## An analysis of the possible applications of Artificial Intelligence Techniques to a Clinical Laboratory Information Management System

Diana Calva1<sup> $\dagger$ </sup> and Mario Lehman2<sup> $\dagger\dagger$ </sup>,

CEMINT Sofilab, Lisboa 14-A Col. Juárez, Mex. D.F.

#### Summary

Laboratory Information Management Systems (LIMS) have been in use since the mid seventies. Today the laboratory needs, have more to do with decision support and decision making. Depending on the problem needed to be solved, a different and unique method of artificial intelligence would have to be implemented that could easily interact with the database search engines each LIMS uses. Here we identify some of these new problems to be solved by a LIMS, the possible alternatives that could be implemented in order to solve them and include an example based on fuzzy logic.

### Key words:

Laboratory Information Management System, Artificial Intelligence, Relational Database, fSQL.

### **1. Introduction**

The amount of information obtained from the daily tasks performed in the clinical laboratory need to be supported with information systems. Also, most of the data collected maintain relations which need to be recorded and tracked in order to provide information useful for the laboratory operation. During the seventies and eighties, what was expected from a clinical laboratory information system, had to do only with storing and retrieving test results, booking patients, knowing the amount of tests performed over a period of time, billing patients, generating and printing test results reports, and performing simple statistics (internal quality control) among others [1]. Today most LIMS use relational databases and structure query language (SQL) for performing most of these tasks which are based on data searching, updating and retrieval. As the amount of data grows and the clinical laboratory finds their new needs could, in a way, be solved by adding new functions to the LIMS, queries become more complex and the structure query language becomes limited. Most of these needs have to do with planning, logistics, aspects related with total quality requirements, obtaining information having to do with the behavior of certain diseases (dengue hemorrhagic fever), the prevalence of different clinical parameters (like high glucose or cholesterol) in a target population, management of

epidemic data or data that could lead to the identification of nosocomial infection (grouped by age, sex, geographical location, etc.) [2], among others. Other needs have to do with including new technologies such as voice controlled tasks; image processing for identification of sample tags or character identification in medical orders; making suggestions based on the history of a particular patient, etc. [3] Also, most of the time, the laboratory stores precise data and needs to retrieve information based on imprecise criteria where the existence of a fuzzy query language or a way of translating the stored data into imprecise information in an effective way might be the best solution for achieving this possibility. All these new needs can be solved by using one or more artificial intelligence techniques. Today, no commercial available database engine provides tools based on artificial intelligence in order to perform these complex queries. In this work we identify and briefly explain some of these needs and propose a solution based on one or more artificial intelligence technique. Also we will illustrate an example of a query based on imprecise criteria on precise data

# 2. Data Retrieval on a Traditional LIMS based on a Relational Database

In relational databases, data retrieval is achieved through the performance and execution of queries, stored procedures, views or functions. Most of the time, one of these methods are enough in order to obtain results and extract useful information from the database. However these is limited for retrieving information that already exists in the database when using as search criteria any information already existing in the database as can be seen in Figure 1, where a store procedure for obtaining the set of patients booked between certain dates is shown. The procedure retrieves the list of the patients including their First Name, Last Name and Patient ID. These data are retrieved from two different tables.

Manuscript received December 5, 2008

Manuscript revised December 20, 2008



Fig. 1 Stored Procedure for retrieving a list of patients booked between a specified range of dates.

A similar query could be performed for retrieving information on a set of patients classified between a range of dates of birth, and this way the set of pediatric patients could be obtained. An example of such query can be seen in Figure 2.



Fig. 2 Query for obtaining pediatric patients based on the assumption that these patients are the ones having 12 years old or less.

However this kind of query will retrieve patients having 12 years old or less and will leave out of the result those that might still be considered as pediatric patients but whose age is slightly greater than 12 years, since the criteria "pediatric" is imprecise. The same will happen when trying to obtain the set of adult patients, since a range containing the ages where a patient is considered adult would have to be established, leaving out of the resulting set of patients those that might be considered as adults but do not belong to the selected range

### 3. Using Fuzzy Logic for achieving imprecise queries

The description above is just a simple example of the limitations of the current solutions a LIMS based on a relational database can present. Imprecise data can more easily be obtained by using fuzzy logic. The best way to achieve fuzzy queries is by implementing fuzzy classification, this way, there is no need of having a database able to support fuzzy logic operations in the structured query language [4]. This process might be achieved by converting all the crisp data into fuzzy data by using additional tables and catalogs and then applying a trapezoidal function that will allow associating a membership degree [5]. An example of this can be seen in Fig. 3, where a table containing the membership degrees for "age", are contained in table Mem\_Deg\_Age, this is obtained after applying f(x,a,b,c,d) to particular set of crisp data.

