

This volume contains 28 papers to be presented in the 8th International Conference on Intelligent and Information Technologies (CITII 2012), which was held on October 22-26, 2012 at the Instituto Tecnológico de Apizaco, México. The conference received 68 submissions by 176 authors from 5 countries for evaluations, which were carefully reviewed and selected by an international programme committee. The papers in this volume are structure into eleven thematic fields representative of the main current areas of interest within the intelligent and information technologies, such as:

- Agents and Multi-Agent Systems
- Human-Computer Interaction
- Computational Intelligence
- Robotics and Intelligent Control
- Artificial Vision
- Databases
- Software Engineering
- Computer Networks and Communications
- Web Based Systems
- Systems Distributed and Parallel
- Machine Learning and Data Mining

The volume will be useful for researches and students working in the respective areas of the intelligent and information technologies, as well as, for all readers interested in intelligent and information applications.

ISSN: 18770-4069
www.cic.ipn.mx/rsc
www.ipn.mx



Special issue: Advances in Intelligent and Information Technologies

J. C. Hernández Hernández
J. F. Ramírez Cruz
A. Cortés Fernández
J. H. Sossa Azuela

Vol. 60



RESEARCH IN COMPUTING SCIENCE

ISSN: 18770-4069

Special issue: Advances in Intelligent and Information Technologies

J. C. Hernández Hernández
J. F. Ramírez Cruz
A. Cortés Fernández
J. H. Sossa Azuela

Vol. 60



Table of Contents

Índice

Page/Pág.

Computational Intelligence

Single-step-ahead and Multi-step-ahead Prediction with Evolutionary Artificial Neural Networks	3
--	---

V. Landassuri-Moreno, C. Bustillo-Hernández and J.J. Carbajal-Hernández

Integración de métodos para la clasificación-reconocimiento de patrones metalográficos a base de inteligencia artificial y procesamiento digital de imágenes 13	
---	--

E. A. Ruelas., J. A. Vázquez-López, J. Yañez-Mendiola, I. López-Juarez, C. F. Bravo-Barrera

DL Prolog: Another Unifying Programming Language.....	23
---	----

José Oscar Olmedo-Aguirre

Clasificador Asociativo de Patrones Iterativo(CAPI)	35
---	----

Sergio Valdez, Raúl Santiago y Humberto Sossa

Análisis de preguntas de cuestionarios con Redes Neuronales	45
---	----

José de Jesús Zamarripa-Topete, Luz Evelia López-Chico y Victor Rafael Velázquez -Mejía

Heurísticas para resolver el problema de la Supersecuencia Común más Corta	55
--	----

Marcela Rivera-Martínez, Luis-René Marcial-Castillo, Lourdes Sandoval-Solís

Identificación de Locutor usando Vectores Acústicos basados en Cuantiles	65
--	----

José-Martin Olguín-Ezpinoza, Pedro Mayorga-Ortiz, Luis Vizcarra

Identificación de personas por medio de la voz utilizando LPC's y reconocimiento por KNN(K-Nearest Neighbors)	77
---	----

Jesús Gustavo Medrano-Romero, Yesenia Nohemí González-Meneses, José Federico Ramírez-Cruz y Blanca Estela Pedroza-Méndez

Redes bayesianas aplicadas a un modelo de Dinámica de Fluidos Computacional para las condiciones climáticas al interior de un invernadero.	87
---	----

Guillermo Alfonso De la Torre-Gea y Enrique Rico-García

Robotics and Intelligent Control

Localización y Reconocimiento de Obstáculos de un Robot Móvil Mediante SLAM-R	99
---	----

Sergio Díaz-Zagal, Citlalih Gutiérrez-Estrada, José Rafael Cruz-Reyes, Rafael Lemus-González y Juan José Rivera-Alemán

Desacoplamiento dinámico y control de robot RRR de arquitectura abierta..... 111

Daniel De la Cruz-Muciño, Miguel Gabriel Villareal-Cervantes, Jesús Saíd Pantoja-García, Nancy Donaji Hidalgo-Vázquez

Aplicación de Técnicas Inteligentes de Control en un Motor de Corriente Directa
Incluyendo Modulación por Ancho de Pulsos..... 121

R. Morales-Caporal, J. I. Paredes-Tabales, O. Sandre-Hernández, E. Bonilla-Huerta y J. F. Ramírez-Cruz

Artificial Vision

Detección y cuantificación de Hemorragias Intracerebrales en Tomografía Axial
Computarizada 133

Enrique Irwin Illán García, María Guadalupe Medina Barrera, Federico Ramírez Cruz y José Juan Hernández Mora

Identificación de displasias en análisis de muestras de Papanicolaou usando ACO 143

Maricela Gress-Roldán, Edmundo Bonilla-Huerta, Roberto Morales-Caporal y José Crispín Hernández-Hernández

Segmentación en imágenes microscópicas..... 153

José A. Márquez, Beatriz A. Sabino, Jesús M. Campos y Manuel I. Martín

Software Engineering

Contratos traslúcidos: Implementación de Diseño por Contrato en el Desarrollo
Orientado a Aspectos 165

Guadalupe Isaura Trujillo-Tzanahua y Ulises Juárez-Martínez

A Generic Model of Deliverable Artifacts through the Development of Graphical
User Interfaces 175

Laura C. Rodríguez-Martínez, Ricardo Mendoza-González, Hector Durán and Mario A. Rodríguez-Díaz

Desarrollo de aplicaciones en las nuevas plataformas móviles para la gestión de la
seguridad del control de accesos 189

Arturo Estudillo-Quilantán, Alberto-Israel Castellanos-Reyes, Erick Cacho-Rojas y Ted Echeverría-Dionisio

Desarrollo y prueba de un Software de Apoyo para la Mejora de la Pronunciación en
Alumnos con Déficit del Habla por Dislalia 201

José Manuel Soto-Corzo, José Juan Hernández-Mora, José Federico Ramírez-Cruz, Blanca Estela Pedroza-Méndez

Computer Networks and Communications

BlueNews: modelo para la propagación de información por medio de la tecnología Bluetooth..... 215

Oscar D. Ozaine, Alberto Urbina, Arnoldo Díaz-Ramírez, Heber S. Henández y Juan F. Ibáñez

Web Based Systems

Uso de Ontologías y otras Fuentes de Datos para la Generación de Currículos Vitae ..227

Kenenias Bethuel Pérez-Bentanzos, Beatriz Alejandra Olivares-Zepahua, Giner Alor-Hernandez, Ana María Chávez-Trejo

Redes Sociales para el Seguimiento de Egresados 239

Briceida González-Jiménez, Gerardo Contreras-Vega, Maria-Karen Cotés-Verdín

Sistema Web de Escenarios de Aprendizaje Virtuales para las Matemáticas de Primer Grado de Nivel Secundaria 251

Gaspar García-González, Carmen Cerón-Garmica y Etefvina Archundia-Sierra

Tecnología móvil en la educación: Educ@-móvil 261

Karina Verastegui Barranco y Paola Nayeli Cortez Herrera

Descubrimiento y Visualización de Conocimiento en la Herramienta HACOPDS usando Tecnología de Web Semántica 271

Luis Alberto Nava-Ordoñez, César Enrique Rose-Gómez y Ana María Chávez-Trejo

Machine Learning and Data Mining

Una aproximación basada en redes neuronales para el problema de implicación textual translingüe 285

Darmes Vilariño, David Pinto, Mireya Tovar y Beatriz Beltrán

Agents and Multi-Agent Systems

Multiagent based decision making in strategy games 297

Diana Fernández-Tirado, Adolfo Pacheco-Nieto y Gilberto G. Flores-Vidaña

Author Index 309

Índice de autores

Track Chairs	311
Coordinadores de Área	
Editorial Board of the Volume	313
Comité Editorial del Volumen	
Additional Reviewers	315
Árbitros Adicionales	

Single-step-ahead and Multi-step-ahead Prediction with Evolutionary Artificial Neural Networks

V. Landassuri-Moreno¹, C. Bustillo-Hernández², J.J. Carbajal-Hernández²

¹Centro Universitario UAEM Valle de México
Universidad Autónoma del Estado de México, C.P. 54500, México.

