: inversion optimization problem.
Data filtering for 2D NMR inversion of real log data from hydrocarbon wells, including test and computation performances and using wavelet transform.

2010 ⋆ Forum Toulouse Technologie, Toulouse, France.
Organization and prospection team.

⋆ Enseeiht, Toulouse, France.
Project manager at ExpoLangue exhibition at Enseeiht.

⋆ Heidelberger Pädagogium, Heidelberg , Germany.
One-month language and cultural stay/ Intensive summer courses.

⋆ Mentor Teaching Maths, Physics and Arab language (Enseeiht, private tutoring).
⋆ Commercial officer (stocktaking).

Fig. 1.