Neuro-symbolic System for Business Internal Control

Juan M. Corchado[1], M. Lourdes Borrajo[2], María A. Pellicer[1], and J. Carlos Yáñez[3]

[1] Deparatamento de Informática y Automática,
University of Salamanca, Plaza de la Merced s/n, 37008 Salamanca, Spain
corchado@usal.es
[2] Department of Computer Science,
University of Vigo, Campus As Lagoas, s/n, 32004 Ourense, Spain
[3] Department of Financial Accounting,
University of Vigo, Campus As Lagoas, s/n, 32004 Ourense, Spain
{lborrajo, jcyanez}@uvigo.es

Abstract. The complexity of current organization systems, and the increase in importance of the realization of internal controls in firms, make it necessary to construct models that automate and facilitate the work of auditors. An intelligent system has been developed to automate the internal control process. This system is composed of two case-based reasoning systems. The objective of the system is to facilitate the process of internal auditing in small and medium firms from the textile sector. The system, analyses the data that characterises each one of the activities carried out by the firm, then determines the state of each activity, calculates the associated risk, detects the erroneous processes, and generates recommendations to improve these processes. As such, the system is a useful tool for the internal auditor in order to make decisions based on the risk generated. Each one of the case-based reasoning systems that integrates the system uses a different problem solving method in each of the steps of the reasoning cycle: fuzzy clustering during the retrieval phase, a radial basis function network and a multi-criterion discreet method during the reuse phase and a rule based system for recommendation generation. The system has been proven successfully in several small and medium companies in the textile sector, located in the northwest of Spain. The accuracy of the technologies employed in the system has been demonstrated by the results obtained over the last two years.

1 Introduction

Nowadays, organization systems employed in enterprises are increasing in complexity. Moreover, in recent years, the number of regulatory norms has increased considerably. As a consequence of this, the need has arisen for periodic internal audits. But the evaluation and the prediction of the evolution of these types of systems, characterized by their great dynamism, are, in general, complicated. It is necessary to construct models that facilitate analysis work carried out in changing environments, such as finance.

P. Perner (Ed.): ICDM 2004, LNAI 3275, pp. 1–10, 2004.

The processes carried out inside a firm can be included in functional areas [19]. Each one of these areas is denominated a "Function". A Function is a group of coordinated and related activities, which are necessary to reach the objectives of the firm and are carried out in a systematic and reiterated way [11]. Functions are divided into activities, which are associated to well defined objectives. The functions that are usually carried out within a firm are: Purchases, Cash Management, Sales, Information Technology, Fixed Assets Management, Compliance to Legal Norms and Human Resources.

In turn, each one of these functions is broken down into a series of activities. For example, the function Information Technology is divided in the following areas: Computer Plan Development, Study of Systems, Installation of Systems, Treatment of Information Flows and Security Management.

Each activity is composed of a number of tasks, for example: register, authorise, approve, harmonise, separate obligations, operate, etc. Control procedures are established in the tasks to assure that the established objectives are achieved. Rule-based systems (RBS) have traditionally been used with the purpose of delimiting the audit decision-making tasks [6]. However, Messier and Hansen [13] found many situations in which auditors resolved problems by referring to previous situations. This contrasts with the very nature of RBS systems, since they have very little capacity for extracting information from past experience and present problems in order to adapt to changes in the environment.

In contrast, case based reasoning systems (CBR) are able to relate past experiences or cases to current observations, solving new problems through the memorization and adaptation of previously tested solutions. This is an effective way of learning, similar to the general structure of human thought. CBR systems are especially suitable when the rules that define a knowledge system are difficult to obtain, or the number and complexity of the rules is too large to create an expert system. Moreover, CBR systems have the capacity to update their memory dynamically, based on new information (new cases), as well as, improving the resolution of problems [14].

However, in problems like those presented in this study, standard techniques of monitoring and prediction cannot be applied due to the complexity of the problem, the existence of certain preliminary knowledge, the great dynamism of the system, etc. In these types of systems it is necessary to use models that combine the advantages of several mechanisms of problem-solving capable to of resolving specific parts of the general problem and attending other parts.