²Centro de Investigación en Computación
Instituto Politécnico Nacional, C.P 07738, México.
vmlandassurim@uaemex.mx, chbustillo004@cic.ipn.mx, juancarvajal@sagitario.cic.ipn.mx
Paper received on 25/07/12, Accepted on 11/09/12.

Abstract. In recent years the Evolutionary Algorithms have been remarkably useful to improve the robustness of Artificial Neural Networks. This study will introduce an experimental analysis of the FS-EPNet algorithm to understand how neural networks are evolved with a steady-state algorithm and compare the Single-step-ahead (SSP) and Multiple-step-ahead (MSP) methods for prediction tasks over two test sets. It was decided to test an inside-set during evolution and an outside-set after the whole evolutionary process has been completed to validate the generalization performance with the same method (SSP or MSP). Thus, the networks may not be correctly evaluated (misleading fitness) if the single SSP is used during evolution (inside-set) and then the MSP at the end of it (outside-set). The results show that the same prediction method should be used in both evaluation sets providing smaller errors on average over the generalization performance.

Keywords: Evolutionary Algorithms, Artificial Neural Networks, EANNs, Single-step-ahead prediction, Multi-step-ahead prediction

1 Introduction.

Artificial Neural Networks (ANNs) are mathematical models inspired by the structural and functional organization of biological neural networks. They are characterized by having input, hidden and output units with interconnection between them, where each connection has an associated weight which is updated during the training phase to allow the network to learn a given task. Since their origin, they have been used to solve control [1], classification [2, 3] and prediction [4] tasks, showing a performance and adaptability superior to those of conventional mathematical models, as it is actually known that one hidden layer with sufficient nodes can approximate any classification decision boundary [5]. Even though neural networks have proved to be a robust method for solving different kinds of problem, they involve several different parameters that need to be chosen appropriately to obtain a functional network. Early studies used to select many of those parameters

by trial and error [5]. Another difficulty is that some of these parameters may change over time, and thus more elaborate methods are needed to adjust them.

ANNs and Evolutionary Algorithms (EAs) have been widely inspired by biological organisms, usually giving them superior performance when both are applied together to solve a problem than when they are applied in separate stages. Thus, Evolutionary Artificial Neural Networks (EANNs), sometimes called neuroevolution, have been remarkably useful at adapting the ANNs' parameters during evolution [2, 6, 7]. Furthermore, EAs can improve the robustness of networks because they are less likely to be trapped on local minima than traditional gradient-based search algorithms [6].

The usage of EAs over ANNs require an extra error-evaluation, e.g. Normalized Root Mean Squared Error (NRMSE) to determine the fitness of each individuals in the population (this metric is used in this work). Therefore, there may be different evaluation sets within an EANN: validation set to discover overtraining; inside-set to obtain the fitness of an individual during evolution, and a final test set called outside-set to evaluate the generalization performance after the evolution has finished. In this way, inside and outside terms are used to make reference to performance evaluation during and after the evolutionary process has been completed. Besides test sets, a prediction metric is needed, e.g. the Single-step-ahead (SSP) and Multiple-step-ahead prediction (MSP) methods. Note that in classification tasks, an extra test set (validation set) is not required like in the prediction case because the validation set uses the same method to evaluate the fitness during evolution as in the generalization test, i.e. like the SSP method. Consequently, the validation set and inside-set use to be the same in this case.

Thus, this paper is aimed to compare SSP and MSP procedures over both test sets (inside and outside) to forecast two chaotic time series (TS): *Lorenz* and *Mackey-Glass*, usually tested in prediction tasks. Thus, the networks may not be correctly evaluated (misleading fitness) if the single SSP is used during evolution (inside-set) and then the MSP at the end of it (outside-set), i.e. both evaluations may be perform in the same terms. Moreover, no previous study has been found explaining such scenario, which should be tested empirically.

The reminder of this paper is organized as follows: In Sec. 2 is presented an overview of the FS-EPNet algorithm aimed at evolving ANNs using a steady-state algorithm based on Evolutionary Programming. Sec. 3 then gives a description of SSP and MSP for TS forecasting. Thereafter, Sec. 4 presents the experimental set-up with the Lorenz and Mackey-Glass TS, while Sec. 5 describes our empirical prediction results for the comparison between SSP and MSP over the inside-set and outside-set. Finally, we provide our conclusions in Sec. 6.

2 The FS-EPNet Algorithm

The EPNet algorithm [3, 8] is based upon the standard Evolutionary Programming (EP) approach, aimed at evolving ANN architectures and weights at the same time as obtaining smaller network topologies. The original algorithm [3] does not tackle the feature evolution, i.e. input adaptation in the same evolutionary process. However, further improvements consider their evolution [8], i.e. Feature Selection

EPNet algorithm (FS-EPNet), being the algorithm used during this empirical study. The FS-EPNet algorithm emphasizes the evolution of ANN behaviours by EP, like node-splitting, which maintains the behavioural (i.e. functional) link between the parent and its offspring. It does not have a crossover operator, nor a genotype to represent the individuals. Instead it carries out the evolutionary process by performing only nine different mutation operations directly on the phenotype as shown in Fig. 1: (1) hybrid training composed of training with the Modified Back Propagation (MBP) algorithm and Simulated Annealing (SA); (2) node deletion; (3) connection deletion; (4) input deletion; (5) delay deletion; (6) connection addition; (7) node addition; (8) input addition; and (9) delay addition. The algorithm performs only one such mutation on the selected individual in each generation. The training in the EPNet algorithm is only a partial training, i.e. the networks are not trained until they converge. This is motivated by computational efficiency, which lets the evolution advance faster, with the individuals improving their fitness through the generations. For a more detailed description of the EPNet algorithm see [3, 8].

