2019 Second International Conference on Artificial Intelligence for Industries (AI4I)

How to increase effectiveness of inference in rule-based systems

1st Andrzej Macioł
Department of Applied Computer Science, Faculty of Management AGH Univeristy of Science And Technology
Kraków, Poland
amaciol@zarz.agh.edu.pl

2nd Stanisław Jędrusik
Department of Applied Computer Science, Faculty of Management AGH Univeristy of Science And Technology
Kraków, Poland
jedrusik@zarz.agh.edu.pl

3rd Andrzej Paliński
Department of Applied Computer Science, Faculty of Management AGH Univeristy of Science And Technology
Kraków, Poland
palinski@zarz.agh.edu.pl

***Abstract*— The control of the inference process is an important, though not always noticed, problem of using rule-based systems built with the use of knowledge-based approach in practice. Therefore, we have developed an inference algorithm allowing to minimize the costs of data acquisition. The algorithm was patented in the USA. The aim of this paper is to present the principles of the developed algorithm and the advantages of using it in scientific research and business practice. The performance of the algorithm is illustrated by using an example of a banking scoring system.**

Keywords*— *rule-based systems, Business Rules Management Systems, reasoning effectiveness, scoring method

I. Introduction

There are two main approaches to the problem of knowledge acquisition in rule-based systems: data driven approach, based on declarative knowledge derived from observation, and knowledge based approach, where knowledge is acquired from domain experts. Recently it has been assumed that the data driven knowledge acquisition approach [1] is much more effective. Clearly, the use of a data-driven approach requires that at least a statistically significant number of patterns (observations) is available. However, there are business and research problems requiring a declarative knowledge approach, for which the data-driven knowledge acquisition approach cannot be applied. These include: designing of new technological, organizational, business, technical (e.g., new materials), etc. solutions; anticipating the behavior of people, communities, organizations, when introducing new ways of interaction or when we have only partial data; designing new solutions based on simulation. For obvious reasons (the source of knowledge are not machine learning methods, but humans judgments), declarative knowledge approach is also applied in Business Rules Management Systems (BRMS). The use of a knowledge-based approach to knowledge acquisition has specific consequences for the reasoning mechanisms. The data-driven approach is used by default for black box inference engines, whereas for white boxes (the knowledge base in the form of IF...THEN rules), the reasoning must be based on classical Horn clauses (application of modus ponens) and one of the strategies of chaining (forward, backward or mixed). Also, in the case of black box systems, the reasoning strategy is imposed by the tool used to acquire knowledge and remains beyond the reach of the user, whereas in the case of classical rule-based systems, the reasoning strategy decides about their effectiveness. Already in the 1970s it was noticed [2] that in the case of large knowledge bases, the key problem is the execution time of the inference process. In many cases of applying the rule-based systems (especially BRMS), knowledge about the facts is not given a priori and must be acquired. In such cases, we face the problem of the cost of data acquisition. The concept of the cost of data acquisition may be understood literally (e.g., assessment of credit worthiness in business problems, ordering laboratory tests in technological issues) or indirectly (e.g., time required to perform the necessary analysis, free but time-consuming consultations with experts). The improvement of inference engines, aiming at minimize the costs, is unfortunately not a commonly perceived problem. However, our experience, related to the implementation of the rule-based systems (BRMS in particular) in business and technological practice, indicate that without considering this problem it is not possible to efficiently apply rule-based systems in business and technology.

The aim of our research was to develop methods and algorithms allowing for significant increase in the effectiveness of rule-based systems, especially BRMS. The effect of this research – the original inference algorithm – was patented in the USA [3].

II. Effectiveness of Reasoning in Rule-Based Systems in Literature

Unfortunately, the problem of improving the effectiveness of the rule-based systems is not widely described in the literature. Most studies focus on optimization of the inference process in networks of rules by applying forward inference, and in later works, also backward. The Rete algorithm, developed by C.L. Forgy [2], should be mentioned as the first. There are also known modifications of this algorithm: object oriented version of Rete – ReteOO [4], Rete II, Rete III, ReteNT), and other competitive solutions, e.g., TREAT [5], LEAPS [6], PHEREAK [7]. These algorithms are applicable if the set of facts is given a priori, however they are not useful if we want to improve the efficiency of inference engines based on other principles of acquiring knowledge about facts. There exist a few studies that refer to the evaluation of the effectiveness of forward versus backward chaining, taking into account the time or number of solutions in solving specific problems [8] and static optimization of decision trees [9], but this also does not solve the problem of controlling the inference process. One can also find in the literature some general considerations on the reasoning mechanisms in the rule-based systems, but no principles of their optimization are described [10].

