**Paweł Weichbroth**
Gdańsk University of Technology, Gdańsk, Poland
e-mail: pwi@zie.pg.gda.pl

# A FRAMEWORK OF RULE BASED EXPERT SYSTEM FOR MARKET BASKET ANALYSIS

**Abstract:** This paper presents a novel approach to discovering association rules from market basket data. For this purpose the expert system with inference engine, visualisation engine, and knowledge base has been elaborated. After a short introduction, a general description of the system's components is given. Then, the functionality of inference engine, visualisation engine and knowledge base is characterized briefly. The whole work is finished by conclusions and indicating future work.

**Keywords:** expert system, association rules, market basket analysis.

## 1. Introduction

Expert systems have dynamically improved since the time of their first commercial application at the beginning of the 1980s. Today, expert systems are widely used in business, science, construction brand and mechanics but they have also many other applications. Roots of expert systems reach many disciplines, especially the branch of human information processing called cognitive science.

In the late 1950s and early 1960s a couple of programs were written in order to solve key problems. The most popular one was the General Problem Solver [Giarratano, Riley 1998] written by Newell and Simon, described in many articles and afterwards in detail defined in monumental, over 920 pages work *Human Problem Solving* [Newell 1972]. The most insightful thesis presented by Newell and Simon was a statement assuming that a great number of human information or cognition problems can be expressed in the form of a rule If – Then. Let us say, if it is going to rain, take an umbrella with you. The thesis corresponds with small, modular piece of knowledge, called a chunk. Chunks are classified in order which leads to associations among chunks of knowledge.

One theory assumes that human memory is simply organized in chunks. To make it more understandable we can base on a well-known chunk of knowledge: If the unemployment increases, the inflation diminishes in short period (ceteris paribus).

Newell and Simon popularized usage of those rules to represent human knowledge. They also showed how the process of inference may be executed in practice.

Expert systems which are common nowadays have their roots in pioneering works of Feigenbaum, Lederberg, Shortliffe and Buchanan from the late 1960s and the early 1970s [Giarratano, Riley 1998]. Alongside with acceptation of knowledge-based paradigm in the 1970s, a couple of good prototypes of expert systems were developed. Those systems could interpret mass spectrograms to indicate chemical constituents (Dendral) [Lindsay et al. 1993], diagnose disorders (Mycin) [Hajek, Valdes 1994], analyze geological data (Dipmeter) to find a bed of dead oil and minerals (Prospector) and automate computer systems configuration (Xcon/R1) [Giarratano, Riley 1998]. Success of computer systems like Prospector, which discovered beds of precious minerals, and the program Xcon, which significantly decreased the general expenses of the Digital Equipment Corporation (DEC), was the main reason for general growth of preoccupation of expert systems, both in industry branch and academic environment. Authors of Dentral for the first time pointed out that proper inference was strictly dependent on expertise knowledge saved in form of rules but not on complicated techniques of searching and inference [Niederliński 2006].

In the 1970s and 1980s expert systems went through crisis. Some factors influenced it really strongly, first of all the naive and ruinous hopes reposed in functionality of expert systems. Secondly, strong belief that expert system will be a happy medium for the hard times of arms race. As a result states like Great Britain and the US resigned from financing the projects connected with artificial intelligence and projects started to be financed from private budgets. What is more, the collapse of Japanese program of the fifth generation computers, which aims were not strictly defined, conduced to pause in research and development of expert systems [Niederliński 2006].

However, recent years brought out many successes to expert systems in specialist applications like medical diagnosis (medicine), optimal routing (computer network), optimal stocks portfolio (finance), valuation of the debtor and money laundering (banking) [Korczak, Oleszkiewicz 2009; Korczak et al. 2008]. Novel approach is the application of expert system [Orłowski, Sitek 2007] to select and valuation the information technologies [Orłowski et al. 2010] in agent architecture [Orłowski, Ziółkowski 2009] with usage of ontology [Czarnecki 2008, 2009].

Motivation to construct the expert systems, illustrated in this work, is to solve the problem of the analysis of market basket through implementation and verification of the method of market basket in the form of expert system, which enables:

1. Detecting and calculating the strength of dependence among commodities in selling transactions.
2. Discovering the patterns of market basket.
3. Graphic visualization of discovered patterns of customers.

The knowledge mentioned above can be successfully used in:

– planning Cross Selling campaigns,
– implementation of Up Selling technique in ecommerce activities,

– optimization of packages of products and services,
– planning of loyalty programs.

The aim of this article is to present and describe the logical model of expert system used in analysis of the market basket accordingly with all components.

## 2. The definition of the expert system

In the work [Niederliński 2006] expert system was defined as "a program which solves problems deputing to experts", other source [Giarratano, Riley 1998] gives definition: "An expert system is a computer program that draws on the knowledge of human experts captured in a knowledge base to solve problems that normally require human expertise", and „An expert system is a knowledge-based system emulates expert thought to solve significant problems in a particular domain of expertise". Edward Feigenbaum from Stanford University, the early pioneer of expert system technology, defined expert system as [Giarratano, Riley 1998] "an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solutions".