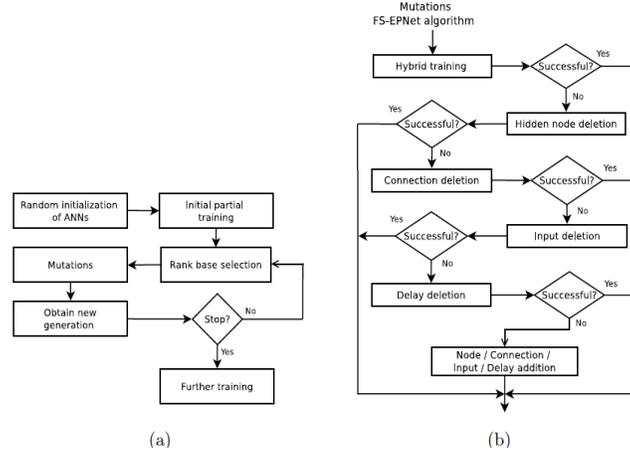


Figure 1. Feature Selection EPNet algorithm (FS-EPNet). General procedure (Fig.1a) and FS-EPNet mutations (Fig. 1b)

Table 1. Single-step prediction

<i>Forecasting</i>	<i>Inputs</i>
y_{t+1}	x_t, x_{t-1}, x_{t-2}
y_{t+2}	x_{t+1}, x_t, x_{t-1}
y_{t+3}	x_{t+2}, x_{t+1}, x_t
y_{t+4}	$x_{t+3}, x_{t+2}, x_{t+1}$

3 Time Series Prediction

For the TS prediction problem with ANNs, it is common to try to use a small subset of recent TS information to perform the prediction. Therefore, we are aiming

to obtain accurate predictions using only a finite segment of previous values up to the point to be predicted. Thus we have:

$$x_{t+1} = F[x_t, x_{t-k}, x_{t-2k}, \dots, x_{t-(d-1)k}] \quad (1)$$

Where d is the number of inputs, k is the time delay and F is the method or algorithm that performs the prediction (the network for this work). There is one condition that needs to be satisfied: given an attractor of dimension D , we must have $d \geq 2D+1$ [9]. There are two general ways to perform the prediction of a TS in terms of the desired number of values to be forecast. Thus, assume the TS X is $[x_1, x_2, \dots, x_t]$, the number of points ahead to predict is n , the test set is $[x_{t+1}, x_{t+2}, \dots, x_{t+n}]$, and the forecast in the same interval is $[y_{t+1}, y_{t+2}, \dots, y_{t+n}]$. In the following examples, we are assuming that the number of inputs (past information) is 3, delays are set at 1 and the prediction step is $\Delta t=1$.

3.1 Single-step-ahead prediction (SSP)

The simplest method is just to predict a value in the future, and we may call this method One-step or Open-loop or Single-step-ahead prediction (SSP). It is called Open-loop forecasting because a pattern is used to predict a value and no feedback is used to continue the predictions as in an autoregressive method. Table 1 shows the single-step prediction method. A sample of previous works that have used (SSP) are [10, 11, 12, 13], where [10, 11] predict the *Lorenz* TS and [12, 13] the *Mackey-Glass* TS.

3.2 Multi-step-ahead prediction

Another interesting prediction method is the Multi-step-ahead prediction (MSP) which uses closed-loop forecasting through an autoregressive method as shown in Table 2.

Table 2. Multiple-step prediction

<i>Forecasting</i>	<i>Inputs</i>
y_{t+1}	x_t, x_{t-1}, x_{t-2}
y_{t+2}	y_{t+1}, x_t, x_{t-1}
y_{t+3}	y_{t+2}, y_{t+1}, x_t
y_{t+4}	$y_{t+3}, y_{t+2}, y_{t+1}$

Note that in Table 2 the predictions are used as input values in subsequent predictions, i.e. it is repeated one-step prediction several times, using the actual prediction to predict the next value. The input vector from the SSP (Table 1) and MSP (Table 2) methods may be seen as a window of d values with k delays that is moved one position ahead every time a value is predicted, to be ready to predict the next value. The real difference between both methods is that the SSP moves the window

input vector over the original data available, meanwhile the MSP starts with the original data, overlap original and predicted data, and finish with predicted values in the window input vector. Previous publications that used the MSP method are [3, 14, 15], where [14, 15] are focused on the Lorenz TS and [3] predicting the Mackey-Glass TS.

3.3 MSP and SSP comparison

For the previous section, it can be said that prediction tasks with SSP are similar to classification tasks as one input vector produce one output vector and there is no feedback in the output as in MSP. Having said that, a standard procedure in the literature is to evaluate an inside-set with the SSP method for classification and prediction tasks to obtain the fitness of individuals. Moreover, any publication has been found so far that uses MSP over the inside-set, that kind of evaluation is never said and it may be assumed to be SSP as it is the standard.

4 Experimental set-up and data sets

As previously remarked, it was decided to use an extra test set during evolution as tasks solved with MSP require to validate the generalization performance with the same method (MSP) when the evolution finish, i.e. validation, inside-set and outside-set. Different to classification task where only the validation set and outside-set is required.

Thus, the networks may not be correctly evaluated (misleading fitness) if the single SSP is used during evolution and then the MSP at the end of it. For example, it can be assume that SSP method is easier than MSP by the feedback in the later, therefore, if a prediction task requiring MSP is evaluated with SSP during evolution (inside-set), it will probably produce a different fitness than if the MSP is used in the same inside-set, which could produce a bias in the selection process with networks not so fit. For that reason it was needed to use an extra test set in prediction tasks, so the validation set is used mainly to evolve the learning rate and the inside-set to measure the fitness as it were a real prediction. However, it may be expected to obtain a smaller fitness error during evolution with SSP than MSP.

There are some common parameters that were fixed for the experiments throughout this study: population size 30, generations of evolution 3000, initial connection density 30%, initial learning rate 0.15, minimum learning rate 0.01, epochs for learning rate adaptation 5, number of mutated hidden nodes 1, number of mutated connections 1-3, temperatures in SA 5, iterations per temperature in SA 100, 1500 epochs of training inside the FS-EPNet, and 2000 of further training at the end of the algorithm. The only stopping criteria was the number of generations. For all the experiments, 30 independent runs were performed to ensure statistical validity of the results. The inside-set was set-up with 30 patterns to perform the prediction and the MSP is performed on the outside-set in all TS tested. All these parameters were set at convenient traditional values and are not intended to be optimal.

Two chaotic TS were used to test the insights presented in this study: a) the first one is the *Lorenz* TS [16] generated with the fourth order Runge-Kutta method as done in [14; **Error! No se encuentra el origen de la referencia.**], i.e. the following values are used to generate the TS: $\Delta t=1$, $\sigma=10$, $r=28$, $\beta=8/3$ and *time step*=0.05 using 1000 values to train and 100 to test; and b) the *Mackey-Glass* TS usually generated with fourth-order Runge-Kutta method as Lorenz TS, where the parameters used here to generate it are: $x(0)=1.2$, $\tau=17$, $\beta=0.2$ and $\beta=-0.1$ as done in [3] using 500 values to train and 500 to test. Therefore, the outside-set for *Lorenz* TS was set to 100 and for *Mackey-Glass* to 500 patterns.

5 Experimental Results

In this section is presented the results from a set of experiments developed to determine if the usage of SSP in the inside-set may degrade the performance of task requiring MSP on the outside-set.

To illustrate this, consider Fig. 2 for the best predictions found for the Lorenz with MSP (Figs. 2a and 2b) and SSP (Figs. 2c and 2d) on the inside-set during evolution and using MSP on the outside-set. Interestingly to note (and as previously expected), the average fitness of the networks evaluated with SSP on the inside-set have a lower error during all the evolutionary process than the fitness obtained from the MSP as can be seen in Fig. 3. The best prediction error on the inside-set at the end of the 3000 generations of evolution were smaller with the SSP than with the MSP as expected too (Table 3, NRMSE Inside-set row). It was also obtained a smaller error with the SSP on the average fitness over all independent trials as shown at the end of generations in Fig. 3.

