An Evolutionary Computing Model to Implement Decision Support Systems for Buying Electronics

G. Santos, R. Vieira, and E. Costa

Abstract— **The representation of data through a search engine in e-commerce systems using decision support systems can be considered a complex process due to lots of information and may grow exponentially, according to research performed by the user for acquisition of equipment electronics. The import information collected by the system may influence the time to answer during the course of the search engine in accordance with the user's request. This article proposes a hybrid approach using database and genetic algorithm based on abstract data types to reduce the data of database dimensionality and helpful in the selection of products.**

Index Terms— **Artificial intelligence, Business, Genetic Algorithms, Knowledge Based Systems.**

I. Introduction

The process of decision support systems (DSS) can be considered as a major source related to the techniques that represent the knowledge based learning experience and can also be applied in almost all diverse areas of knowledge [1]. Recent works like [2], [3] have applied knowledge based learning experience in medicine. It consists of examining the basis of old cases of and therefore selecting them with the current ones according to similar characteristics aiming at reviewing the proposed solutions in which the most successful among them will be part of the basis of knowledge or cases. In agreement with Tsai [4], intelligent systems are used for inferring the main parameters of a productive process applied to the manufacture of printed-circuit board, involving lengthy production activities consisting of a number of variables that can influence the quality of product. However, other areas were notable when applying DSS in the internet environment such as addressing the construction of intelligent agents able to interact directly in products and services negotiations. Multi-agent systems are hereby used for the purpose of automating negotiation process constructed for autonomy and power to perform simple or complex tasks [5],[6], [7].

The background of artificial intelligence (AI) using multi-agent systems is directly related to the processes of automated trading, where the agent might be able to perform automated negotiations similar to those made in the negotiations of traditional commerce. In most cases the problems of automated trading can be complex and also involve sophisticated solution, providing the analysis of certain issues so that they are resolved in an integrated way [5]. Other applications addressing DSS have been employed in literature addressing a single area, normally based on auctions [8]. In agreement with Turban [9], the auctions may employ an environment of negotiations and discuss various models mainly those that are related to the Dutch auction, for example eBay, in which sellers determine the price for an initial item and the number of commodities for sale.

Manuscript submitted July 17, 2009. This work was funded by CAPES through the project PROCAD 1493/2007.

G. N. Santos is master's student in Computer Modeling of Knowledge, Computer Institute, Federal University of Alagoas - UFAL, Alagoas, Zip Code: 57072-970, BR (e-mail: genildo16@gmail.com).

R. V. Vieira is with the Department of Computer Institute, Federal University of Alagoas - UFAL, (e-mail: rv2l@hotmail.com).

E. B. Costa is with the Department of Computer Institute, Federal University of Alagoas - UFAL, (e-mail: ebcosta@gmail.com).

II. Methodology

The experiments used in this article are obtained from e-commerce environments, where the customer applies a search engine available on site and automatically the web information system shows a list of products surveyed. A major drawback of this environment is related to purchasing good quality products at low price level, traditional search engines may inhibit the customer to buy a good price quality product owing to the large amount of information generated during consultation. However, DSS uses this search to select cases of best interest (those of low price and good quality) by generating a knowledge database in which its exponential growth is able to influence the process of analysis. Some works were related and proposals approaches were also developed for reducing dimensionality in databases using traditional genetic algorithms (GA). In agreement with Qi [10], the evolutionary algorithms are applied to reduce an infinite population size of infinite size in a data space using crossing techniques. Rockett [11] and Raymer [12], shows the use of GA for complex problems involving data of high-dimensionality.

However, traditional genetic algorithms are seldom used in literature to reduce database knowledge using DSS. Furthermore, traditional GA may not support any kind of information derived from these database, and may influence the formation of chromosomes. In agreement with Vieira [13], new GA approaches can be seen in literature using genetic algorithm based on abstract data types (GAADT). Genetic algorithm includes a subset of the computational study of evolution and natural selection, Darwin-based techniques for optimization problems focusing on the application of selection, mutation, recombination, and for a population of solutions to concurrent problems [12]. The

GAADT is a computational genetic model, where individuals will be represented in according to the genetic material (chromosomes) which has its foundations in basic training units [13]. The experiments carried out in this article show that, in applying GAADT you can have a decrease in records in the database without losing the information essence from the basis of initial cases. This article proposes hybrid architecture for library cases of computation system of negotiation with virtual actors using genetic algorithm based on abstract types of data reducing the size of the database knowledge.

