THE INFERENCE EFFICIENCY PROBLEM IN BUSINESS AND TECHNOLOGICAL RULES MANAGEMENT SYSTEMS

Barbara Baster and Andrzej Macioł
Faculty of Management, AGH University of Science and Technology, ul. Gramatyka 10, 30-067 Kraków, Poland

Keywords: Business Rules Management Systems, Technology, Rule-based Systems, Inference Efficiency, Databases.

Abstract: In the following paper we present the results of our works related to development of rule engine for automated interpretation of business rules. Our experiences and experimental results show that knowledge description for this purpose may be stored in the form of relational databases. The aim of the experimental research presented in this paper was to determine degree in which organization of knowledge base and assumed inference strategy influence the efficiency of inference process itself. Experiments proved that owing to the application of mechanism characteristic for relational data bases, knowledge base can be easily arranged so as to maximize efficiency of inference. The efficiency of inference is strongly influenced by preliminary knowledge transformation from the set of examples or random rules into arranged form.

1 INTRODUCTION

In the paper we present the results of our ongoing works related to development of rule engine for automated interpretation of business rules. Our researches are connected with the project cofounded by the European Union. The works are aimed at development of a tool which will operate in accordance with general BRM idea and will simultaneously include specific technological knowledge.

Recent researches highlight the particular specificity of issues of the borderline of management and technology. Because of this specificity, the business rule management systems that are currently available on the market are very rarely applied in the fields related to process and technology management. One of the factors that contributes to the current state of matter is a necessity to apply hybrid system in this field – connection of rule-based system with procedural computations. Such situation takes place e.g. in case of the technology development for a new product, or process management in emergency situations. In such instances, a rule-based system would control the selection process of conditions that should be verified depending on the results of preceding computations. The verification of individual conditions would be based on execution of computations in accordance with predefined algorithms (procedural computations). Similarly, hybrid systems could support the processes related to the disposal of faulty products.

During research works it was also found that typical decision problems connected with technology issues do not require preparation of complex knowledge models. Nevertheless, the number of rules may in some cases grow to hundreds and complicate the acquisition of knowledge simultaneously raising the requirements above the limits settled by inference systems.

In this publication we present the results of research which was aimed at assessment of possibilities to improve the inference procedure from the point of view of processed queries and time consumption. Moreover, one of our goals is to prove that concept which utilizes relational database as vocabulary carrier gives wide range of possibilities in the field of improvements of knowledge structure and the efficiency of inference process.

2 INFERENCE EFFICIENCY PROBLEM

At present, many forward-chaining systems use inference networks to capably match rule conditions to facts. The most efficient rules engines are designed by the Rete algorithm. It is a widely-used match algorithm for comparing a large numbers of

Baster B. and Macioł A. (2010).
THE INFERENCE EFFICIENCY PROBLEM IN BUSINESS AND TECHNOLOGICAL RULES MANAGEMENT SYSTEMS.
In *Proceedings of the 5th International Conference on Software and Data Technologies*, pages 10-15
DOI: 10.5220/0002931600100015

patterns to a large collection of slowly changed data without iterating over the sets (Forgy, 1982). The algorithm was originally developed by Charles Forgy (1981) for OPS5 and it has soon become the basis for many popular system shells, including for example NASA's CLIPS (Giarratano, 1989), Jess or Drools.

The Rete algorithm is designed to increase the speed of forward-chaining rule systems. Systems based on Rete construct a network of nodes. Each node corresponds to a pattern occurring in the condition part of a rule (the left-hand-side) and contains a memory of facts which fulfil that pattern. When facts are changed, they propagate along the network and cause nodes to be annotated if the fact matches the pattern. The output of Rete is a set of new matched patterns, and a set of formerly matched patterns no longer matching (Lee and Schor, 1990).

Research involving Rete algorithm has been focused on improving its performance, as in (Kang and Cheng, 2004), (Zhou et al., 2008), (Tuttle and Eick, 1991). Tuttle and Eick (1991) propose a generalization of the Rete network, called a historical Rete network which can store and maintain a run's historical information.