III. Method for Improving Inference Process

Our experience comes from the implementation of an original BMRS called Rebit. The Rebit system is based on the propositional calculus extended through the use of variables and functions of these variables, where the functions are

978-1-7281-4087-2/19/$31.00 ©2019 IEEE
DOI 10.1109/AI4I.2019.00035

understood just as in algorithmic languages (rather than in predicate calculus). Similarly as in other rule-based systems based on FOL, in order to prove the fact that appears in the conclusion of a rule, the respective premises of the rule must

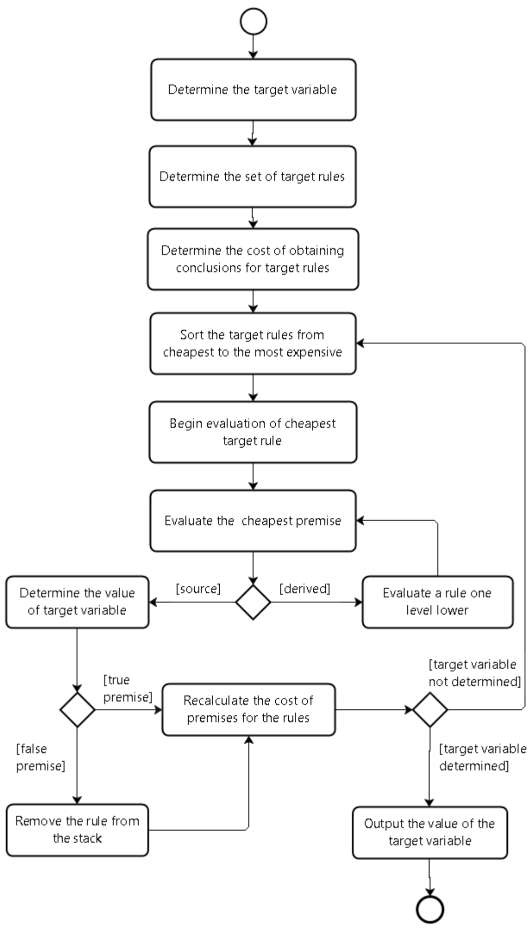

Fig. 1. Reasoning algorithm minimazing cost of data acquisition

be satisfied. In the Rebit system, the logical value of a premise depends on the values of variables or/and functions. The values of variables can be obtained from an environment or can be determined by the inference mechanism on the basis of the relevant rules. In the analyzed case, the general goal of the inference mechanism is to determine the value of some target variable. In order to achieve this goal, we must know the values of all variables used in the inference chain. The cost of obtaining the value from the environment can be determined directly. However, in the case of derived variable values, the cost of obtaining them depends on the cost of determining the values of all variables involved in the deduction process. In the case of the knowledge model used in the Rebit system, variables appear directly in the rules as components of premises, but also as arguments of functions in premises and arguments of all functions nested in them. The Rebit system also uses functions on the conclusion side. When calculating the costs associated with the evaluation of a rule, the cost of obtaining the values of all variables that are arguments of the conclusion function and all functions nested in it should be taken into account.

The key issue now is to decide which variable values must be first determined at each inference step. Fig. 1 presents a diagram of how our inference system works. In the first step, the user determines the target variable. Then, one selects the set of rules that are required to determine the value of the target variable (just like it is in the case of the backward chaining). Next, for each rule, the cost of determining its conclusion is computed according the procedure depicted in Fig. 2.

We take the sum of costs of obtaining all values of the variables involved in the considered rule. The inference system checks whether the given variable is an input variable or a derived variable. The cost of an input variable is read from the database. In the case of a derived variable, we compute (by using iteratively the procedure from Fig. 2) the cost of all rules determining the value of this variable, and finally we take the average of the computed costs. The process continues until the value of all involved variables are verified. Once the costs of all rules, analyzed at this stage, are determined, they are sorted in increasing cost order and put on the stack of tasks for further processing by the inference engine. The inference engine evaluates successive rules according to the predetermined order. The premises of each evaluated rule are assessed in the order of the lower cost. This means that if the premise contains a source variable, its cost is directly determined. If the premise contains a derived variable, it evaluates the rule that assigns value to that variable. If a premise is not true, the rule is removed from the stack. Next,