III. Implementation and Results

Table I is a sample within a field of shopping online that originally contained more than one thousand registered sales records for a base *B*, thus, a base *B* can be considered a set of all genetic sample units and can be used in shaping the genetic material of the chromosomes of a population.

We have made many tests with data extracted from a search site Buscapé (http://www.buscape.com.br, price comparison site) through the existing mechanism in the prototype software created for illustrating the operation of the models presented in this paper. Fig. 1 illustrates the cited mechanism. To illustrate the operation of the model presented in this paper, we'll use the example of the problem represented in Table I which involves 10 records, where each record is formed by n characteristics (period of consultation or purchase, description, price and the product model). Each character will have its identification, value, type and a code to distinguish from other characteristics. For example, if case 1 has a characteristic identified as *Period*, valued *07/24/2007*, with *Date* type (a value represents the number of days that have elapsed, for example, since 12/30/1899.) and the code *1.1*. Another feature, still case 1, is the characteristic *Price*, valued *199.*^{00}, with *Numeric* type and the code *1.3*. To indicate the characteristics as the bases should be grouped to form a given character and if a law is needed to provide training. The law training of characteristics will be represented by the number of axioms of gene formation (AGF) [14]. The properties of this definition can only be completed during the genetic algorithm instantiation to a particular problem [15].

TABLE I
Example of Cases Stored

Ref	Period	Description	Price	Model
1	07/24/2008	Pressure Device	199,^{00}	BP33AA1
2	05/08/2008	Pressure Device	99,^{00}	MS-918
3	05/08/2008	Pressure Device	162,^{00}	Microlife
4	05/08/2008	Pressure Device	139,^{00}	BP3AF13
5	05/08/2008	Oximetro	650,^{00}	With wire
6	09/20/2007	Otoscope	307,^{00}	Pocket Jr
7	05/09/2007	Otoscope	329,^{00}	Pocket Jr
8	10/10/2007	Oximetro	1.845,^{00}	Wireless
9	10/15/2007	Otoscope	259,^{00}	Control
10	10/20/2007	Oximetro	1.295,^{00}	Wireless

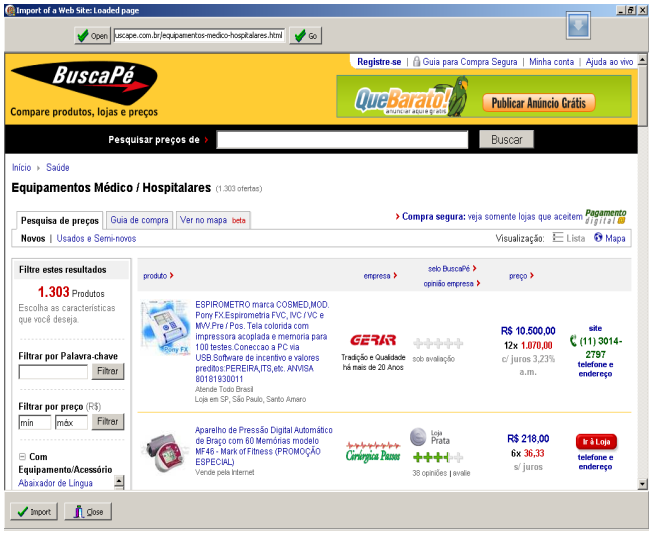

Fig. 1. Import screen

In agreement with [12] and [14], we can give the following definitions to the problem of support for buying electronics: *A* is known as the set of n characteristics, where such characteristics are formed by a code, an identifier (or description) a type and cost corresponding to the type. The code is what distinguishes a characteristic from the other and a possible set of codes will be within the natural numbers (i.e., code $\in \mathbb{N}$). The role of the identifier is associated with a label id, a feature and a possible set of identifiers will be within the alphabetical symbols (i.e., id $\neq \emptyset$). The type talks about the domain represented by the characteristic of the case where all possible sets will be *Numeric*, *Data* and *Alphabetical*. The possible sets of values according to the type chosen are defined below:

TABLE II
Model of Function $value(v)$

Title	Definition
Function $value(v)$	Set of possible values according to the type chosen.

$$value(v) = \begin{cases} v \in \mathbb{R}\ if\ type = Numerical \\ v = \ mmddaa\ if\ type = Date \\ v \in \mathbb{N}\ if\ type = Alphabetical \end{cases}$$

Where

$$dd, mm, aa \in \mathbb{N} | 1 \leq dd \leq 31, 1 \leq mm \leq 12$$

TABLE III
Model of Base *B* of the Environment

Title	Definition
Base *B* of the environment	Representing the basis *B* of the gene for the environment *A*.

$$B = B_{codes} \cup B_{identifiers} \cup B_{types} \cup B_{values}$$

Where

$$B_{codes} = \{ code \in N | code \neq \emptyset \}$$

$$B_{identifiers} = \{ i \in Alphabetical \}$$

$$B_{types} = \{Numerical, Date, Alphabetical\}$$

$$B_{values} = \{v \in \mathbb{R} \vee v = mmddaa | mm, dd, aa \in N | 1 \leq mm \leq 12, 1 \leq dd \leq 31 \vee v \neq \emptyset\}$$

TABLE IV
MODEL OF GENE *G* OF THE ENVIRONMENT

Title	Definition
Gene *G* of the environment	Representation of the set of genes belonging to environment *A*.
$G = \{g =< cod, id, t, v > \mid cod, id, t, v \in B\}$	

TABLE V
MODEL OF AGF THE ENVIRONMENT

Title	Definition
AGF the environment	Representation of the set of axioms gene training belonging to the *A* environment.
$AFG = \{g =< cod, id, t, v > \mid \forall(cod, id, t, v) \exists(cod \neq 0 \wedge id \neq \emptyset \wedge t \neq \emptyset \wedge v \neq \emptyset)\}$	

TABLE VI
MODEL OF CHROMOSOMES *C*

Title	Definition
Chromosomes *C* of the environment	Representation of all the chromosomes of the environment *A*.
$C = \{< g_1, g_2, \dots, g_n > \mid \forall x(1 \leq x \leq n) g_x \in G\}$	

In Table VI, we have *C* as the chromosome and g_n as the gene by the GAADT model in *A*, equivalent to a particular case. That is, *C* represents the case and the g_n characteristics of this case. The set of axioms formation of chromosome (AFC) from the *A* environment is defined in Table VII.

For the experience of Table I, as result, we would have the following set of genes for the chromosome of case 1 in Table VIII.

For each case contained in Table I a result is achieved in a corresponding chromosome implementation. To carry out the operations, definition of the determined degree function is necessary as follows in Table IX.

TABLE VII
MODEL OF AFC THE ENVIRONMENT *A*

Title	Definition
AFC the environment	Representation of the set of axioms formation of environmental chromosome *A*.
$AFC = \{C_i =< g_1, g_2, \dots, g_x >, C_j =< g_1, g_2, \dots, g_y > \mid \forall C_i, C_j \in C \wedge \forall m, n(1 \leq m \leq x, 1 \leq n \leq y) g_m, g_n \in G\ (cod_m \neq cod_n \wedge\ id_m = id_n)\}$	

TABLE VIII
CHROMOSOME IMPLEMENTATION FOR CASE 1

code	*id* (identifiers)	*t* (types)	*v* (value)
1.1	Period	Date	07/24/2007
1.2	Description	Alphabetical	Pressure Device
1.3	Price	Numerical	$199,^{00}$
1.4	Model	Alphabetical	BP33AA1

TABLE IX
MODEL OF FUNCTION $degree(g)$

Title	Definition
Function $degree(g)$	Representing the set of axioms training of environmental genes *A*.

$$degree(g) = \begin{cases} v \rightarrow \frac{1}{v}, if\ v \in\ Numerical \\ v \rightarrow \frac{1}{current\ date - v}, if\ v \in\ Date \\ v \rightarrow \ln r \mid \end{cases}$$

Where

$$r \equiv representation \in\ \mathbb{N}, if\ v \in Alphabetical$$

TABLE X
MODEL OF FUNCTION $domi(g_1, g_2)$

Title	Definition
Function $domi(g_1, g_2)$	Dominant functions between two genes.

$$domi(g_1, g_2) = \begin{cases} g_\lambda\ if\ tipo(g_1) \neq type(g_2) \\ g_1\ if\ tipo(g_1) = type(g_2) \wedge degree(g_1) > degree(g_2) \\ g_2\ if\ tipo(g_1) = type(g_2) \wedge degree(g_2) > degree(g_1) \end{cases}$$

Table X describes the dominant role between the two genes to the environment *A*, returning the dominant gene.