Based on quoted definitions, the system presented in this paper is dedicated to analyse similarities between variables in large databases. It should be noticed that no expert system can replace a human-expert and has some disadvantages and advantages in comparison. Boundaries concerns inference engine which is not able to group all association rules independently to user profiles. On the other hand, implemented algorithm is fast and effective enough [Mikulski, Weichbroth 2009], processing a great amount of data per second, which is unreachable for a human. Additionally, the amount of stored information is only limited by hardware resources. It is reasonable to think that these are objective reasons to design and develop an expert system, not only because of the aforementioned reasons.

## 3. Market basket analysis

This paragraph gives a brief description of frequent pattern mining for the discovery of interesting associations and correlations between itemsets in transactional databases.

Frequent patterns are patterns (such as itemsets, subsequences, or substructures) which appear in a data set frequently. For example, a set of items, such as eggs and milk, which appear frequently together in a transaction data set is a frequent itemset. A subsequence, such as buying first a PC, then a removable disk and then a printer, if it occurs frequently in a shopping history database, is a (frequent) sequential pattern [Han, Kamber 2006]. Frequent pattern mining searches for recurring relationships in a given data set.

Association rule is simply, probabilistic statement about co-occurrence of variables (items) in large databases. If we assume that all variables are binary, association rule will have the following form:

if ($A$ = 1 and $B$ = 1) then $C$ = 1 with probability $p$,

where $A$, $B$, $C$ are binary variables and $p = p(C = 1| A = 1, B = 1)$ is conditional probability which states for $A = 1$ and $B$ =1 then $C = 1$ [Hand et al. 2001]. Conditional probability $p$ is defined as accuracy or confidence of the rule. In turn, $p(A = 1, B = 1, C = 1)$ is defined as support. Assumptions of the thesis above, are comparatively easy to understand and to interpret which effects in clear understanding the whole approach.

Commonly, the goal of the market basket analysis is to find all the association rules, which satisfy the constraint in figure of minimal values $p_s$ (support) and $p_c$ (confidence), often defined as cut-off [Weichbroth 2009] (for example, cut-off equals $p_s \geq 0.05$ and $p_c \geq 0.7$). Such sets are called frequent item sets. On the other hand, the definition "frequent" is based on given frequency from user, called support [Wick, Wagner 2006].

The presented approach to finding association rules originate in applications for market basket analysis. Data which can be given to analyse, often come from glossary market, telecom company or nowadays more often from multi-branch Internet market. This kind of data may be considered in a matrix category with $n$ rows (corresponding to baskets) and $p$ columns (corresponding to items). Matrix like this may be very large, counts rows in millions and columns in thousands, highly sparse, especially when we take into consideration the fact that common basket contains only a few items. Application of association rules enables to find simple patterns in such a type of data in relatively efficient computational manner [Hand et al. 2001].

The classic example of market basket analysis is retail suggestive sell in e-commerce solutions, like this using by online shopping stores like amazon.com (USA) or merlin.pl (Poland). In this case, the server analyses a very large transactions database in order to find sets of items, which were bought frequent together.

## 4. The model of the expert system

Expert system for market basket analysis (shortly ESMBA) has got all attributes of a typical expert system. The system's core is inference engine, possessing two inference modes: forward chaining and backward chaining. Visualization of the discovered knowledge, depicted in the form of association rules, is processed by visualization engine. Knowledge base is responsible for gathering and sharing knowledge to domain expert. Users' queries and answers are stored in knowledge repository.

The idea of the expert system is to discover knowledge from market basket database and inference from this knowledge. The action's sequence, necessary to inference process, is dynamically synthesized for every knowledge base by inference engine. This means it is not overtly programming during creation of the knowledge base. Additionally, in declared system's functionality, independent component of market basket transactions database, was added (Figure 1).

Our expert system is a set of programs to solve problems deputing to experts. It has functional structure which consists of the following components: inference engine, visualization engine, knowledge repository, knowledge base, knowledge base editor, user and expert interface and market basket database (Figure 1).

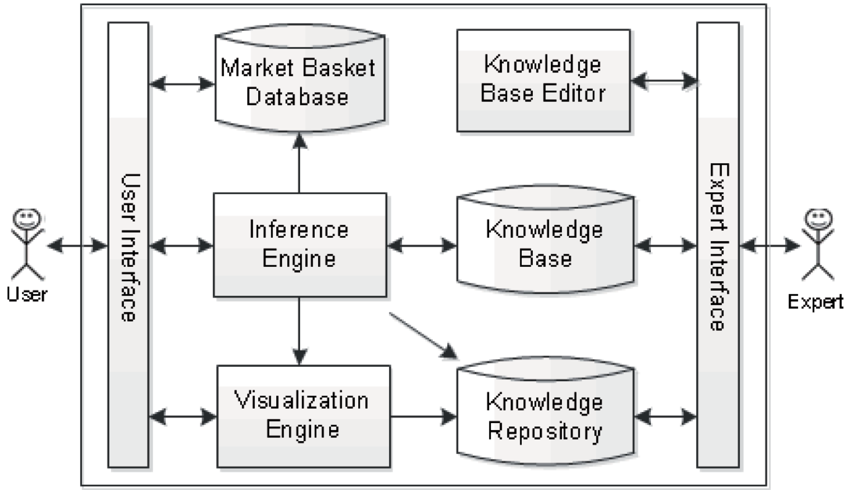


**Figure 1.** Functional architecture of a system expert

Source: based on [Abraham 2005].