In this sense, an adaptive system has been developed. The system possesses the flexibility to behave in different ways and to evolve, depending on the environment in which it operates. The developed system is composed of two fundamental subsystems:

- Subsystem IEA (Identification of the State of the Activity) whose objectives are:
  1. to identify the state or situation of each one of activities of the company.
  2. to calculate the risk associated with this state.
- Subsystem GR (Generation of Recommendations), whose goal is:
  1. generation of recommendations from the detection of inconsistent processes. These recommendations will allow the positive evolution of the internal processes of the company.

Both subsystems are implemented with the use of two CBR systems (one for each subsystem). Each one of the CBR systems is used as a basis for the integration of symbolic and connectionist models, used in different steps of the reasoning cycle.

The rest of this article is structured as follows: firstly, to explain the concept of internal control (IC) and describe its importance within the modern company; secondly, the basic concepts that characterize case based reasoning are presented, an explanation given of how the system has been constructed; finally, the initial results will be presented.

2 Internal Control

Small to medium enterprises require an internal control mechanism in order to monitor their modus operandi and to analyse whether they are achieving their goals. Such mechanisms are constructed around a series of organizational policies and specific procedures dedicated to giving reasonable guarantees to their executive bodies. This group of policies and procedures are named "controls", and they all conform to the structure of internal control of the company. The establishment of objectives (which is not a function of the Internal Control) is a previous condition for control risk evaluation, which is the main goal of the Internal Control.

In a wide sense, the administration of a firm has three large categories of objectives when designing a structure for internal control [2]:

1. Reliability of financial information.
2. Efficiency and effectiveness of operations.
3. Fulfillment of the applicable rules and regulations.

The internal auditor must monitor the internal controls directly and recommend improvements on them. Therefore, all the activities carried out inside the organization can be included, potentially, within the internal auditors' remit. Essentially, the activities of the auditor related to IC can be summarised as follows:

- To be familiar with and possess the appropriate documentation related to the different components of the system that could affect financial aspects.
- To assess the quality of internal controls in order to facilitate the planning of the audit process with the aim of obtaining necessary indicators.
- To assess internal controls in order to estimate the level of error and reach a decision on the final opinion to be issued in the memorandum on the system under consideration.

As a consequence of the great changes in firms brought about by current technological advances, considerable modifications have taken place in the area of auditing, basically characterized by the following features [16]:

- Progressive increase in the number and level of complexity of audit rules and procedures.
- Changes in the norms of professional ethics, which demand greater control and quality in auditing.

- Greater competitiveness between auditing firms, consequently resulting in lower fees; the offer of new services to clients (e.g. financial or computing assessment...).
- Development of new types of auditing (e.g. operative management auditing, computer auditing, environmental auditing...).

Together, these circumstances have made the audit profession increasingly competitive. Consequently, the need has arisen for new techniques and tools, which can be provided by information technology and artificial intelligence. The aim is to achieve more relevant, more suitable information, in order to help auditors make decisions faster and thereby increase the efficiency and quality of auditing.

3 Case Based Reasoning Systems: An Overview

A case based reasoning system (CBR) solves a given problem by means of the adaptation of previous solutions to similar problems [1]. The CBR memory stores a certain number of cases. A case includes a problem and the solution to this problem. The solution of a new problem is obtained recovering similar cases stored in the CBR memory.

A CBR is a dynamic system in which new problems are added continuously to its memory, the similar problems are eliminated and gradually new ones are created by combination of other several existent ones. This methodology is based on the fact that human use the knowledge learned in previous experiences to solve present problems.

CBR systems record past problem solving experiences and, by means of indexing algorithms, retrieve previously stored problems with their solutions (cases), and match and adapt them to a given situation. This means that the set of cases stored in the memory of CBR systems represents the knowledge concerning the domain of the CBR. As discussed below, this knowledge is updated constantly.