For the function type used in Table X, it can be expressed in the definition in Table XI.

Thus the modeling of creating the genetic algorithm capable of generating further buying experience with cardinality less than the current record and reduce the space of the database to increase the efficiency of recovery without loss of information within an e-commerce environment for equipment purchase pertaining to a hospital. The exposed description in the Table XII reinforces the model in study.

Assuming the records contained in Table I as population, in applying the algorithm in Table XII and in future the results could be viewed in Table XIII.

Table XIII is the assistance given by the system in selecting equipment electronics. For example, regarding the item described by Otoscope of the Table I, the purchase of the model Pocket Jr is recommended at a price of R$ *288,*[50].

TABLE XI
MODEL OF FUNCTION $type(g)$

Title	Definition
Function $type(g)$	Gene function type.

$$type(g) = \begin{cases} Numerical\ if\ (g \in B_{codes}) \vee (g \in \mathbb{R}) \vee (g \in \mathbb{N}) \\ Alphabetical\ if\ g \in B_{identifiers} \vee g \in B_{types} \\ \text{Date if } g = mmddaa \end{cases}$$

Where

$$mm, dd, aa \in \mathbb{N} \mid 1 \leq dd \leq 31, 1 \leq mm \leq 12$$

TABLE XII
MODEL OF FUNCTION $GA(P_n)$

Title	Definition
Function $GA(P_n)$	Genetic algorithm function of environment A.

$$GA(P_n) = \begin{cases} P_n = P_{otm} & if\ adap_m(P_n) > t \\ P_{n+1} & if\ n+2 = k \\ GA(P_{n+1}) & \end{cases}$$

Where

- P_n is the number of records within the moment n. P_o is the number of initial records coming from the database;
- P_{otm} is the number of records optimized or ideal;
- $adapt_m(P_n) > t$ is a condition for stopping to ensure that a certain number of records have a greater adjustment than a tax value t;
- $n + 2 = k$ is a condition for stopping to ensure that even if the algorithm did not converge for a C_{otm} it will stop in a number of k iterations;

Fig. 2 shows the software screen, created to illustrate the operations of the models presented in this paper, referring to compression of existing records on the database with registered sales of *10,000*, approximately.

IV. CONCLUSION

The manipulation of data using decision support systems process based past experiences can be considered an intuitively attractive strategy because it is similar to the behavior of the solution to human problems. People use past experience to solve new problems, and this approach becomes appropriate, and effective and often, softens the load analysis of a specific field. This leads to the advantage that intelligent systems can be based on superficial knowledge and does not require a significant effort in terms of knowledge created when compared to other approaches based on rules. We describe a genetic algorithm able to optimize a database in DSS environment which it was tested in an online shopping sample area that originally contained more than 10 thousand registered sales in which the model presented converged to 850 records, approximately. With the aid of the examples in the previous section, it was possible to illustrate the reduction of the database without information loss and describe a relational model that meets most of the representations and mechanisms for evolutionary systems.

TABLE XIII
NEW CASE BASE

Ref.	Period	Description	Price	Model
11	05/08/2008	Pressure Device	119,^{00}	BP3AF13
12	10/15/2007	Otoscope	288,^{50}	Pocket Jr
13	05/08/2008	Oximetro	1.110,^{50}	Wireless