NAME	AGE
Ulises	24
Ana	10
Carlos	55

(a)

TAG	а	b	с	d
Adult	28	36	45	53
Young	10	20	25	35
Child	0	0	8	14
Old	45	55	70	70
1 \				

(b)



NAME	AGE	MEMBER DEGREE
Ulises	24	0.8
Ana	10	0.4
Carlos	55	0.3

**Fig. 3.** Fuzzy Classification over a conventional database. a) Table containing crisp data, b) Table containing the different limits

for f(x, a, b, c, d), c) Function applied for obtaining the property degree for each data d) Table containing Membership Degrees after applying f(x, a, b, c, d)

This way, several tables containing Membership Degrees for other important classification criteria may be added as tables in a conventional database entity relation model, and fuzzy searches could then be performed even by combining the results for other different fuzzy classifications resulting from applying the corresponding function and fuzzy set operations to these new tables as can be seen in Figure 4. Having a fuzzy logic as a tool for retrieving imprecise data from a LIMS adds functionality and facilitates the implementation of epidemiologic applications, follow up and evaluation of tests results for a particular treatment being applied to an individual or group of individuals



**Fig. 4.** (A) SQL Query analyzer performs a query over crisp data, (B) Results from this query contain the needed information for fuzzyfication, (C, D) A particular function is applied for obtaining membership degrees, (E) Membership degrees are stored in a table, (F) After performing store procedures, triggers or views on the query analyzer fuzzy information is obtained, (G, H)Fuzzy set operations can be performed with the results obtained from the second query and an imprecise result can be obtained

# 4 Complex Querying needed for solving specific clinical laboratory problems

The above is just an example of what might be achieved by including complex queries in a LIMS. Here, we identify some others and discuss other artificial intelligence techniques we believe might result useful if combined with SQL queries.

One of these techniques is Bayesian Networks which have proved to be useful in many applications dealing with assisted clinical diagnostics [6]. In the clinical laboratory decisions have to be made when a result or set of results do not fall between the reference values of a particular test. When this happens, the laboratory personnel must decide, depending on the available information of the patient, if the obtained result is expected or not, when it is not evident that the result was expected to be out of range, the laboratory technician has to run the test a second time under the same conditions and if the result is still out of range, then it must take decisions on whether to run the test again under a different technique, review if the patient has past results on that particular test, or contact the doctor responsible for ordering the tests in order to gather more information. A Bayesian Network combined with the tuples obtained form SQL queries, could assist this decision, by providing the laboratory technician a suggestion based on knowledge that can be obtained from the same database.

Another important task in the clinical laboratory has to do with internal quality control; this task is based on a statistical follow up of the results of known control tests and multirules using a combination of decision criteria, or control rules, to decide whether an analytical run is incontrol or out-of-control. The well-known Westgard multirule QC procedure uses 5 different control rules to judge the acceptability of an analytical run. By comparison, a single-rule QC procedure uses a single criterion or single set of control limits, such as a Levey-Jennings chart with control limits set as either the mean plus or minus 2 standard deviations (2s) or the mean plus or minus 3s. "Westgard rules" are generally used with 2 or 4 control measurements per run, which means they are appropriate when two different control materials are measured 1 or 2 times per material, which is the case in many chemistry applications. Some alternative control rules are more suitable when three control materials are analyzed, which is common for applications in hematology, coagulation, and immunoassays [7].

 Table 1. Other problems that could be solved in a LIMS by using an Artificial Intelligence Technique.

Application	Artificial Intelligence
11	Technique
Login using biometry (voice or	Neural Networks
fingerprint recognition)	
Test ordering with voice	Neural Networks
processing	
Type mistmatches, autocorrection	Bayesian Networks
Reference Values identification	Fuzzy logic, Deductive
	database
Identification of commonly	Bayesian Networks
ordered set of tests (depending on a	-
specific specialist or doctor)	

(d)

The decision on which set of rules to use, as well as the results interpretation and the decisions the laboratory personnel has to do based on such results, could very well be aided by obtaining the data on the obtained results with simple SQL queries and then by using a neural network for decision support, since neural networks are well suited for decision support of well known patterns as is the case of the Westgard multirules. Table 1, shows other probable applications which could be solved by using other artificial intelligence techniques.