A typical CBR system is composed of four sequential steps which are recalled every time a problem needs to be solved [9, 1, 17]:

1. *Retrieve* the most relevant case(s).
2. *Reuse* the case(s) in order to solve the problem.
3. *Revise* the proposed solution if necessary.
4. *Retain* the new solution as a part of a new case.

Like other mechanisms of problem solving, the objective of a CBR is to find the solution for a certain problem. A CBR is a system of incremental learning, because each time a problem is solved, a new experience is retained, thereby making it available for future reuse.

CBR systems have proven to be an effective method for problem solving in multiple domains, for example, prediction, diagnosis, control and planning [10]. This technology has been successfully used in several disciplines: law, medicine, diagnosis systems etc. [17].

The case based reasoning can be used by itself, or as part of another conventional or intelligent system [12]. Although there are many successful applications based on CBR methods alone, CBR systems can be improved by combining them with other technologies [8]. Their suitability to integration with other technologies, creating a global hybrid reasoning system, stems from the fact that CBR systems are very flexible algorithms, capable of absorbing the beneficial properties of other technologies.

4 Hybrid Neuro-symbolic System for Internal Control

This section describes the intelligent system in detail. The objective of the system is to facilitate the internal control process in small to medium sized enterprises. After analyzing the data relative to each activity that is developed within the firm, this system determines the state of each activity and calculates the associated risk. It also detects any erroneous processes and generates recommendations for improving these processes.

In this way, the system helps the internal auditor make decisions based on the risk associated to the current state of each one of the activities in the firm.

The cycle of operations of the developed case based reasoning system is based on the classic life cycle of a CBR system [1, 18]. This cycle is executed twice, since the system bases its operation on two CBR subsystems (subsystem IEA-Identification of the State of the Activity and subsystem GR-Generation of Recommendations). Both subsystems share the same case base (Table 1 shows the attributes of a case) and a case represents the “shape” of a given activity developed in the company.

Table 1. Case structure

	IDENTIFICATION	DESCRIPCTION
PROBLEM	*Case number*	Unique identification: positive integer number.
	Input vector	Information about the tasks (n sub-vectors) that compose an industrial activity: $(GI_1,V_1),(GI_2,V_2),...,(GI_n,V_n)$) for n tasks. Each task sub-vector has the following structure (GI_i,V_i): • GI_i: importance rate for this task inside the activity. Can only take one of the following values: ▪ VHI (Very high importance). ▪ HI (High Importance). ▪ AI (Average Importance) ▪ LI (Low Importance) ▪ VLI (Very low importance) • V_i: Value of the realization state of a given task. Positive integer number (between 1 and 10).
	Function number	Unique identification number for each function
	Activity number	Unique identification number for each activity
	Reliability	Percentage of probability of success. It represents the obtained percentage of success using the case as a reference to generate recommendations.
	Degree of membership	$((n_1 ,\mu_1), (n_2, \mu_2), \ldots, (n_k, \mu_k))$ • n_i represent the ith cluster • μ_i represent the membership value of the case to the cluster n_i
SOLUTION	*Activity State*	Degree of perfection of the development of the activity, expressed by percentage.

In the subsystem IEA the state or situation of each company activity is predicted, in real time, and the associated risk to this state is calculated, as can be seen in Table 2 and Figure 1. First, the more similar cases of the case base are grouped, using fuzzy clustering [3, 5]. When a new problem case is presented, the cluster to which it belongs is identified, than all *k* cases from this cluster which present a high degree of ownership to the case problem are retrieved.

The retrieved cases are used in the reuse phase to train an RBF network (Radial Basis Function) [7, 4]. The goal of the network is to build a generic solution from the *k* retrieved cases. The network determines the company state.