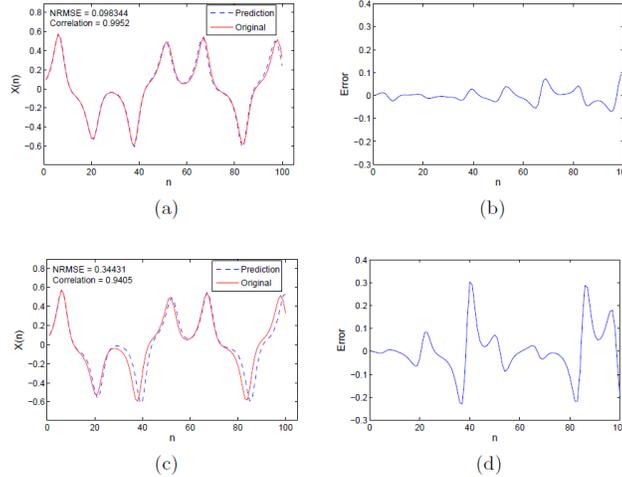


Figure 2. Best predictions found for Lorenz TS with MSP (Figs. 2a and 2b) and SSP (Figs. 2c and 2d) on the inside-set during evolution and using MSP on the outside-set after 3000 generations. Figs. 2b and 2d present the error in terms of $y_i(t) - z_i(t)$, where $y_i(t)$ is the prediction

at time t and $Z_i(t)$ is the original data to be predicted (outside-set).

At the end of the evolution, even the network that uses MSP have a bigger fitness error in the inside-set, it obtained the smallest generalization error, with statistical significant having a p -value < 0.01 (Fig. 2a), because the selection mechanism during evolution was in the same terms as the generalization measurement.

Table 3 presents the individual parameters evolved for the Lorenz TS, showing how the NRMSE over the inside-set is smaller when the MSP is used during evolution than SSP, but the generalization performance is smaller when the MSP is used in both test sets. Also note that using SSP produce the convergence of delays (Table 3 and graphically in Fig. 4d) as it is easier to predict with SSP than MSP for the feedback in the latter as previously remarked. Fig. 4 shows the evolution of hidden nodes (Fig.4a), connections (Fig. 4b), input nodes (Fig. 4c) and delays (Fig. 4d) for the Lorenz TS. There can be seen that evolving those parameter were similar in both cases for hidden, input nodes and connections, i.e. using the SSP or MSP on the inside-set. It may be worth to say too, that an incorrect prediction method over the inside-set may produce poorer networks by the deviation of it parameters as can be seen in Fig. 4d.

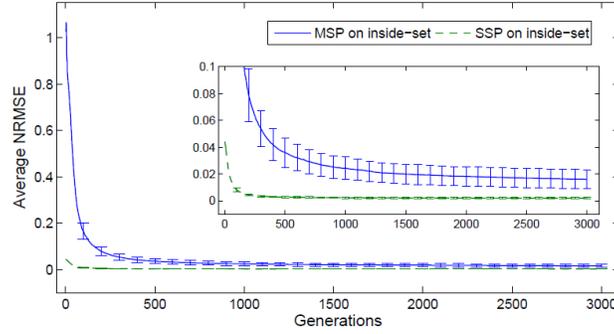


Figure 3. Average fitness value of *Lorenz* TS with MSP and SSP over the inside-set

Table 3. Average fitness value of *Lorenz* TS with MSP and SSP over the inside-set

Parameter	MSP – MSP				SSP – MSP			
	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Number of Inputs	6.70000	1.91455	3	10	6.13333	1.79526	3	9
Number of Delays	2.46666	0.93710	1	4	1	0	1	1
Number of Hidden Nodes	13.33333	4.25346	5	26	12.8	3.28423	6	18
Number of Connections	108.466	56.1989	39	351	92.9333	35.5614	43	172
NRMSE Validation Error	0.02475	0.02848	0.00062	0.12563	0.00145	0.00104	0.00035	0.00447
NRMSE Inside-set	0.01657	0.00775	0.00676	0.03480	0.00189	0.00084	0.00123	0.00477
NRMSE Outside-set	0.52301	0.26162	0.09834	0.92639	0.73256	0.23340	0.34431	1.20320

Comparing these results against results found in the literature, Dudul [14] obtain a NRMSE for the best individual with a State-space 8th order of NRMSE = 0.1822 and a NRMSE = 0.7325 with a Regularized ANN-ARMA, while the FS-EPNet obtain a NRMSE = 0.09834 (Table 3) for the best individual found.

Table 4. Mackey-Glass time series individual results

Parameter	MSP – MSP				SSP – MSP			
	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Number of Inputs	8.33333	2.39732	5	13	8.10000	1.26899	5	9
Number of Delays	3.56667	0.85836	2	6	3.16667	0.74664	2	4
Number of Hidden Nodes	16.0666	6.17522	6	31	12.5666	3.94517	7	23
Number of Connections	150.600	78.1499	52	376	114.866	51.2214	55	263
NRMSE Validation Error	0.00567	0.00835	0.00049	0.04147	0.00182	0.00112	0.00061	0.00466
NRMSE Inside-set	0.00330	0.00152	0.00119	0.00706	0.00377	0.00191	0.00163	0.01075
NRMSE Outside-set	0.16810	0.04626	0.07397	0.30965	0.17183	0.31555	0.01696	1.41379

To finalize the results of this work, in Table 4 is presented the individual results of evolving the Mackey-Glass TS. It can be seen how the usage of MSP over the inside-set produced smaller average errors than SSP, however in this case there where no statistically significance in the results. Comparing these results against previous studies, e.g. [17], it was found that the FS-EPNet could not improve them, requiring a more analysis and further research.

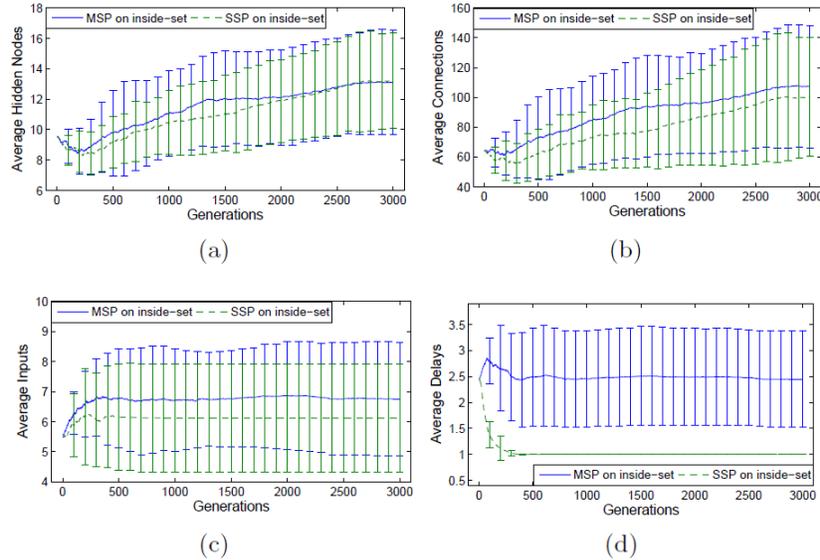


Figure 3. Best predictions found for the *Lorenz* TS with MSP and SSP over the inside-set for hidden (Fig. 4a, connections 4b), input (Figs. 4c and delays 4d) during 3000 generations

6 Conclusions and Discussion

In this work was presented a comparison between two prediction methods, Single-step-ahead (SSP) and Multi-step-ahead, during the evolution of Artificial Neural Networks (ANNs) for time series (TS) prediction. The experiments were carried out using the EPNet algorithm designed to evolve ANNs architectures and weights simultaneously through and steady-state procedure. From two chaotic TS tested (Lorenz and Mackey-Glass), it was determined that tasks that use SSP will use SSP for the fitness during evolution and to evaluate the generalization performance. Contrary, tasks that use MSP will use the same MSP method in both parts of the process. Further research is required to test a broad range of TS to generalize these results, however, from this study it can be expected that the same method should be used in both stages.