As can be seen on Figure 1, in order to edit knowledge base, only expert can do this using knowledge base editor. Inference engine can only have access to the knowledge repository, saving queries declared by users and answers to them, received from inference engine.

Another component of the presented system is sales transactions database. It will be used to find all frequent itemsets and based on them to generate association rules. Discovered knowledge will be saved in knowledge base. Performing operations on it is only permitted by using knowledge base editor. Thus, user interface enables communication with inference engine since only expert interface with knowledge base.

### 4.1. Inference engine

The main task of inference engine is to process analysis of sales transactions stored in market basket database. For this purpose, modified Apriori [Agrawal, Srikant 1994] algorithm was implemented, detailed described in [Weichbroth 2009; Mikulski, Weichbroth 2009] and empirically verified on www server logs [Mikulski, Weichbroth 2009]. The output data is a set of association rules with support and confidence for each rule. The last task is to save them to knowledge base, with indicating the date, hour and number of processed records.

Additionally, the inference engine is able to objectively group some rules to a pattern. It is possible to use two criteria: antecedent or consequent. For the demonstration of such described functionality, we use a small set of artificial association rules, presented on Figure 2.

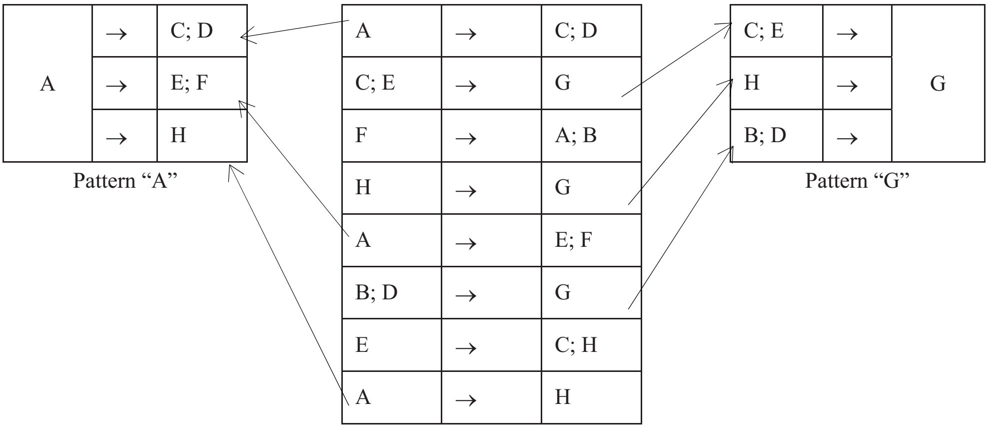


**Figure 2.** The process of grouping association rules to market basket patterns

Patterns "A" and "G" have one common product "G". The process of joining concerns finding common antecedent or consequent between them. As a result of joining patterns, the structuralised client model is created (Figure 3).

<table>
<tr><td rowspan="3">A</td><td>→</td><td>C; D</td><td>C; E</td><td>→</td><td rowspan="3">G</td></tr>
<tr><td>→</td><td colspan="2">H</td><td>→</td></tr>
<tr><td>→</td><td>E; F</td><td>B; D</td><td>→</td></tr>
</table>

**Figure 3.** The effect of joining patterns to form of a client model (profile)

On the ground of joined patterns (Figure 3), we can observe that promotion should touch the H product. This Product joins the pattern "A" and pattern "G". Taking into consideration mentioned above, when the customer with profile "A" feels a need of purchasing product G, also the need to purchase product G can appear and vice versa. In other words, promotion campaign should involve the products from “the middle” with the greatest amount of “incoming” and “out-coming” links. It is that because moving up and down on the pattern tree increase the chance to increase selling transactions.

### 4.2. Visualization engine

The concept of functionality of the components responsible for visualization of association rule was presented in the work [Weichbroth 2010]. The component GraphMiner was implemented in the ActionScript3 language and can be handled on each web browser (thin client), supporting Macromedia Flash technology. It contains three scripts which import data from: knowledgebase; second script generates data to XML files and the last one generates data charts (nodes and links). Input data for the importing script are components of association rules (antecedent and consequent) and their measures: support and confidence. Below we present the diagrams plotted to visualize discovered dependences among data (Figures 4, 5).