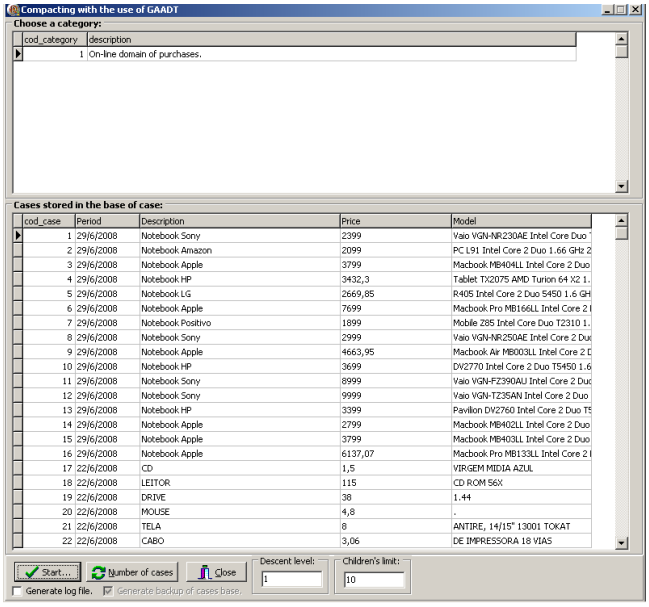

Fig. 2. Compression screen of prototype software

From this study an algorithm capable of generating new records of lesser cardinality was developed to reduce the database area, and increase the records recovery efficiency without losing information within a shopping online area.

For future work, we plan to create a multi-agent evolutionary system to contemplate the main theories involving the line of knowledge discovery research to automate negotiation processes. Parallel to this task, we will create a paper illustrating the software operations and behavior with a large amount of information in the knowledge database.

REFERENCES

[1] S. Pal, and S. Shiu, *Foundations of Soft Case-Based Reasoning*. New Jersey: Wiley, 2004.

[2] I. Bichindaritz, Stefania Montani, and Luigi Portinale, "Special issue on case-based reasoning in the health sciences," vol. pg. 207- 209, 2007.

[3] S. Montani, "Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support," vol. 275-285, 2006.

[4] C. Y. Tsai, and C. C. Chiu, "A case-based reasoning system for PCB principal process parameter identification," *Expert Systems with Applications*, vol. pg 1183–1193, 2007.

[5] S. Kraus, *Strategic negotiations in multiagent enviroments*. Massachustts: Bradford, 2001.

[6] J. A. R. P. Sardinha, MAS-School e ASYNC: Um Método e um Framework para Construção de Agentes Inteligentes, 2005.

[7] B. Costin, M. Ganzha, and P. Marcin, *Developing a Model Agent-based E-Commerce System*. Berlin: Springer, 2007, vol. 37.

[8] A. P. Silva, and R. Oliveira, "Um ambiente de suporte a negociações eletrônicas automatizadas utilizando ontologias e regras de produção," *Simpósio Brasileiro de Engenharia de Software*, pp. 49-59, 2006.

[9] E. Turban, and D. King, *Comércio Eletrônico - Estrátégia e Gestão*. São Paulo: Pearson Education do Brasil, 2004.

[10] X. Qi, and F. Palmieri, "Theoretical Analysis of Evolutionary Algorithms With an Infinite Population Size in Continuous Space Part 11: Analysis of the Diversification Role of Crossover," TRANSACTIONS ON NEURAL NETWORKS, vol. Vol 5, pg 120-128, no. nº 1, 1994.

[11] R. Kumar, and P. Rockett, "Multiobjective Genetic Algorithm Partitioning for Hierarchical Learning of High-Dimensional Pattern Spaces: A Learning-Follows-Decomposition Strategy," TRANSACTIONS ON NEURAL NETWORKS, vol. Vo. 9, no. ed. 5, 1998.

[12] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Kuhn, and A. K. Jain, "Dimensionality Reduction Using Genetic Algorithms," IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, vol. 4, 2000.

[13] R. V. Vieira, and M. A. Lopes, "Uma Abordagem Teórica Inicial para os Algortimos Genético através de tipos Abstratos de Dados," Universidade Federal de Pernambuco, Recife, Relatório Técnico RT-D1/UFPE, N° 002/99, 1999.

[14] R. V. Vieira, Um Algoritmo Genético Baseado em Tipos Abstratos de Dados e sua Especificação em Z, 2003.

[15] H. Ahn, K. Kim, and I. Han, "Hybrid genetic algorithms and case-based reasoning systems for customer classification," vol. 23, no. 3, 2006.