References

1. Stanley, K.O., Miikkulainen, R.: Evolving neural networks through augmenting topologies. *Evolutionary Computation* 10(2) (2002) 99–127
2. Bullinaria, J.A.: Understanding the emergence of modularity in neural systems. *Cognitive Science* 31(4) (2007) 673–695
3. Yao, X., Liu, Y.: A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks* 8(3) (1997) 694–713
4. Cholewo, T.J., Zurada, J.M.: Sequential network construction for time series prediction. *International Conference on Neural Networks* 4 (Jun 1997) 2034–2038
5. Bishop, C.M.: *Neural Networks for Pattern Recognition*. Oxford University Press (November 1995)
6. Yao, X.: Evolving artificial neural networks. *Proceedings of the IEEE* 87(9) (1999) 1423–1447
7. Bullinaria, J.A.: Evolving neural networks: Is it really worth the effort? In: *Proceedings of the European Symposium on Artificial Neural Networks*, Evree, Belgium: d-side (2005) 267–272
8. Landassuri-Moreno, V., Bullinaria, J.A.: Feature selection in evolved artificial neural networks using the evolutionary algorithm EPNet. In: *Proceedings of the 2009 UK Workshop on Computational Intelligence*. UKCI '2009, Nottingham, UK: University of Nottingham. (July 2009)
9. Beldjoudi, J., Contreras, D.: Recurrence plots in nonlinear time series analysis: Free software. *Journal of Statistical Software* 7(9) (2002)
10. Rojas, I., Pomares, H., Bernier, J.L., Ortega, J., Pino, B., Pelayo, F.J., Prieto, A.: Time series analysis using normalized pg-rbf network with regression weights. *Neurocomputing* 42(1-4) (2002) 267–285
11. Gholipour, A., Araabi, B.N., Lucas, C.: Predicting chaotic time series using neural and neurofuzzy models: A comparative study. *Neural Processing Letters* 24(3) (2006) 217–239
12. Müller, K.R., Smola, A.J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., Vapnik, V.: Using support vector machines for time series prediction. (1999) 243–253
13. Müller, K.R., Smola, A.J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., Vapnik, V.: Predicting time series with support vector machines. In: *ICANN '97: Proceedings of the 7th International Conference on Artificial Neural Networks*, London, UK, Springer-Verlag (1997) 999–1004

14. Dudul, S.V.: Prediction of a lorenz chaotic attractor using two-layer perceptron neural network. *Applied Soft Computing* 5(4) (2005) 333–355
15. Guerra, F.A., dos S. Coelho, L.: Multi-step ahead nonlinear identification of lorenz's chaotic system using radial basis neural network with learning by clustering and particle swarm optimization. *Chaos, Solitons & Fractals* 35(5) (2008) 967 – 979
16. Lorenz, E.N.: Deterministic nonperiodic flow. *Journal of atmospheric Science* 20 (1963) 130–141
17. Assaad, M., Boné, R., Cardot, H.: A new boosting algorithm for improved time-series forecasting with recurrent neural networks. *Information Fusion* 9(1) (2008) 41–55 Special Issue on Applications of Ensemble Methods.2. De Mol, L.: Tracing Unsolvability: A Mathematical, Historical and Philosophical analysis with a special focus on Tag Systems. PhD thesis, University Gent (2007)

Integración de métodos para la clasificación-reconocimiento de patrones metalográficos a base de inteligencia artificial y procesamiento digital de imágenes

E. A. Ruelas¹, J. A. Vázquez-López², J. Yañez-Mendiola¹, I. López-Juárez³,
C. F Bravo-Barrera⁴

¹ CIATEC, León, Gto. Mexico
edruelas@itesi.edu.mx
jyanez@ciatec.mx

² Instituto Tecnológico de Celaya, Celaya, Guanajuato., México
antonio.vazquez@itcelaya.edu.mx

³ Centro de Investigación y de Estudios Avanzados del IPN – Unidad Saltillo, Ramos Arizpe,
Coahuila., Mexico
ismael.lopez@cinvestav.edu.mx

⁴ Laboratorio de Pruebas Equipos y Materiales. Irapuato, Guanajuato., Mexico
fernando.bravo01@cfe.gob.mx

Paper received on 24/07/12, Accepted on 11/09/12.

Resumen. Este artículo propone un método para el reconocimiento de patrones en materiales de acero al carbón mediante la red neuronal artificial Fuzzy ArtMap y procesamiento digital de Imagen. Un aspecto importante que se aborda es dejar ver la viabilidad para la integración de estos dos métodos para desarrollar un sistema de reconocimiento de diversos estados de aceros al carbón en base a imágenes obtenidas de microscopio usando la red neuronal fuzzy ArtMap.

Palabras clave: Red neuronal artificial (RNA), procesamiento digital de imágenes.

1 Introducción

La metalografía es la disciplina que estudia microscópicamente las características estructurales de un metal o de una aleación. Sin duda, el microscopio es la herramienta más importante del metalurgista tanto desde el punto de vista científico como desde el técnico. Es posible determinar el tamaño de grano, forma y distribución de varias fases e inclusiones que tienen gran efecto sobre las propiedades mecánicas del metal. La microestructura revelará el tratamiento mecánico y térmico del metal y, bajo un conjunto de condiciones dadas, podrá predecirse su comportamiento esperado.

El comportamiento esperado del material y la falla de mismo presentan patrones de comportamiento que pueden ser analizadas por un sistema basado en la clasifica-

ción-reconocimiento de patrones por medio de redes neuronales y procesamiento digital de imágenes. La red neuronal artificial (RNA) clasificara el patrón de comportamiento que arroje el análisis de la microestructura del material, eliminando la subjetividad del operador en el sistema a través del entrenamiento exhaustivo de la RNA.

Las herramientas que se utilizaran serán: RNA, Lógica difusa y procesamiento digital de imágenes. La primera se emplea para reconocer y clasificar patrones de comportamiento dentro de la microestructura del material, la lógica difusa se incorpora en la arquitectura de la RNA para obtener una decisión más robusta en cuestión al reconocimiento y clasificación de los patrones observados en la microestructura, el procesamiento digital de imágenes es un punto clave, ya que por medio de este se alimenta al sistema enfocado a la clasificación-reconocimiento de patrones. Una parte importante del problema es la complejidad un sistema que integre los métodos y que además sirva como método de diagnóstico basado en la clasificación-reconocimiento de patrones y procesamiento de imágenes. La otra parte del planteamiento del problema se refiere a la necesaria optimización de uso del sistema, de tal forma que se mantengan bajas las probabilidades de los errores tipo I.