If the internal auditor, in the analyzed company, thinks the initial solution is coherent, the problem, together with his solution (the state of the activity identified by the system) is stored in the case base, as it is a new case, a new piece of knowledge. The addition of a new case to the case base causes the redistribution of the clusters. Also, from this solution, the level of risk inherent to the state of the activity is calculated. The subsystem GR generates the recommendations for improving the state of the analyzed activity, as represented in Figure 1 and Table 3. In order to recommend changes in the execution of the processes in the firm, the subsystem compares the obtained state of the activity to the cases belonging to the cluster used in the initial phase of subsystem IEA, which reflect a better situation for this activity in the firm. Therefore, in the retrieval phase, a selection is made of those cases whose solution or state of activity is higher (in an interval between 15% and 20%) than the

Table 2. Summary of technologies employed in the IEA subsystem

CBR's Phases	Technology	Input	Output
Retrieve	Fuzzy Clustering	Problem case	K similar cases
Reuse	RBF network	Problem case K similar cases	Initial solution: state of the activity
Revise	RBS	Problem case Initial solution: state of the activity	Revised solution: state of the activity level of risk inherent to the state of the activity
Retain	RBF network	Problem case Final solution	New Case

Table 3. Summary of technologies employed in the GR subsystem

CBR's Phases	Technology	Input	Output
Retrieve	Fuzzy Clustering	Problem case	K similar cases
Reuse	method ELECTRE	K similar cases	Most favourable case(s)
Retain	Manual	Most favourable case(s) (used to generate recommendations)	Modified case(s)

solution generated as output in the subsystem IEA. In the reuse phase, the multicriteria decision-making method Electre [15] is used to obtain the most favorable case of all, depending on the degree of relevance of the tasks.

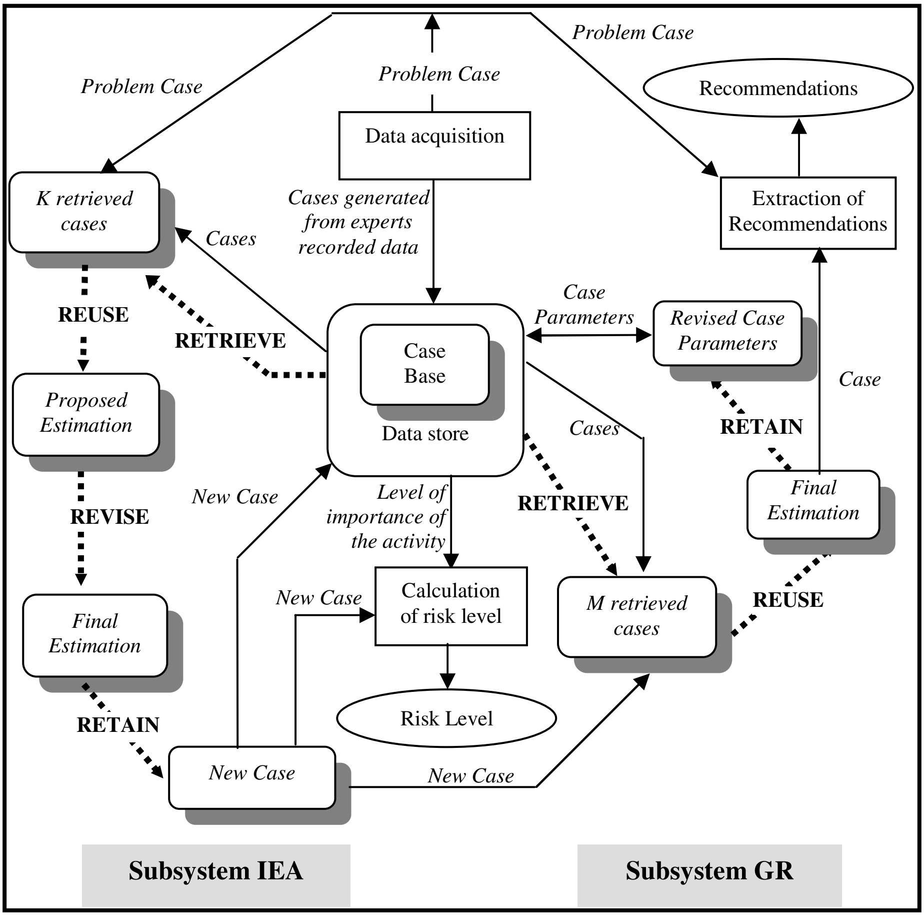

Fig. 1. System reasoning process

The selected case is compared with the initial case problem, using a simple rule based system to determine the order in which the recommendation should be taken to minimize or eliminate the erroneous processes.