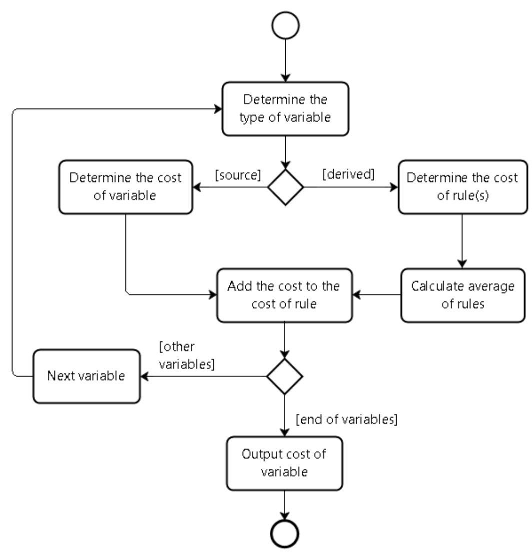

Fig. 2. Algorithm for calculating the cost of rules

the costs of all variables and all rules related to the recently evaluated variable are modified by setting the cost of this variable to zero (as it was already evaluated). Next, we update the order of the rules and continue to the analysis for the rule with the lower cost. If the value of the target variable has been determined, it is returned as the result of the inference process.

IV. Example of application of the method

Below we present the description of the algorithm managing the inference process in scoring analysis, i.e., in the assessment of creditworthiness of enterprises carried out by banks.

Any bank grants loans only to creditworthy persons. Therefore, before granting a loan, a bank analyzes the risk associated with loan default. One of the possible ways to asses a potential borrower is to make a quantitative assessment based on a point evaluation of the customer's characteristics, the so-called *scoring model*. Scoring models for evaluation of credit risk with respect to retail loans have been developed since the 1970s. Initially, scoring models were developed using statistical methods such as logistic regression and discriminant methods [11]. The enormous progress in the research on machine learning over the last 20 years has contributed to an increase in the importance of machine learning based methods in the process of building scoring models [12].

Scoring systems can obviously play only a supplementary role, as the creditworthiness of a company depends not only on its current financial condition, but mainly on the growth prospects of the company and the industry in which it operates. Hence, market and product/service analysis offered by the company play an important role in the credit risk analysis process.

Market analysis requires hiring a marketing specialist and is associated with significant costs, so, whenever possible, a bank should attempt to conduct the risk assessment based on own data.

Another important element of credit risk analysis is to determine the value of the collateral and the reliability of the guarantor in the case of personal collateral. A borrower holding collateral with a high liquidation value signals low credit risk [13]. At this stage of the analysis hiring an expert may also be required, especially when evaluating a property. The cost of this stage is high but lower than market research.

Comparing the scoring models for bulk retail loans with the rule models developed for individual business risk analysis, attention should be paid to the problem of cost in the process of model learning and reasoning. The construction of scoring models is related to the classical problem of unbalanced datasets, in which the minority class is a "positive" one [14]. Risk-bearing borrowers are a minority class in the learning set, and therefore they might be classified improperly, which in turn may result in significant losses for the lender. One of the methods of dealing with this problem, making the learning process safe, is to assign a high cost to an improperly classified risky borrower.

In the reasoning mechanism presented in the article, the situation is completely different. The rule model is built on the basis of expert knowledge, and the result produced by this model depends on the data obtained in the inference process. Since data acquisition is associated with different costs, hence the inference system attempts to minimize the cost when evaluating the rules. Such situation fully corresponds to the credit risk valuation related to individual business entities. The cost problem does not arise in the model learning process, but in the model implementation process.

Below, we present a simplified model of credit risk analysis with the use of a rule-based system. The complete credit risk evaluation system differs from this simplified system by the greater number of financial ratios, detailed valuation of each piece of collateral, more advanced analysis of the market situation, and, additionally, by the assessment of the enterprise management system. Simplified rules take into account the essence of credit risk assessment and the estimated scale of costs of external data acquisition. Internal data held by a bank are data contained in the credit application forms, which are mainly financial data verified by an auditor. It is assumed that the cost of this data is 1. The cost of expert's assessment of the value of a collateral is 3, while the cost of evaluation of guarantor's credibility is 2. It is highly expensive to assess the market position of a company. The assessment of the degree of dependence on customers, as well as the analysis of signed contracts with customers and the assessment of the degree of dependence on suppliers reach the cost of 2. The cost of analysis of market share is 3.