Se han realizado investigaciones donde comprueban la viabilidad y beneficios de la integración entre RNA y procesamiento digital de imágenes enfocados al reconocimiento de patrones. Se propone un sistema para la resolución de píxeles basados en múltiples problemas de clasificación de imágenes espectrales [8]. El tema de integración de ambos métodos se ha probado en el campo de la medicina, la técnica de procesamiento de imágenes y redes neuronales artificiales para detectar y clasificar la tuberculosis en imágenes de las diapositivas de tejidos [7]. Otra aplicación exitosa fue desarrollada por [10], sistema mediante el cual una red neuronal de propagación hacia atrás con un enfoque para clasificar defectos a partir de imágenes por sus convexos y brillo. En la ref. [9], se presenta el uso de redes neuronales art empleadas para el reconocimiento de imágenes en cuanto al color y textura. La arquitectura presentada logra una buena segmentación y reconocimiento, en comparación con otros métodos. Siguiendo con la aplicación de la RNA's en la ref. [3], se estudian diversas aplicaciones de la red neuronal Fuzzy ARTMAP, en reconocimiento de patrones dentro del procesamiento de señales de diferentes fuentes, imágenes, voz y texto.

La colaboración entre las redes neuronales artificiales y lógica difusa puede ser utilizado para mejorar o resolver algunas de las limitaciones que presentan cada una de ellas; estos nuevos sistemas híbridos, llamados sistemas neuro-difusos, desarrollan las propiedades y ventajas propias de cada tecnología en beneficio de la otra tecnología complementaria, obteniendo una mejora importante en el comportamiento global del sistema [2].

Dentro del ambiente de sistemas que combinan redes neuronales y lógica difusa se encuentran trabajos enfocados al ramo alimenticio. En la ref. [1], presenta una herramienta visual desarrollada sobre el entorno Matlab dedicada al control de calidad del proceso. En la ref. [6], propone una metodología de integración entre redes neuronales y Lógica difusa para hacer evaluaciones y predicciones que conllevan a la optimización de fabricación de arrabio. Otra aplicación se desarrolla en la implementación y evaluación de filtros tanto para el pre-proceso como para la detección de bordes en imágenes digitales, los cuales se basan en lógica difusa [4]. Un sistema

de procesamiento de imagen se compone de una serie de subsistemas que operan sobre una escena con el objetivo de interpretar alguna característica notable. La materia prima del procesamiento de imágenes y la visión son las imágenes, las cuales se consideran como una representación del mundo físico que tiene información importante, la cual es captada mediante un proceso de muestreo, generalmente por medios electrónicos.

De acuerdo con [5], una imagen puede estar definida con una función de dos dimensiones denotada $f(x, y)$, donde x y y denotan la posición de una muestra en donde se mide la intensidad de la luz. Cuando se muestrea tanto espacialmente como en amplitud una señal bidimensional, hablamos de una imagen digital, al tratamiento que se le da a la imagen digital a través de la información que se obtiene por el par ordenado, se le denomina procesamiento digital de imágenes. Las imágenes son representadas por medio de matrices con M renglones y N columnas. Donde la coordenada $f(0,0)$ de la imagen corresponde a $f(1,1)$ de la matriz y así sucesivamente. Se ha demostrado la efectividad del software MATLAB para el procesamiento de imágenes, mediante la herramienta de simulink [5].

2 Fuzzy ArtMap

La red neuronal artificial utilizada para la investigación es la Fuzzy ARTMAP la cual cuenta con una arquitectura perteneciente a la familia de ART (Adaptive Resonance Theory). Esta teoría fue desarrollada por Grossberg y Carpenter en la Universidad de Boston. Fuzzy ARTMAP puede efectuar aprendizaje supervisado. Esta red neuronal crea varias neuronas de acuerdo al número de patrones presentados en su entrenamiento y a las diferencias entre ellos. Esta red tiene dos módulos Fuzzy ART, uno para manejar las entradas y el segundo para las salidas (ARTa y ARTb). También tiene un campo de correspondencia (F^{ab}) que liga las clases de entrada con las de salida. Ver figura 1.

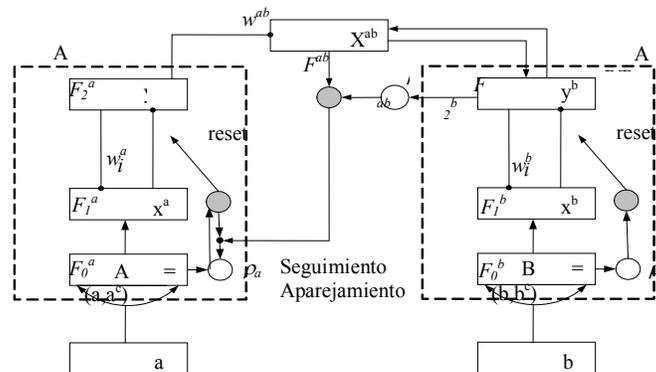


Figura 1. Estructura de la RNA Fuzzy ARTMAP.

La forma de operación de un módulo básico (ARTa ó ARTb) opera como un módulo básico ART como se ilustra en la figura 3. Dicho modulo consta de 2 sub-

sistemas, atención y orientación. El primer subsistema consta de dos capas de nodos hacia arriba, F1 y F2 que representan la memoria de corto plazo (STM), que a su vez están unidas por los pesos ajustables de largo plazo (LTM).

Así, la información de salida del elemento de procesamiento reverbera hacia atrás y hacia adelante entre las dos capas. Si una resonancia estable toma lugar entonces puede ocurrir aprendizaje o adaptación. El subsistema de orientación reinicia el subsistema de atención cuando ocurre un evento no común. Si el patrón de entrada no es reconocido inmediatamente entonces la red entrará a un estado de resonancia, después de lo cual un nuevo patrón es almacenado por primera vez. De este modo la red responde rápidamente a datos aprendidos previamente.

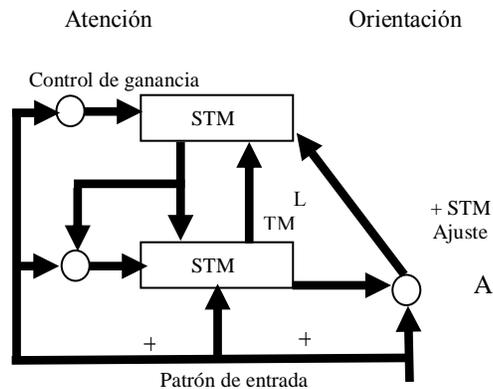


Figura 2. Funcionamiento de la RNA Fuzzy ARTMAP.

El parámetro de vigilancia de la red (ρ) mide la diferencia permitida entre los datos de entrada y los patrones almacenados. Por tanto este parámetro es el determinante para afectar a la selectividad o granularidad de la predicción de la red. Este parámetro tiene 3 factores: vigilancia a la entrada (vigilancia base) (ρ_a), vigilancia de salida (ρ_b) y vigilancia de mapeo (ρ_{ab}).

3 Metodología Experimental

El caso de aplicación para esta investigación esta siendo desarrollado dentro del Laboratorio de Pruebas de Equipos y Materiales (LAPEM) de la comisión federal de electricidad (CFE).