This is a decision support system that facilitates the auditing process for internal auditors. After the time necessary for correcting the errors detected, the firm is evaluated again. Auditing experts consider that three months are enough to evolve the company towards a more favourable state. If it is verified that the erroneous processes and the level of risk have diminished, the retention phase is carried out, modifying the case used to generate the recommendations. The reliability (percentage of successful identifications obtained with this case) of this case is thereby increased. In contrast, when the firm happens not to have evolved to a better state, the reliability of the case is decreased.

5 Results

The developed hybrid system has been tested in several small to medium companies in the textile sector, located in the northwest of Spain. Previously, surveys were carried out by auditors and experts in different functional areas of the firms within the sector. These surveys have provided the necessary prototype cases in order to construct the case bases of the system.

Results obtained demonstrate that the application of the recommendations generated by the system causes a positive evolution in firms. This evolution is reflected in the reduction of erroneous processes. The system has been tested in 22 companies (12 medium-sized and 10 small). The best results occurred in the companies of a smaller size. Figure 2 presents a graphical representation of the companies evolution. This is due to the fact that these firms have a greater facility to adapt and adopt the changes suggested by the system's recommendations., because of their smallest size. 15 of them evolved successfully, 5 of them did not improve their results and 2 of them reduced their business. These results demonstrate the suitability of the techniques used for their integration in the developed intelligent control system.

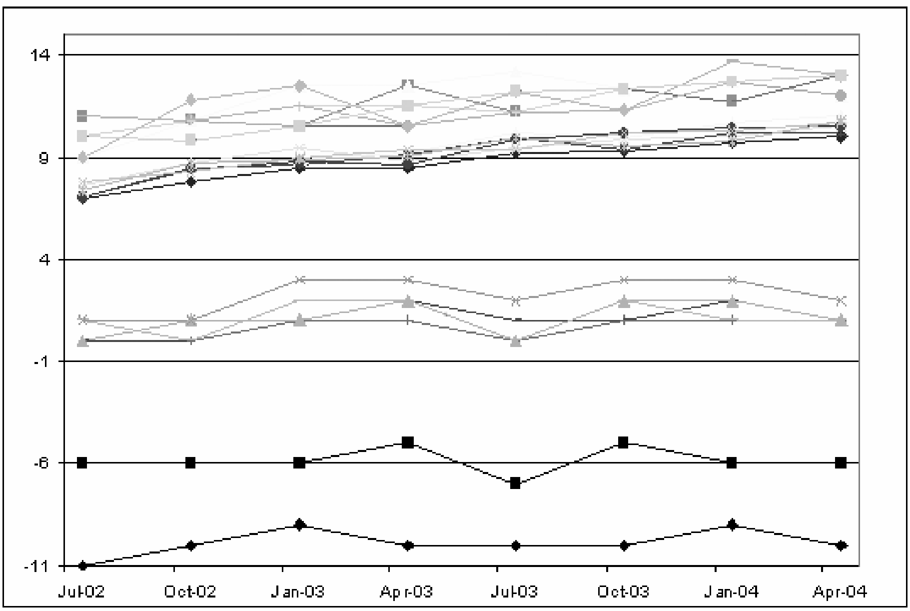

Fig. 2. Firms' evolution

6 Conclusions

This article presents a neuro-symbolic system that uses two CBR systems employed as a basis for hybridization of a multicriteria decision-making method, a fuzzy clustering method, and an RBF net. Therefore, the developed model combines the complementary properties of the connectionist methods with the symbolic methods of Artificial Intelligence.

The used reasoning model can be applied in situations that satisfy the following conditions:

1. Each problem can be represented in the form of a vector of quantified values.
2. The case base should be representative of the totality of the spectrum of the problem.
3. Cases must be updated periodically.
4. Enough cases should exist to train the net.