In the literature except for Rete we can also find different production rule condition testing algorithms. Wang and Hanson (1993) present one of them. Their paper shows the results of a simulation study comparing the join-condition testing efficiency of Rete and TREAT (a variation of Rete) in the context of a database production rule system. Actually TREAT is very similar to Rete. The way that both algorithms test selection conditions is identical. The difference lies in how they test joins. According to the results of Wang and Hanson's research TREAT algorithm in many ways outperforms Rete. It requires less storage than Rete, and is less sensitive to optimization decisions. Based on these results, authors conclude that TREAT is the preferred algorithm for testing join conditions of database rules, however, they agree that both of these two match algorithms can be successively used to improve the speed of the match process in DBRSs.

3 APPLICATION OF RELATIONAL DATA BASES AS A KNOWLEDGE MODELS FOR BRE

Our experiences indicate that the most effective solution, capable of direct cooperation with majority of industrial information systems which simultaneously provides decidability, is a combination of relational model with inference system that utilizes attributive logics. Our solution, named Inference with Queries (IwQ) (Macio¤, 2008), has been developed as a knowledge model and an inference engine for formulation of Business Rules Management Systems. Knowledge storage is realized in accordance with principles of Variable Set Attribute Logic (VSAL).

An assumption was made that in rule-based decisive system, attributes as well as variables will be stored in form of relations. In our solution, we have introduced data metamodel. Concept of data metamodel comes down to partial transfer of intensional database structure to its extensional part.

Solution based on a relational database and metamodel allows for very simple and flexible development of both data structure and inference process. Owing to that fact final effects resemble those obtained by means of Rete algorithm, yet some specific redundancy characteristic for this solution is excluded.

4 TECHNOLOGICAL PROBLEMS EXAMPLES

Our research works concentrated on the field of metal processing in its broad sense, but their final results may be generalized. Despite vast diversity and uniqueness of business processes that are realized in the analysed field, there exist multiple processes that can be relatively easily automated. Those processes are conducted in a cyclical manner and realized on the basis of relatively simply definable quantitative and qualitative criterions. In the case of non-measurable criterions a common practise is to asses point-like scores. These processes can be described with the use of rule-based systems. The processes which are repeatable and their frequency of realization is relatively high should be automated in the first place.

In spite of the fact that similar problems are solved by means of the BRM systems which are widely available on the market, in our case we face the problem of connecting data that originates from different sources. In dependence of facts' source, we may predict different level of costs, hence the inference strategy becomes an important matter. Institutions that verify contractor's credibility constitute an exemplary case of such 'costly' source of information. It becomes obvious that rules which require such type of data should be avoided at all cost as long as remaining rules allow for final settlement (e.g. in case when quality of material offered by the contractor – supplier is not satisfactory, there is no use in verification of his credibility).

5 EXPERIMENT

5.1 Aim and Assumptions

The aim of the experimental research was to determine the degree in which organization of knowledge base and assumed inference strategy influence the efficiency of inference process itself.

In the first step an analysis was conducted whether structure of knowledge base, understood as number of rules and their connection pattern, influences inference time. BRM systems most often do not provide mechanism of knowledge generation on the basis of examples (which arises from the assumption, that knowledge in rules management is *a priori* defined). In our opinion, in considerations where decision tables are utilized we deal with a set of examples, that can be transformed by means of machine learning methods into set of rules in simpler form than the set corresponding to every column of the table. Owing to this transformation the number of rules can be reduced, which entails streamlining of the inference process.

In our research works we analysed to what extent does the transformation of exemplary sets described below to the form of decision table will allow to increase efficiency of inference. Transformations were performed by means of selected algorithm of decision tree learning (ID3).

Method of knowledge description in the form of database, which was elaborated by our team, allows for extremely simple manipulations on accessibility conditions to the rules and facts, without a necessity of changes in the inference algorithm. Such changes may be introduced by altering the parameters of these SQL queries which pass information to the inference engine. In our research we were looking for the answer to the question whether modifications in access sequence for the rules and facts does increase efficiency of inference in case of such structure of the knowledge base. For this purpose we conducted experiments both for unsystematic knowledge as well as different variants of its systematisation. We did analyse what are the consequences of systematization within the tables storing rules and facts.

In a vast majority of currently available BRM systems a forward chaining inference is used. The inference engine receives either a single piece of information or a set of information items about facts and begins with classic forward chaining. In case when the data is not complete and firing of the rule is impossible, the engine queries for missing premises.

Our system's character resembles inference system in a mixed manner, it works in accordance with classical scheme which combines forward chaining and backward chaining for the search of missing premises. When the general definition of inference target is prepared, a start point must be defined, i.e. attribute which values must be inputted explicitly at the very beginning of the process.