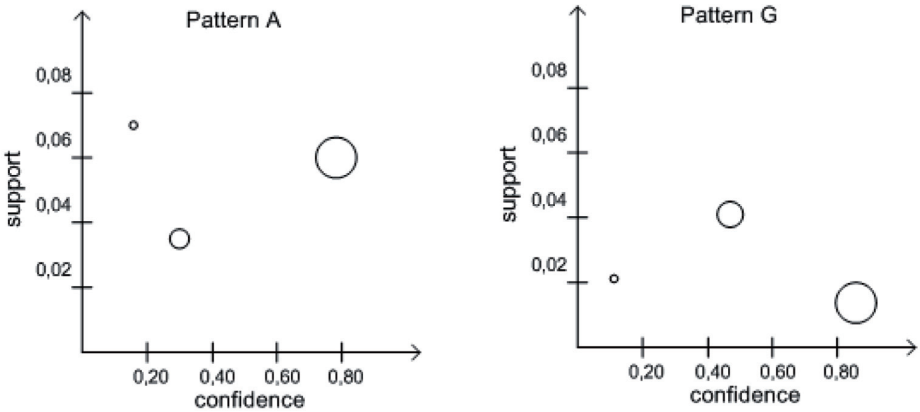


**Figure 4.** Visualisation of artificial patterns “A” and “G”

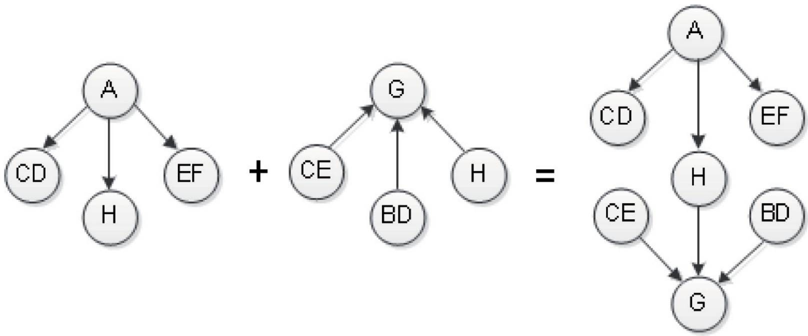


**Figure 5.** The process of joining patterns to one client model (profile)

Visualization of the knowledge written in form of association rules can be presented in other, more sophisticated form. If the user defines his minimal support or (and) confidence or (and) depicts precise products, the selection of rules can be conducted. Basing on those rules we may (Figure 6) plot the graph constructed with nodes and links. Strength of correlation among them (in the form of an arrow) is represented by its thickness – the higher the support of the association rule, the thicker the line. Moreover, arrows are coloured on the base of confidence and defined in advance by users and their partitions. Node dimension is dependable on centrality measure which contains first-order centrality measure (amount of "outcoming" links) and second-order measure (amount of "incoming" links) [Łapczyński 2009]. A good illustration of this functionality is presented below, in Table 1 and Figure 6.

**Table 1.** The parameters of visualization link analysis

| Support | | Confidence | | Node dimension | |
|---|---|---|---|---|---|
| Partition | Line thickness | Partition | Colour | Connections | Radius |
| ⟨0.01; 0.02) | ¼ pt | ⟨0; 0.1) | yellow | 1 | 0.75 cm |
| ⟨0.02; 0.04) | ¾ pt | ⟨0.1; 0.2) | orange | 3-5 | 1.25 cm |
| ⟨0.04; 0.06) | 1½ pt | ⟨0.2; 0.3) | red | 6-8 | 1.75 cm |
| ⟨0.06; 0.08) | 1¾ pt | ⟨0.3; 0.4) | dark red | 9-11 | 2.5 cm |
| ⟨0.08;1⟩ | 2½ pt | ⟨0.4; 0.5) | light green | 12 and more | 3 cm |
| – | – | ⟨0.5; 0.6) | green | – | – |
| – | – | ⟨0.6; 0.7) | light blue | – | – |
| – | – | ⟨0.7; 0.8) | dark blue | – | – |
| – | – | ⟨0.8; 0.9) | purple | – | – |
| – | – | ⟨0.9; 1⟩ | black | – | – |

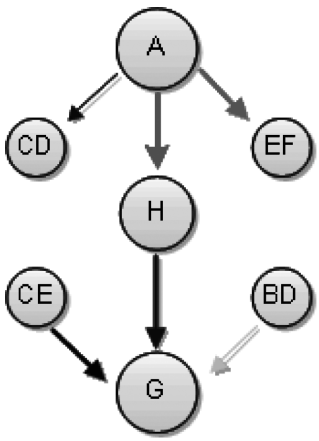


**Figure 6.** Rich link analysis of association rules

During designing visualization engine assumed that application should have following properties [Korczak, Dudycz 2009]:
– efficiency (on user demand, knowledge visualization is processing in real-time),
– focalization (visualization engine uses focus and context techniques like dynamic zooming and distortion),
– interactivity (data manipulation techniques like filtering, saving, loading, exporting and importing).

### 4.3. Knowledge repository

The term "knowledge repository" appears commonly in the literature of knowledge management, especially in association with commercially available knowledge management products. It refers to a system or system architecture that houses and manages a collection of corporate intellectual assets [Lin, Qin 2002]. A knowledge repository is a computerized system that systematically captures, organizes and categorizes the knowledge. The repository can be searched and data can be quickly retrieved. Knowledge repository provides high quality technical documentation relating to the expert system. This information is public, read-only and is intended for use in various aspects of education, training, design and operation.