Figure 3 shows the hierarchy of the rules, and Table 1 shows the selected rules together with their cost given in parentheses. The premises of the rules being the conclusions of other rules, have a cost equal to the arithmetic mean of the costs of the premises of these rules.

In the considered case, the target variable is "Credit decision", so the system should start from the cheapest rule with the target variable in the conclusion, i.e., Rule 5 (4.71). As can be seen, its premise contains the variable "Collateral". The cheapest rule, whose conclusion contains this variable, is Rule 8 (3). In order to evaluate it, the value of collateral with a cost of 3 should be valued. If the value of the variable "Collateral" is "high", then the premise of the rule will not be met and the system will start evaluation of another cheapest rule, Rule 3 (6.75), which requires checking the value of the variable "Economic condition".

First of all, Rule 14 (2) will be activated, which only requires checking the value of the variable "Liquidity ratio". If the value of the variable "Liquidity ratio" is lower than the critical value included in the condition, we immediately obtain the credit decision "deny".

The proposed cost-based reasoning system avoids in many cases "expensive" rules that require verification of the values of the variables "Collateral" and "Market position". Moreover, decisions leading to a denial of granting a credit turn out to be "cheap" so that a bank can avoid real high sunk costs. Positive decisions will allow higher costs of risk analysis to be covered by commission or interest margin charged to the customer after the loan has been granted.

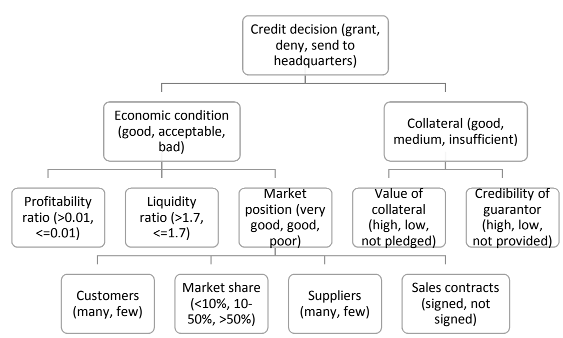

Fig. 3. Hierarchy of rules used in the scoring model

TABLE I. LIST OF RULES

RULE 1: (11.46) IF Economic condition = "good" (6.75)
AND Collateral = "good" (4.71)
THEN Credit decision = "grant"

RULE 2: (11.46) IF Economic condition = "good" (6.75)
AND Collateral = "sufficient" (4.71)
THEN Credit decision = "grant"

RULE 3: (6.75) IF Economic condition = "bad" (6,75)
THEN Credit decision = "deny"

RULE 4: (11.46) IF Economic condition = "acceptable" (6,75)
AND Collateral = "sufficient" (4.71)
THEN Credit decision = "send to headquarters"

RULE 5: (4.71) IF Collateral = "insufficient" (4.71)
THEN Credit decision = "deny"

RULE 6: (11.46) IF Economic condition = "acceptable" (6,75)
AND Collateral = "good" (4.71)
THEN Credit decision = "grant"

RULE 7: (5) IF Value of collateral = "low" (3)
AND Credibility of guarantor = "low" (2)
THEN Collateral = "sufficient"

RULE 8: (3) IF Value of collateral = "high" (3)
THEN Collateral = "good"

RULE 9: (5) IF Value of collateral = "not pledged" (3)
AND Credibility of guarantor = "high" (2)
THEN Collateral = "good"

RULE 10: (5) IF Value of collateral = "not pledged" (3)
AND Credibility of guarantor = "low" (2)
THEN Collateral = "insufficient"

RULE 11: (5) IF Value of collateral = "low" (3)
AND Credibility of guarantor = "not provided" (2)
THEN Collateral = "insufficient"

RULE 12: (5) IF Value of collateral = "not pledged" (3)
AND Credibility of guarantor = "not provided" (2)
THEN Collateral = "insufficient"

RULE 13: (5) IF Value of collateral = "low" (3)
AND Credibility of guarantor = "high" (2)
THEN Collateral = "sufficient"

RULE 14: (2) IF Profitability ratio <= 0.01 (1)
AND Liquidity ratio <= 1.7 (1)
THEN Economic condition = "bad"

…

RULE 32: (9) IF Customers = "many" (2)
AND Market share = "<10%" (3)
AND Suppliers = "many" (2)
AND Sales contracts = "signed" (2)
THEN Market position = "very good"

V. Conclusions and Further Works

The presented algorithm has been tested many times. The results show that the proposed solution strongly reduces the cost of data acquisition. All of these experiments concerned problems with a relatively small number of rules, which is specific for technological problems and BRMS systems. How this algorithm scales when the number of rules increases requires further research.