Específicamente enfocado en el área de materiales. Para los intereses de la empresa y la investigación se pretende desarrollar un sistema experto con las siguientes características:

- Reconocer y clasificar patrones de comportamiento en el deterioro del material (aceros al carbón) afectado por temperatura.
- Identificar el tipo de material y predecir el tiempo de vida de dicho material después del uso en campo en una escala de 1 a 10.

Los patrones detectados presentes en la microestructura para el deterioro del material, con afectación por temperatura son los siguientes:

- Crecimiento de grano.
- Crecimiento de espesor magnetita.
- Afectaciones Cristalográficas.
- Granos de Perlita en proceso de Esferoidización.
- Grafitización.
- Descarborización.

Cabe mencionar que los datos obtenidos acerca de patrones de comportamiento, tipo de material y afectaciones son a partir de imágenes. Por lo cual el procesamiento digital de imágenes es una base teórica primordial dentro de esta investigación.

Las pruebas realizadas se limitan solo a los patrones de granos dispersos, granos de perlita en proceso de esferoidización y granos finos de Perlita Laminar. Cada uno de ellos con características propias generados a partir de que el material sufrió afectaciones por temperatura en campo.

A continuación se muestran ejemplos de materiales afectados por temperatura en pruebas de campo y su microestructura:

i. Granos dispersos



Figura 3. Material afectado por temperatura (Granos dispersos).

Tabla 1. Características generales del patrón granos dispersos (Fig 3).

Nombre	Tamaño	Bytes	Clase
F	235x290x3	211344	uint 8

Tabla 2. Características generales del patrón esferoidización (Fig 4).

Nombre	Tamaño	Bytes	Clase
A	235x247x3	174876	uint8

ii. **Granos de Perlita en proceso de Esferoidización**



Figura 4. Material afectado por temperatura (Esferoidización).

iii. **Granos finos de Perlita Laminar**



Figura 5. Material afectado por temperatura (Perlita Laminar).

Tabla 3. Características generales del patrón perlita laminar (Fig 5).

Nombre	Tamaño	Bytes	Clase
B	171x212x3	108756	uint8

3.1 Procesamiento de la imagen

Se presenta un diagrama de flujo indicando las operaciones de manera detallada en el proceso de extracción y limpieza de ruido de la imagen metalográfica obtenida de los materiales a partir de microscopio.

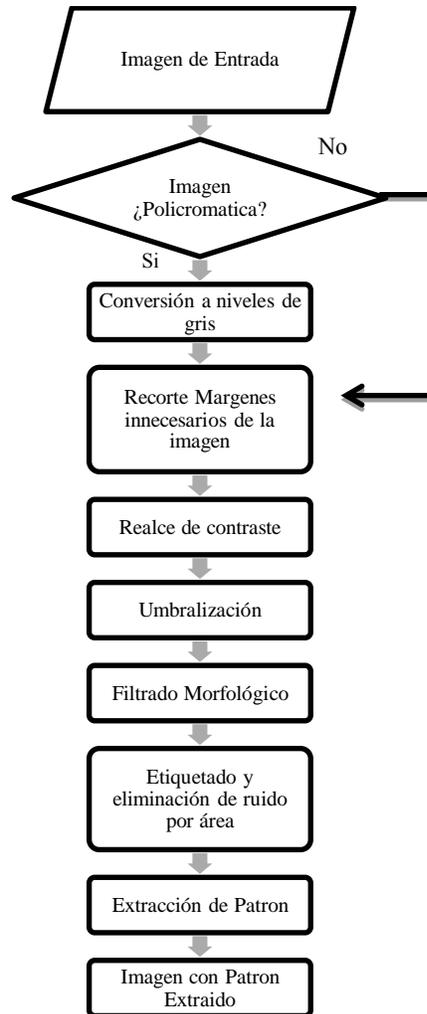


Figura 6. Diagrama de proceso para la extracción del patrón presente en la Imagen.

El procesamiento digital de la imagen es un paso necesario llevado a cabo en este trabajo mediante técnicas ya conocidas dentro de esta disciplina. El procesamiento se muestra en la Fig. 7. Este Proceso surge de la necesidad de hacer más evidente el patrón de comportamiento y de esta forma esperar una mejor eficiencia en la RNA en el reconocimiento del patrón.

3.2 Fase de Entrenamiento

En la presente investigación se plantean dos técnicas de entrenamiento para la red neuronal Fuzzy ArtMap que son: Matriz desacoplada y por familia. Y de esta forma medir la eficiencia de cada una en cuanto a la clasificación-reconocimiento de patrones que para este caso es en el área de metalografía.

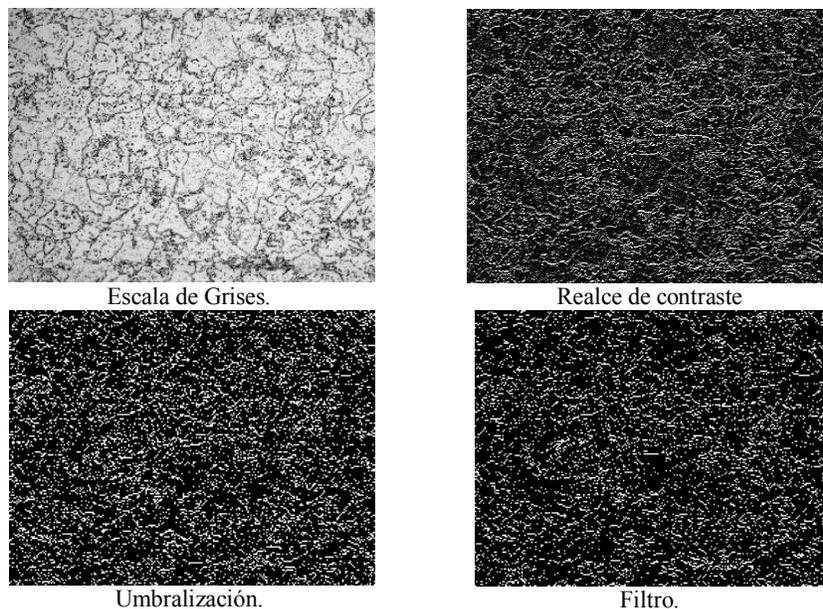


Figura 7. Procesamiento digital de la Imagen.

Cabe mencionar que una de las limitantes que presenta la RNA Fuzzy ArtMap es que para la fase de entrenamiento solo acepta a lo más matrices de $n \times 50$. De aquí surgen las dos alternativas para el entrenamiento de la RNA que cumplan con la restricción de su arquitectura.

- Matriz desacoplada.

El término surge por la idea de convertir la matriz de entrenamiento de 5×10 . A un vector de entrenamiento de 1×50 . Que cumple con la restricción de tamaño que presenta la RNA.

- Familia de Vectores.

Otra alternativa de entrenamiento evaluada es la de tomar familias de vectores de 1×10 . Que al darles un mismo vector de salida la RNA tomará como el mismo elemento o patrón. Las familias se realizaron de cinco vectores para formar una matriz de 5×10 . Con esta forma de entrenamiento también se desea probar que el vector de entrenamiento no debe ser del mismo tamaño que el vector de prueba. Los parámetros de funcionamiento para esta fase de entrenamiento son los que se muestran en la tabla 4.