Knowledge repository stores answers which were given to the user and corresponding results. It also keeps log files of accessing the knowledge base with information "4W1H" about "What, Who/Whom, When, Where and How". Moreover, it is an image library for particular knowledge base and specified constraints.

### 4.4. Knowledge base

Expert systems have got different knowledge bases, designed and developed in a domain-driven way [De Hoog 1997]. The concept is to use a separate knowledge base that could be edited or redefined for new problems while retaining all the same code for interpreting and using that knowledge [Lindsay et al. 1993]. Presented framework assumed that knowledge base will contain:
– logical knowledge: frequent itemsets, knowledge base, patterns, relations, models (profiles),
– procedural knowledge: methods,
– explanatory knowledge: knowledge origin.

Knowledge base resources come from two sources:
– subjective (arbitrary): human-expert who has expertise knowledge about domain, knowledge achieved during work or other resources,
– objective (heuristic): market basket analysis, results delivered from inference engine.

The role of the expert is to verify discovered patterns and models of clients. On the other hand, iterated market basket analysis in precisely defined cycle is the basis for determining which of the models are local and which are global. One of the possible criteria to use is geographical localization in case of international multi-branch stores. If we assume approximately almost the same assortment, empirical client model from Warsaw can be completely different from the one which comes from London. There are a lot of reasons for such diversity like culture, habits or nationality. Additionally, periodic comparison of patterns and models enables discovering seasoning products, whose sales fluctuations are dependent on days of the week, summer holidays or holidays in general.

Taking into account a well-known classification [Niederliński 2006] based on colligated criteria (applying rule's conclusions or relational model's conclusions as conditions for other rules) and types of logic, knowledge base described in such a way, can be classified as elementary (colligation is not accompanied with negation) and precise (inference process utilize from classic dual-based Aristotle's logic) [Niederliński 2006].

### 4.5. Knowledge base editor

The component available exclusively from the level of expert interface is called the editor of knowledge base. It is used to read, formulate and modify a knowledge base. It is a tool for graphical browsing, editing and formulating knowledge bases across the multiple tables in uniform manner. It offers users an intuitive interface, in which objects and items are represented as nodes in graph (Figure 6), with the relationships between them forming the edges. The Expert is able to create models in subjective manner, change them, delete and verify them. A model created in abstractive manner can be verified through analyze research, conducted on real (existing) or new data, storing in market basket database.

### 4.6. User and expert interface

Communication between expert and user was performed basing on mechanism of Web portal, using Apache server and a Web browser. The division of the interface into the user interface and the expert interface was conducted in order to isolate the components. User's authentication is performed by the Apache server. Correctly defined and located system's interfaces make it possible to exchange information easily among system components.

### 4.7. Market basket database

The aim of inference engine is to analyze the historic data, stored in a database. To make it appropriately, it is necessary to perform data preprocessing, to adapt the primary format of data to a form which is required by the expert system. The essential issue during designing and implementing the database is an issue of data granularity which means how the data should be detailed when we save it in the database. The process of analyzing market basket needs such attributes as: date of transaction, list of positions, price, and form of payment. The advanced model of analysis assumes that some additional data about customers should exist thanks to the use of loyalty cards or regular customer cards store cards.

## 5. Threats to validity

Rules comprise relatively weak forms of knowledge – they are just accurate relation of variables, rather than reliable thesis, which describe and characterize population. In fact, common meaning of the term "rule" means the common interpretation "from left to right side". What follows that rule, the term association rule can be a confusion, since the rules will be verified – which means they are no longer accidental.

The classical model of exploration the association rules avails the support mechanism which treats each transaction in the same manner. To the contrary, in real aggregations each transaction has different wage. The object of research in analyzing market basket is the collection of transactions, in which each one brings profit or loss. As yet, many works dedicated to extraction of association rules with preassigned weight have been published, for example [Cai et al. 1998; Ramkumar et al. 1998; Tao et al. 2003; Wang et al. 2000]. The importance of the number of items in transaction should be also noticed. Transactions with higher number of items should be treated as more important than those with lower number of positions. Implemented algorithm is not capable of estimating the weights of transactions such as profit and validity. In the work of Sakshi and Akkiraju [2010] a new measure of weighed support (w-support) for item sets with exclusively binary attributes was presented.

Higher complexity of applications generated the new needs. First of all, evolution of understanding how to deal with key problems like knowledge representation, inference and risk management. Secondly, development of the expert system shell and its programming languages. These two factors potentially minimize effort needed to design and deploy inference engine and knowledge base for dedicated domain.

As far, the inference and visualization engines were implemented and verified, which gives us hope for success of the whole project. Basing on software engineering approach, proposed in work [Ricordel, Demazeau 2000], which consists of four stages: analysis, design, and deployment, in the phase of design the diagrams of data flow and use cases in UML notation will be used [Wrycza et al. 2005].