We have made an attempt to define rules that would allow to determine which of the source attributes (that are not the result of inference) should be selected as the starting point. For that purpose we conducted simulations which enabled us to settle the dependencies between selection strategy of starting point and the inference efficiency.

Simulation analysis of inference systems requires preliminary settling of the rules for answering these questions, which constitute a basis for verification of facts correctness. In case of a system that is founded on attribute logic, this process is based on selection of particular feature's value of the attribute from the domain of this feature. In the basic research we made an assumption, that ever value of the feature is equally probable. This assumption corresponds to a situation when a knowledge engineer does not have an opportunity to asses variability of attribute values. These assumptions were taken for initial works.

Because of the fact that feature values of attributes are not uniformly distributed, we have also analysed impact of this phenomena on efficiency of inference process.

5.2 Knowledge Base

The inference schema is shown in Figure 1. It is based on relational database which contains the assessment criteria as it shows Figure 2. The three-degree and two-degree qualitative scale has been used to estimate individual premise (good, average, bad or yes and no).

Sample rule is as follows:

```
Rule:  5
if
Certainty = bad
Quality   = yes
Logistics = good
```

```
then Supplier := average
```

5.3 Research Plan

Procedure of the experiments was aimed at observation of inference process in above-described database for randomly generated responses. Number of queries of the inference engine as well as process duration time were the objects of observation. Since the number of rules was significantly limited, additional retarding element was introduced which prolonged answering time of queries (without it, firing times of about 1 millisecond were observed in extreme cases).

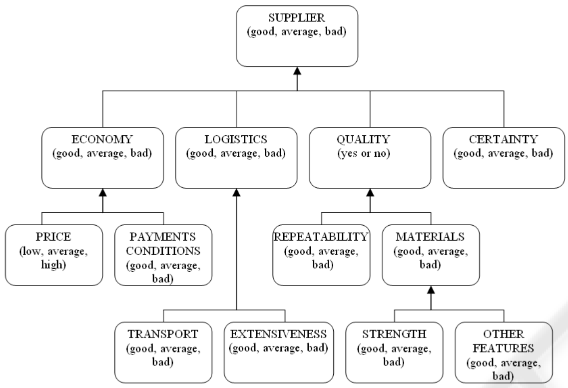

Figure 1: An inference schema.

Supplier - Conclusions				
Economy	Logistics	Quality	Certainty	SUPPLIER
good	good	no	good	average
good	good	no	average	average
good	good	no	bad	average
good	good	yes	good	good
good	good	yes	average	good
good	good	yes	bad	average
good	average	no	good	average
good	average	no	average	average

Figure 2: Criteria of supplier selection (partial).

Considering the aim of the research, the following schedule of experiments was prepared:

1. Inference in a raw knowledge base, in which the number of rules corresponded to the number of examples (E1)
2. Inference in a knowledge base after its transformation into decision tree by means of ID3 algorithm (the number of rules was reduced from 90 to 60) (E2)
3. Inference in the transformed knowledge base and after rearrangement of the rules according to increasing number of premises, for each separate attribute that constituted conclusion, for 3 randomly selected variants of fact arrangement inside the database (E3, E4, E5)

All of the above-mentioned experiments were conducted under assumption of equal probability of each value that an attribute can take.

Moreover, one additional experiment was performed for the most favourable variant among all previously described, but under assumption of dispersed probability distribution of attribute values (E6).

5.4 Experimental Results

Table 1: Comparison of results for the experiments E1 and E2.

Attribute	Number of queries		
	E1	**E2**	**BP***
TRANSPORT	8	7,30	8,75%
EXTENSIVENESS	8	7,16	10,50%
CERTAINTY	8	7,14	10,75%
STRENGTH	8	7,14	10,75%
OTHER FEATURES	**8**	**7,12**	**11,00%**
PRICE	8	7,18	10,25%
PAYMENT CONDITIONS	8	7,36	8,00%
RECURRENCE	8	7,66	4,25%
MEAN	**8**	**7,26**	**9,28%**
	Process duration		
	E1	**E2**	**BP***
TRANSPORT	621,25	580,00	7,11%
EXTENSIVENESS	618,59	563,13	9,85%
CERTAINTY	684,69	621,25	10,21%
STRENGTH	569,06	506,25	12,41%
OTHER FEATURES	**572,97**	**506,56**	**13,11%**
PRICE	624,22	566,56	10,18%
PAYMENT CONDITIONS	619,38	577,81	7,19%
RECURRENCE	624,53	593,13	5,30%
MEAN	**684,69**	**621,25**	**9,42%**