### **5** Discussion

Information Technologies applied to the clinical laboratory have evolved since the first Laboratory Information Systems (LIS) of the eighties, to the current LIMS. Many of the laboratory needs are covered with the functions current LIMS, based on the storage and retrieval of data in and from relational databases offer. On the other hand, artificial intelligence has played an important role in applications dealing with supporting clinical diagnostics but not many things have been done regarding applications involving clinical laboratory. Many authors have dealt with the problem relational databases and more specifically SQL present when trying to implement artificial intelligence techniques and have proposed several ways of approaching and solving these problems [4, 8, 9]. In this work we present an example of an application dealing with imprecise search criteria and how it was solved by using fuzzy logic and information retrieved from a relational database. We also mention several other problems that might be solved with the data stored in a relational database in the clinical laboratory, but which need the implementation of an artificial intelligence technique in combination with the use of SQL. The main issue here is, not only identifying the problems that might be solved with artificial intelligence, but also to decide which is the most appropriate technique to use in order not to compromise performance and time response when performing complex queries supported on any artificial intelligence technique.

### 6 Conclusion

The application of artificial intelligence techniques is proposed for performing several functions that, today result difficult to develop with simple SQL queries. We show an example of how imprecise queries that are useful for the clinical laboratory can be achieved. Future work involves evaluating the different artificial intelligence techniques, their implementation with simple SQL queries and which are best suited for adding functionality to current LIMS which do not count with these kinds of applications and tools.

#### Acknowledgments

This work was supported by Software Integral para Laboratorio S. A. de C. V. (México DF, México) through the Research Project NeuroSofilab (Ref. SOF-971021-I75/2001-1) from Consejo Nacional de Ciencia y Tecnología (CONACyT) and Secretaría de Hacienda y Crédito Público (México)..

### References

- [1] Christopher MM, Hotz CS. Medical informatics in the veterinary clinical laboratory: an integrated subspecialty in diagnostics, resident training and clinical research. Revue Méd Vét. 2000;151(7):569-76.
- [2] Lamma E, Mello P, Nanetti A, Riguzzi F, Storari S, Valastro G. Artificial Intelligence Techniques for Monitoring Dangerous Infections. Information Technology in Biomedicine, IEEE Transactions on. 2006;10(1):143-55.
- [3] Winkel P. The application of expert systems in the clinical laboratory, Clinical Chemistry. 1989;35(8):1595-600.
- [4] Werro N. Concept and Implementation of a Fuzzy Classification Query Language. Department of Informatics, University of Fribourg; 2005.
- [5] Medina JM, Vila MA, Cubero JC, Pons O. Towards the implementation of a generalized fuzzy relational database model. Fuzzy Sets and Systems. 1995;75(3):273-89.
- [6] Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence In Medicine. 2001;23(1):89-109.
- [7] Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry.. Clinical Chemistry. 1981;27(3):493-501.
- [8] Chang JY, Tao SQ, Deng C. Design and implementation of a deductive database system. Beijing Gongye Daxue Xuebao(J Beijing Polytech Univ). 2004;30(4):505-9.
- [9] Taskar B, Segal E, Koller D. Probabilistic classification and clustering in relational data. Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence. 2001:870–6.



**Diana Calva** received the B.S. degree in Biomedical Engineering from Universidad Iberoamericana (México) in 1992. During 1994-1995 she taught the course Clinical Measurements and Laboratory at the Universidad Iberoamericana, and since 1995 she has worked in different Institutions in the area of Biomedical Engineering. Actually, she is a partner founder of the company Sofilab SACV and works as

Manager of the Engineering and Technology division. She is currently taking the PhD in Industrial Engineering (with a degree in Information Technologies) at Universidad Anáhuac (México). Her tesis is related with intelligent systems and the information processing in biomedicine. She is interested in areas such as: intelligent systems, data minning and information security, and their applications in biomedicine, image processing and computational optics.



Mario Marcelo Lehman received the B.S. degree from UNICEN (Argentina) and the PhD in Physics from Universidad Nacional de La Plata (Argentina) in 1999. He was visiting scientist at the Complex Media Laboratory, at Penn University (USA), Centro de Investigación Científica y Estudios Superiores de Ensenada (CICESE, México) and Instituto

Nacional de Astrofísica, Optica y Electrónica (INAOE, México). From 2000 he is working as Director of Research and Development in the company Sofilab SACV (México), with projects approved by the Consejo Nacional de Ciencia y Tecnología (CONACYT). Her areas of interest are: networks and optical systems, computacional optics and metrology, image processing and statistical physics