## 6. Related work

Market basket analysis is still interesting subject of interest for many researchers, especially in the scope of association rules. This discipline of data mining has been successful in many commercial solutions like Statistica [Łapczyński 2009], SAS or Microsoft SQL Server. Brief description of this applications can be also found in [Weichbroth 2010].

A few articles presented how cross-category relationships can be merged with purchase models (see e.g. [Russell et al. 1997, 1999] and [Seetharaman et al. 2004]). Two main research approaches can be distinguished. The first can be described as data-driven using data mining. It is dominated by techniques like pair-wise association (e.g. [Hruschka et al. 1999]), association rules (e.g. [Agrawal, Srikant 1994]), vector quantisation, neural networks (e.g. [Decker, Monien 2003]) and collaborative filtering (e.g. [Mild, Reutterer 2001, 2003]).

The second approach is more like explanatory driven. It tries to identify and quantify cross-category choice effects of marketing-mix variables. In this case, two general methods can be identified. The multivariate probit approach (e.g. [Ainslie, Rossi 1998; Chib, Seetharaman 2002; Deepak et al. 2004; Manchanda et al. 1999]) is an extension of the standard probit approach (e.g. [Train 2003]) for one category. It is built on a disaggregate level and based on Random Utility Theory. The error distribution is assumed to be normal. Alternatively, the multivariate logit approach (e.g. [Hruschka et al. 1999; Russell, Petersen 2000; Singh et al. 2004]) can be used, which is an extension of the multinomial logit model (e.g. [Guadagni, Little 1983]) which is also based on Random Utility Theory. The error term of the multivariate logit approach is assumed to be Gumbel distributed.

Another interesting approach is presented in [Boztug, Hildebrandt 2005] where in the model, the assumption is stated that consumers make their category choices in some fixed order, which is not observed by the researcher. This lack of information causes that the choice in each category is modelled conditional upon known choices in all other categories.

## 7. Conclusions and future research

Above we briefly mentioned the problem of the market basket analysis and methods serving to solve it. In our framework, one of the earliest methods was applied, proposed in [Agrawal, Srikant 1994].

The model presented by Newell and Simon [Giarratano, Riley 1998] to solving human information problems in categories: long-term memory (rules), short-term memory (working memory) and cognitive processor (inference engine), is the fundament of expert systems nowadays. This paper introduces a framework which is built based on independent components (layers), but together they create a coherent system. In other words, every layer is an independent application but also a part of the entire system [Maciaszek 2008].

Presented expert system assumes functionality which enables to conduct market basket analysis and save its results to knowledge base. Discovered knowledge can be successfully used to achieve the targets, mentioned in the introduction. The first step to achieve this goal is to design the logical model of the entire system including most important relations between tables.

Still, there are a couple of things to be verified. Firstly, we need to connect inference engine with visualization engine and knowledge base if the system has to work in real-time mode. Secondly, knowlege and repository base need a schema of tables, necessary to store information from the inference engine and user respectively. The project will be closed after performing a few simulations in order to verify every component and data flow between them. Moreover the results gained from the inference engine (association rules, patterns and profiles) need to be verified in the scope of real-life common sense.

## References


Abraham A. (2005), Rule-based expert systems, [in:] *Handbook of Measuring System Design*, Eds. P.H. Sydenham, R. Thorn, John Wiley & Sons, New York.

Agrawal R., Srikant R. (1994), Fast algorithms for mining association rules, [in:] *Proceedings of the Twentieth International Conference on Very Large Data Bases*, Morgan Kaufmann, San Francisco.

Ainslie A., Rossi P.E. (1998), Similarities in choice behavior across product categories, *Marketing Science*, Vol. 17, No. 2, Hanover, pp. 91-106.

Boztug Y., Hildebrandt L. (2005), *A Market Basket Analysis Conducted with a Multivariate Logit Model*, SFB 649 Discussion Papers SFB649DP2005-028, Sonderforschungsbereich 649, Humboldt University, Berlin.

Cai C.H., Fu A.W.C., Cheng C.H., Kwong W.W. (1998), Mining Association Rules with Weighted Items, [in:] *Proceedings of the 1998 International Symposium on Database Engineering & Applications*, IEEE Computer Society, Washington, pp. 68-77.

Chib S., Seetharaman P.B. (2002), Analysis of multi-category purchase incidence decisions using IRI market basket data, [in:] *Econometric Models in Marketing*, Eds. P.H. Franses, A.L. Montgomery, Elsevier Science, New York, pp. 57–92.

Czarnecki A. (2008), Model zarządzania ontologiami w środowisku oceny technologii informatycznych, [in:] *Zarządzanie wiedzą i technologiami informatycznymi*, Eds. C. Orłowski, Z. Kowalczuk, E. Szczerbicki, Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk, pp. 413-422.

Czarnecki A. (2009), Wykorzystanie ontologii przy ocenie złożoności projektu informatycznego, [in:] *Zastosowanie technologii informatycznych w zarządzaniu wiedzą*, Eds. C. Orłowski, Z. Kowalczuk, E. Szczerbicki, Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk, pp. 179-188.