** betterment in percent*
the best betterment result in bold

Transforming the set of examples into non-arranged decision tree results in reduction of number of queries by 9.28% on average and reduces the inference consumption time by 9.42%. Obviously, the number of asked queries in the raw knowledge base is not dependent on the start point. Even after its transformation into decision tree, no influence of the start point selection on the number of queries asked can be noted. On the other hand, selection of the starting point does influence inference time in the raw as well as transformed knowledge base.

In the table 2. an juxtaposition is included, where results of experiments for the raw knowledge base are compared with the mean result obtained after the arrangement of rules for three different arrangement patterns of fact table.

Arrangement of the rules has a very distinct influence on efficiency of inference process. Independently on the selection of start point, a mean time of inference is reduced by 22,38%, and the number of asked queries falls by 19,32%. Also, the most favourable start point is clearly distinguishable. It is the fact which directly co-decides about final conclusion and simultaneously reveals itself to be certainty fact. If the inference is started from this very point, the inference time may be reduced by 23,55% with respect to the best possible time obtained for the raw knowledge base. At the same time, number of asked queries decreases by 30,25% with regard to the maximum possible number.

Table 2: Comparison of results for the experiments E1 and mean result of experiments E2, E4, E5.

Attribute	Number of queries		
	E1	**E3,E4,E5**	**BP***
TRANSPORT	8	6,22	22,25%
EXTENSIVENESS	8	6,30	21,25%
CERTAINTY	8	**5,58**	30,25%
STRENGTH	8	6,97	12,88%
OTHER FEATURES	8	6,88	13,96%
PRICE	8	6,12	23,50%
PAYMENT CONDITIONS	8	6,09	23,88%
RECURRENCE	8	7,47	6,62%
MEAN	**8**	**6,45**	**19,32%**
	Process duration		
	E1	**E3,E4,E5**	**BP***
TRANSPORT	621,25	469,06	24,50%
EXTENSIVENESS	618,59	476,56	23,29%
CERTAINTY	684,69	460,57	25,86%
STRENGTH	569,06	493,54	20,56%
OTHER FEATURES	**572,97**	479,48	22,82%
PRICE	624,22	457,76	26,32%
PAYMENT CONDITIONS	619,38	**438,02**	29,49%
RECURRENCE	624,53	582,60	6,22%
MEAN	**684,69**	**482,20**	**22,38%**

** betterment in percent*
the best result in bold

In the table 3. an juxtaposition is included, which depicts comparison of results for the experiments conducted on the knowledge base with arranged rules for three different arrangement patterns of the fact table.

The results of experiments imply that the arrangement patter of facts does not have significant influence on mean inference time, number of asked queries or the most profitable selection of start fact.

Table 3: Comparison of results for the experiments E3, E4 and E5.

Attribute	Number of queries		
	E3	**E4**	**E5**
TRANSPORT	6,40	6,30	5,96
EXTENSIVENESS	6,40	6,18	6,32
CERTAINTY	**5,42**	**5,62**	**5,70**
STRENGTH	6,91	7,08	6,92
OTHER FEATURES	6,91	6,90	6,84
PRICE	6,08	5,94	6,34
PAYMENT CONDITIONS	6,17	6,10	6,00
RECURRENCE	7,61	7,40	7,40
MEAN	**6,49**	**6,44**	**6,44**
Attribute	Process duration		
	E3	**E4**	**E5**
TRANSPORT	489,69	480,00	437,50
EXTENSIVENESS	496,56	459,06	474,06
CERTAINTY	**447,97**	461,88	471,88
STRENGTH	503,13	493,75	483,75
OTHER FEATURES	489,69	477,50	471,25
PRICE	462,03	435,94	475,31
PAYMENT CONDITIONS	449,06	**435,31**	**429,69**
RECURRENCE	594,06	571,88	581,88
MEAN	**491,52**	**476,91**	**478,16**

** betterment in percent*
the best result in bold

Table 4. contains comparison of average values for different experiments. One for arranged database, conducted under assumption of uniform distribution probability of attribute values and the remaining for the same database, but with dispersed probability.