The results of our research proved that developed algorithm for controlling the inference process in hierarchical rule-based systems can be successfully used in business practice, where the order and cost of confirming the facts is very important. The use of our patented solution creates new opportunities for rule-based systems, especially for BRMS.

However, until now there is still no solution to the problem of inference control in the case of hierarchical fuzzy reasoning systems (this does not apply to BRMS of course). In the case of fuzzy reasoning systems, the cost of acquiring knowledge about certain facts is related to the significance of this premise in establishing the logical value of the conclusion. The algorithm presented in our patent cannot be used directly to solve this problem. The specificity of fuzzy reasoning systems inference is also noticed by researchers dealing with Rete and similar algorithms (e.g. [15]). The aim of our further research will be to extend the presented algorithm by formulas allowing to take into account the evaluation of the value of facts from the point of view of the effectiveness of the control of the inference process.

Acknowledgment

This work is financed by grants for the maintenance and development of research capacities.

References

[1] E. Hüllermeier, "From knowledge-based to data-driven modeling of fuzzy rule-based systems: A critical reflection." [Online]. Available: https://www.semanticscholar.org/paper/From-knowledge-based-to-data-driven-modeling-of-A-Hüllermeier/282e10aed878aa25def32a67938834904c2ac1fa. [Accessed: 12-Apr-2019].

[2] C. L. Forgy, "Rete: A fast algorithm for the many pattern/many object pattern match problem," *Artif. Intell.*, vol. 19, no. 1, pp. 17–37, 1982.

[3] S. Jedrusik and A. Maciol, "Rule-based inference system and an inference method," US9489628B2, 2013.

[4] "ReteOO Algorithm," *jboss.org Community Documentation*. [Online]. Available: https://docs.jboss.org/drools/release/6.5.0.Final/drools-docs/html/ch05.html#ReteOOSection. [Accessed: 30-Apr-2019].

[5] D. P. Miranker, "TREAT: A Better Match Algorithm for AI Production Systems; Long Version," University of Texas at Austin, Austin, TX, USA, 1987.

[6] D. Batory, "The LEAPS Algorithm," University of Texas at Austin, Austin, TX, USA, 1994.

[7] "5.1. PHREAK Algorithm," *Red Hat, Drools*. [Online]. Available: https://access.redhat.com/documentation/en-us/red_hat_jboss_bpm_suite/6.2/html/development_guide/chap-rule_algorithms#PHREAK_Algorithm. [Accessed: 06-Mar-2019].

[8] S. Kamley, S. Jaloree, and R. S. Thakur, "Performance Comparison between Forward and Backward Chaining Rule Based Expert System Approaches Over Global Stock Exchanges," *Int. J. Comput. Sci. Inf. Secur.*, vol. 14, no. 3, pp. 74–81, 2016.

[9] A. Nowak-Brzezińska, R. Simiński, and A. Wakulicz-Deja, "Inference Algorithms for Hierarchical Knowledge Bases," *Recent Advances in Intelligent Information Systems*, 20-Jun-2010. [Online]. Available: https://www.researchgate.net/publication/255029073_Inference_Algorithms_for_Hierarchical_Knowledge_Bases. [Accessed: 16-Mar-2019].

[10] A. Nowak-Brzezińska, "Enhancing the Efficiency of a Decision Support System through the Clustering of Complex Rule-Based Knowledge Bases and Modification of the Inference Algorithm," *Complexity*, vol. 13, pp. 1–14, Dec. 2018.

[11] D. Andriosopoulos, M. Doumpos, P. M. Pardalos, and C. Zopounidis, "Computational approaches and data analytics in financial services: A literature review," *J. Oper. Res. Soc.*, pp. 1–19, May 2019.

[12] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *Eur. J. Oper. Res.*, vol. 247, no. 1, pp. 124–136, 2015.

[13] A. Palinski, "Loan Payment and Renegotiation: The Role of the Liquidation Value," *Argumenta Oeconomica*, vol. Nr 20 (40), pp. 225--252, 2018.

[14] H. He and E. A. Garcia, "Learning from Imbalanced Data," IEEE Trans. Knowl. Data Eng., vol. 21, no. 9, pp. 1263–1284, 2009.

[15] N. Wulff and D. Sottara, "Fuzzy Reasoning with a Rete-OO Rule Engine," *Lect. Notes Comput. Sci.*, vol. 5858, pp. 337–344, Nov. 2009.