De Hoog R. (1997), Methodologies for building knowledge based systems: Achievements and prospects, [in:] *The Handbook of Applied Expert Systems*, Ed. J. Liebowitz, CRC Press, Boca Raton.

Decker R., Monien K. (2003), Market basket analysis with neural gas networks and self-organising maps, *Journal of Targeting, Measurement and Analysis for Marketing*, Vol. 11, No. 4, Houndmills, pp. 373-386.

Deepak S.D., Ansari A., Gupta S. (2004), *Investigating consumer price sensitivities across categories*, Working Paper, University of Iowa, Iowa City.

Giarratano J., Riley G. (1998), *Expert Systems Principles and Programming*, PWS Publishing Company, Boston.

Guadagni P.M., Little D.C. (1983), A Logit Model of Brand Choice Calibrated on Scanner Data, *Marketing Science*, Vol. 2, No. 3, Hanover, pp. 203-238.

Hajek P., Valdes J.J. (1994), An analysis of MYCIN-like expert systems, *Mathware and Soft Computing* (1), Universidad de Granada, Granada.

Han J., Kamber M. (2006), *Data Mining. Concepts and Techniques*, Elsevier, San Francisco.

Hand D.J., Mannila H., Smyth P. (2001), *Principles of Data Mining*, Prentice Hall, The Massachusetts Institute of Technology Press, Cambridge, MA.

Hruschka H., Lukanowicz M., Buchta C. (1999), Cross-category sales promotion effects, *Journal of Retailing and Consumer Services*, Vol. 6, Elsevier, New York, pp. 99-105.

Korczak J., Dudycz H. (2009), Approach to visualisation of financial information using topic maps, [in:] *Information Management*, Eds. B.F. Kubiak, A. Korowicki, Gdańsk University Press, Sopot, pp. 86-97.

Korczak J., Merchelski W., Oleszkiewicz B. (2008), A new technological approach to money laundering discovery using analytical SQL Server, [in:]: *Advanced Information Technologies for Management – AITM 2008*, Eds. J. Korczak, H. Dudycz, M. Dyczkowski, Research Papers No. 35, Wrocław University of Economics, Wrocław, pp. 80-104.

Korczak J., Oleszkiewicz B. (2009), Data warehouse structures for AML applications, [in:] *Proceedings of CEE Symposium on Business Informatics in Central and Eastern Europe*, Vienna.

Kowalczuk Z., Wszołek J. (2009), *Sieciowy monitor obiektu wspierający pracę eksperta*, [in:] *Inżynieria Wiedzy i Systemy Ekspertowe*, Eds. A. Grzech, K. Juszczyszyn, H. Kwaśnicka, N.T. Nguyen, Akademicka Oficyna Wydawnicza Exit, Warsaw, pp. 329-337.

Lin X., Qin J. (2002), Building a topic map repository, [in:] *Proceedings of Knowledge Technologies Conference*, Washington.

Lindsay R.K., Buchanan B.G., Feigenbaum E.A., Lederberg J. (1993), DENDRAL: A case study of the first expert system for scientific hypothesis formation, *Artificial Intelligence*, Vol. 61, No. 2, Elsevier Science Publishers, Essex.

Łapczyński M. (2009), *Analiza koszykowa i analiza sekwencji – wielki brat czuwa*. *Seminaria zastosowania statystyki i data mining*, Statsoft Polska, Warszawa.

Maciaszek L.A. (2008), *Modelowanie i rozwój adaptacyjnych złożonych systemów informatycznych*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.

Manchanda P., Ansari A., Gupta S. (1999), The "shopping basket": A model for multicategory purchase incidence decisions, *Marketing Science*, Vol. 18, No. 2, Hanover, pp. 95-114.

Mikulski Ł., Weichbroth P. (2009), Discovering patterns of visits on the Internet web sites in the perspective of associative models, *Polish Journal of Environmental Studies*, Vol. 18, No. 3B, Olsztyn.

Mild A., Reutterer T. (2001), Collaborative filtering methods for binary market basket analysis, [in:] *Active Media Technology*, Eds. J. Liu, P.C. Yuen, C.H. Li, J. Ng, T. Ishada, Springer-Verlag, Berlin, pp. 302-313.

Mild A., Reutterer T. (2003), An improved collaborative filtering approach for predicting cross-category purchases based on binary market data, *Journal of Retailing and Consumer Services*, Vol. 6, No. 4, New York, pp. 123-133.

Mirończuk M., Kreński K., *Koncepcja systemu ekspertowego do wspomagania decyzji w Państwowej Straży Pożarnej*, [in:] *Inżynieria wiedzy i systemy ekspertowe*, Eds. A. Grzech, K. Juszczyszyn, H. Kwaśnicka, N.T. Nguyen, Akademicka Oficyna Wydawnicza Exit, Warsaw 2009, pp. 699-712.

Newell A. (1972), *Human Problem Solving*, Prentice Hall, New York.