The system is able to estimate or identify the state of the activities of the firm and their associated risk. Furthermore the system generates recommendations that will guide the internal auditor in the elaboration of action plans to improve the processes in the firm.

The estimation in the environment of firms is difficult due to the complexity of the environment from where the prediction should be obtained and the great dynamism of this environment. However, the developed model is able to produce a prediction with enough precision, and within the limitations of time, imposed by the nature of the problem.

Nevertheless, the system will produce better results with data from firms belonging to the same sector. This is due to the dependence that exists between the processes in the firms and the sector where the company is located.

Although we haven't had the opportunity to test these techniques in big firms, we think that they would be satisfactorily applicable, although changes would take place more slowly than in small and medium firms.

References

1. Aamodt A. and Plaza E. (1994). Case-Based Reasoning: foundational Issues, Methodological Variations, and System Approaches. AICOM. Vol. 7. Nº 1, Marzo 1994.
2. American Institute of Certified Public Accountants (AICPA), Statements on Auditing Standards No. 78 (SAS No. 78) (1996). Consideraciones de la Estructura del Control Interno en una Auditoría de Estados Financieros. (Amendment to SAS núm. 55), The Auditing Standards Executive Committee, New York.
3. Bezdek J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.
4. Corchado, J.M., Díaz, F., Borrajo, L. and Fdez-Riverola F. (2000). Redes Neuronales Artificiales: Un enfoque práctico. Departamento de publicaciones de la Universidad de Vigo.
5. Dave, R.N. (1992). Generalized fuzzy C-shells clustering and detection of circular and elliptic boundaries. Pattern Recogn. 25, 713-722.
6. Denna, E.L., Hansen, J.V. and Meservy, R. (1991). Development and application of expert systems in audit services. Transactions on Knowledge and Data Engineering.

7. Fritzke, B. (1994). Fast Learning with Incremental RBF Networks. Neural Processing Letters. Vol. 1. No. 1. pp. 2-5.
8. Hunt, J. and Miles, R. (1994). Hybrid case-based reasoning. The Knowledge Engineering Review. Vol. 9:4. pp. 383-397.
9. Kolodner J. (1993). Case-Based Reasoning. San Mateo. CA, Morgan Kaufmann. 1993.
10. Lenz M., Bartsch-Spörl B., Burkhard D. and Wees S. (eds.) 1998. Case-based Reasoning Technology: From Fundations to Applications, Springer Verlag, LNAI 1400.
11. Mas, J. and Ramió, C. (1997). La Auditoría Operativa en la Práctica. Ed. Marcombo, Barcelona.
12. Medsker L. R. (1995). Hybrid Intelligent Systems. Kluwer Academic Publishers.
13. Messier, W.F. and Hansen, J.V. (1988). Inducing rules for expert systems development: an example using default and bankruptcy data. Management Science, 34, No. 12, December, 1403-15.
14. Riesbeck, C.K. and Schank, R.C. (1989). Inside case-based reasoning. Lawrence Erlbaum Associates. Hillsdale, NJ.
15. Romero, C. (1993) Teoría de la decisión multicriterio: Conceptos, técnicas y aplicaciones. Alianza Editorial. ISBN: 84-206-8144-X
16. Sánchez, A (1995): "Los Sistemas Expertos en la Contabilidad", Biblioteca Electrónica de Contabilidad, Vol. 1, N. 2.
17. Watson I. (1997). Applying Case-Based Reasoning: Techniques for Enterprise Systems. Morgan Kaufmann.
18. Watson, I. and Marir, F. (1994). Case-Based Reasoning: A Review. The Knowledge Engineering Review. Vol. 9. No. 4. pp. 355-381.
19. Yáñez, J.C., Borrajo, L. and Corchado, J.M. (2001). A Case-based Reasoning System for Business Internal Control. Fourth International ICSC Symposium. Soft Computing and Intelligent Systems For Industry. Paisley, Scotland, United Kingdom, June 26-29, 2001.