As it can be clearly seen, the variable character of attribute values does not influence the inference time or the number of asked queries. It may be concluded from the fact, that optimization of knowledge base structure can be executed during its design phase. Thus there is no necessity to analyze the character of described phenomena in a statistical point of view.

6 CONCLUSIONS

Experiments conducted during our research have proven that owing to application of mechanism characteristic for relational data bases, knowledge base can be easily arranged so as to maximize the efficiency of inference. It also became clear that the efficiency of inference is strongly influenced by preliminary knowledge transformation from the set of examples or random rules into arranged form (e.g. form of decision tree). Arrangement of the set of rules as well as selection of the start point are also significant issues. By proper arrangement of the rules and selection of the favourable start point, inference time can be reduced by 23,55% and the number of asked queries decreased by 30,25%.

Table 4: Comparison of the mean result of experiments E2, E4, E5 and experiment E6.

Attribute	Number of queries		
	E3,E4, E5	**E6**	**BP***
TRANSPORT	6,22	6,22	0,00%
EXTENSIVENESS	6,30	5,96	5,40%
CERTAINTY	**5,58**	**5,86**	-5,02%
STRENGTH	6,97	7,00	-0,43%
OTHER FEATURES	6,88	6,74	2,03%
PRICE	6,12	6,06	0,98%
PAYMENT CONDITIONS	6,09	6,72	-10,34%
RECURRENCE	7,47	7,04	5,76%
MEAN	**6,45**	**6,45**	**0,00%**
Attribute	Process duration		
	E3,E4,E5	**E6**	**BP***
TRANSPORT	469,06	464,06	1,07%
EXTENSIVENESS	476,56	**446,25**	6,36%
CERTAINTY	460,57	496,25	-7,75%
STRENGTH	493,54	488,75	0,97%
OTHER FEATURES	479,48	470,63	1,85%
PRICE	457,76	457,81	-0,01%
PAYMENT CONDITIONS	**438,02**	492,81	-12,51%
RECURRENCE	582,60	546,88	6,13%
MEAN	**482,20**	**482,93**	**-0,15%**

** betterment in percent*
the best result in bold

The results presented in this article prove that the possibility to improve the inference efficiency without necessity of pattern matching algorithms application exists. Even though, they should be treated as an introduction to further research on this matter. In the experiments described above, the simulation was applied as a tool for finding of quasi-optimal solutions. Nonetheless, the simulation itself can not be conducted in every specific case. Therefore, the aim of our further research will be to create the rules formulation that will allow to achieve the best parameters for the structure of knowledge base.

ACKNOWLEDGEMENTS

This research has been partially supported by the Innovative Economy Operational Programme EU-founded project (UDA-POIG.01.03.01-12-163/08-01).

REFERENCES

Forgy, C.L. (1982). Rete: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem. *Artificial Intelligence*. 19, 17-37.

Forgy, C.L. (1981). *OPS5 User's manual*. Technical Report. Department of Computer Science, Carnegie-Mellon University.

Giarratano, J. (1989). *CLIPS User's Guide, Version 4.3*, Artificial Intelligence Section, Lyndon B. Johnson Space Center, COSMIC, 382 E. Broad St., Athens.

Kang, J.A., Cheng, A.M.K. (2004). Shortening Matching Time in OPS5 Production Systems. *IEEE Transactions on Software Engineering*, 30.

Lee, H.S., Schor, M. (1990). Dynamic Augmentation of Generalized Rete Networks for Demand-Driven Matching and Rule Updating. *The 6th Conf. of A.I. Applications 1990 IEEE*.

Lee, Y.H., Yoo,S.I: (1995). *A Rete-based Integration of Forward and Backward Chaining Inferences*. Monterey.

Macioł, A. (2008). An application of rule-based tool in attributive logic for business. *Expert Systems with Applications*. 34, 1825–1836.

Tuttle, S.M., Eick, C.F. (1991). *Historical Rete Networks for Debugging Rule-Based Systems*. San Jose, CA (1991).

Wang, Y.W., Hanson, E.N. (1993). *A Performance Comparison of the Rete and TREAT Algorithms for Testing Database Rule Conditions*. Washington.

Zhou, D., Fu., Y., Zhong, S., Zhao, R. (2008). The Rete Algorithm Improvement and Implementation. *2008 International Conference on Information Management, Innovation Management and Industrial Engineering*.