Niederliński A. (2000), *Regułowe systemy ekspertowe*. Wydawnictwo PKJS, Gliwice.

Niederliński A. (2006), *Regułowo-modelowe systemy ekspertowe RMSE*, Wydawnictwo Pracowni Komputerowej Jacka Skalmierskiego, Gliwice.

Orłowski C., Sitek T. (2007), Ocena technologii informatycznych – koncepcja wykorzystania systemów inteligentnych, [in:] *Komputerowo zintegrowane zarządzanie*, Ed. R. Knosala, vol. II, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole.

Orłowski C., Ziółkowski A. (2009), Wytwarzanie architektury korporacyjnej jako środowisko weryfikacji systemu agentowego do oceny technologii informatycznych, [in:] *Zastosowanie technologii informatycznych w zarządzaniu wiedzą*, Eds. C. Orłowski, Z. Kowalczuk, E. Szczerbicki, Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk, pp. 285-294.

Orłowski C., Ziółkowski A., Czarnecki A. (2010), Validation of an agent and ontology-based information technology assessment system, *Cybernetics and Systems*, Vol. 41, No. 1, Taylor & Francis, London.

Ramkumar G. D., Ranka S., Tsur S. (1998), Weighted association rules: Model and algorithm, [in:] *Proceedings of Fourth ACM International Conference of Knowledge Discovery and Data Mining*, ACM.

Ricordel P.M., Demazeau Y. (2000), From analysis to deployment: A multi-agent platform survey, [in:] *ESAW'00: Proceedings of the First International Workshop on Engineering Societies in the Agent World*, Springer-Verlag, London, pp. 93-105.

Russell G.J., Bell D., Bodapati A., Brown C.L., Chiang J., Gaeth G., Gupta S., Manchanda P. (1997), Perspectives on multiple category choice, *Marketing Letters*, Vol. 8, No. 3, Berlin, pp. 297-305.

Russell G.J., Petersen A. (2000), Analysis of cross category dependence in market basket selection, *Journal of Retailing*, Vol. 76, No. 3, New York, pp. 367-392.

Russell G.J., Ratneshwar S., Shocker A.D., Bell D., Bodapati A., Degeratu A., Hildebrandt L., Kim N., Ramaswami S., Shankar V.H. (1999), Multiple-category decision-making: Review and synthesis, *Marketing Letters*, Vol. 10, No. 3, Berlin, pp. 319-332.

Rutkowski L. (2006), *Metody i techniki sztucznej inteligencji*. Wydawnictwo Naukowe PWN, Warszawa.

Sakshi S., Akkiraju L.J. (2010), A new approach to association rule mining, *Lecturs Notes in Computer Science*, Springer, Berlin [forthcoming].

Seetharaman P.B., Chib S., Ainslie A., Boatwright P., Chan T., Gupta S., Mehta N., Rao V., Strijnev A. (2004), *Models of Multi-Category Choice Behavior*, Working Paper, Rice University, Houston.

Singh V.P., Hansen K., Gupta S. (2004), *Modeling Preferences for Common Attributes in Multi-category Brand Choice*, Working Paper, Carnegie Mellon University, Pittsburgh.

Tao F., Murtagh F., Farid M. (2003), Weighted association rule mining using weighted support and significance framework, [in:] *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, pp. 661-666.

Train K.E. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Wang W., Yang J., Yu P.S. (2000), Efficient mining of Weighted Association Rules (WAR), *Proceedings of ACM SIGKDD'00*, ACM, New York, pp. 270-274.

Weichbroth P. (2009), Odkrywanie reguł asocjacyjnych z transakcyjnych baz danych, [in:] *Rynek usług informatycznych*, Eds. A. Nowicki, I. Chomiak-Orsa, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 82, Informatyka Ekonomiczna 14, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.

Weichbroth P. (2010), The visualization of association rules in market basket analysis as a supporting method in customer relationship management systems, [in:] *Proceedings of Knowledge Acquisition and Management Conference*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław [forthcoming].

Wick M.R., Wagner P.J. (2006), Using market basket analysis to integrate and motivate topics in discrete structures, [in:] *Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education*, ACM, New York, pp. 323-327.

Wrycza S., Marcinkowski B., Wyrzykowski K. (2005), *Język UML 2.0 w modelowaniu systemów informatycznych*, Helion, Gliwice.

## ARCHITEKTURA REGUŁOWEGO SYSTEMU EKSPERTOWEGO DO ANALIZY KOSZYKA ZAKUPÓW

**Streszczenie:** Artykuł ten prezentuje nowatorskie podejście do odkrywania reguł asocjacyjnych w analizie koszyka zakupów. Do tego celu opracowany został system ekspertowy z maszyną wnioskującą i wizualizacyjną oraz bazą wiedzy. Po krótkim wprowadzeniu, w ogólnym zakresie opisano poszczególne komponenty systemu. W dalszej części scharakteryzowano funkcjonalność maszyny wnioskującej i wizualizacyjnej oraz bazy wiedzy. Pracę kończy podsumowanie oraz wskazanie dalszych kierunków badań.