Tabla 4. Entrenamiento de la RNA.

Parámetros de entrenamiento	
Art A	50,10
Base Vigilance	0.2
Rho Map	0.7
Learn Rate	1
Lower Weight Bound	0.5
Art B	3
Vigilance	0.9

3.3 Fase de Prueba

Para realizar esta actividad es necesario establecer la región de prueba en la imagen, dentro de esta región se analizará si el comportamiento descrito en la fase de entrenamiento está presente por medio de la RNA.

Se seleccionaron dos regiones de prueba al azar dentro de toda la imagen en cada una de las microestructuras analizadas en la investigación. El tamaño de cada una de las regiones es de $n \times 50$.

Tabla 5. Entrenamiento de la RNA.

Parámetros de Prueba	
Art A	50,10
Base Vigilance	0.2
Rho Map	0.7
Learn Rate	1
Lower Weight Bound	0.5

4 Resultados

En la tabla 6 y 7 se muestra la eficiencia de la red neuronal para detectar el patrón de comportamiento granos dispersos, esferoidización y perlita laminar presentado en las regiones de prueba de su respectiva microestructura.

Tabla 6. Entrenamiento Matriz desacoplada.

	Prueba 1	Prueba 2
Granos dispersos	89.0%	91.5%
Esferoidización	95.4%	97.4%
Perlita Laminar	92.5%	94.3%

Tabla 7. Entrenamiento Familia de Vectores.

	Prueba 1	Prueba 2
Granos dispersos	93.9%	92.7%
Esferoidización	91.9%	90.2%
Perlita Laminar	85.7%	84.8%

5 Conclusión

El modulo de la RNA Fuzzy ArtMap tiene una eficiencia aceptable en la clasificación-reconocimiento de patrones explorados en este artículo a pesar de realizar la fase entrenamiento con vectores cortos. Con lo que se puede concluir que es posible establecer protocolos de clasificación-reconocimiento de patrones metalográficos por medio de redes neuronales artificiales que posibilite un análisis predictivo a partir de un procesamiento digital de imágenes. Es imprescindible desarrollar mejores técnicas de operación y dimensionalidad de la RNA para poder ofrecer un mejor diagnóstico de la imagen y de esta forma lograr su implementación en campo.

Referencias

- 1 Al-Hadithi, B., Rueda., M., Muro, J. (2006). Interfaz visual para el prototipado rápido de clasificadores de gajos de mandarina basados en redes neuronales.
- 2 Basogain O. X., (2004). Redes neuronales artificiales y sus aplicaciones. Escuela superior de ingeniería de Bilbao.
- 3 Boaz L. and Guterman H. (2008). Advanced Developments and Applications of the Fuzzy ARTMAP Neural Network in Pattern Classification. Volume 137/2008, 77-107.
- 4 Cardona., C., (2004). Evaluación de algoritmos basados en lógica difusa aplicados al pre-proceso y detección de bordes en imágenes digitales. Universidad Nacional de Colombia.
- 5 Gonzalez R. C; Woods R. E y Eddins S. L. (2009). Procesamiento digital de imagen utilizando MATLAB. Editorial Pearson.
- 6 Hussen, Ghidan (2008). Aplicación de las redes neuronales y la lógica difusa a la optimización del proceso de la fabricación del arrabio en un horno alto. Universidad Alcala.
- 7 Osman M. K., Mashor M. Y. and Jaafar H. (2011). Hybrid Multilayered Perceptron Network Trained by Modified Recursive Prediction Error-Extreme Learning Machine for Tuberculosis Bacilli Detection. Volume 35, Part 15.
- 8 Reid. P., Neal H., Simon P., James T., Steven B., Jeffrey B., Maya G and Jonh S., (2003). Optimizing Digital Hardware Perceptrons for Multi-Spectral Image Classification. Volume 19, Number 2.
- 9 Rodríguez M., González D., Díaz-Pernas F. J., Martínez-Zarzuela M. and J. F. Díez-Higuera. (2012). Color-texture image segmentation and recognition through a biologically-inspired architecture. Volume 22, Number 1.
- 10 Shih-Wei L., Shuo-Yan C., and Shih-Chieh. (2007). Irregular shapes classification by back-propagation neural networks. Volume 34, Numbers 11-12.

DL Prolog: Another Unifying Programming Language

José Oscar Olmedo-Aguirre

Departamento de Ingeniería Eléctrica, Cinvestav, Ciudad de México, D.F., México
oolmedo@cinvestav.mx

Paper received on 30/07/12, Accepted on 16/08/12.

Abstract. Programming language research has been guided by the hope of identifying a minimal set of expressive enough programming concepts that can describe uniformly and consistently most computational problems along with their solutions. DL Prolog unifies the logic, functional and imperative programming paradigms by the introduction of dynamic logic modalities into pure Prolog with equational unification. The imperative fragment of DL Prolog uses equations between terms to define logical variables and a novel assignment to redefine imperative variables. The contribution of this paper is in introducing these two operations in the context of the imperative fragment of DL Prolog that are formally described in the structured operational semantics style. A working research prototype has been built upon this design and it is available on request from the author.

Keywords: Programming Language, Logic Programming, Dynamic Logic, Imperative Programming, Unification, Prolog.

1 Introduction.

Unified programming research attempts to identify the minimal set of the simplest and most fundamental programming concepts that coherently can be applied in as many as diverse areas of application and still can lead to reasonable efficient implementations. Among the most successful efforts in conciliating such a diversity of concepts, the functional logic programming paradigm is capable of embedding the notion of state that characterizes the imperative programming, along with constraint satisfaction, concurrent processing and object orientation, among others. Nonetheless, there is neither symmetry nor balance in their integration in the sense that a Java programmer, for example, needs to learn a number of sophisticated concepts that are uncommon in her experience to start programming even relatively simple programs. In this respect, an extension to pure Prolog with dynamic logic modalities, named DL Prolog [8], has been designed by the author to provide a more balanced approach in the sense that the same Java programmer has not to learn too many concepts if she does not need to. Instead of placing the imperative programming at the top of the functional logic programming paradigm as in Curry [4], the unified approach of DL Prolog stands upon the imperative, functional and logic programming paradigms. Unfortunately, due to lack of space, the previous claim will

only be justified by discussing some program examples that will also help to show the programming style of the imperative fragment of DL Prolog.

This article is organized as follows. In section 2, a succinct review of the related work is presented. In section 3, the programming examples describe the use of unification. In section 4, a brief account of the programming style of the imperative fragment of DL Prolog is shown. In section 5, a formal description in the structured operational semantics is given for the imperative fragment. Finally in section 6, some further research directions and concluding remarks are given.

2 Related work.

Classical papers on the integration of functional, equational and logic programming were collected in [1]. Among them, Kahn's Uniform language is the closest to ours in its use of unification though his proposal was developed for the functional language LISP, whereas DL Prolog applies unification for its imperative fragment [2]. Uniform is an AI language that uses augmented unification to solve equational problems that cannot be solved by syntactic unification. Probably the most serious difficulty faced in the design of Uniform was the computational complexity of finding a solution in presence of equality theories that lead to large chains of equalities. In fact, solving equational problems in presence of equality theories is known to be a NP complete problem [3]. In DL Prolog the combinatorial production of equalities has been drastically reduced by (i) incorporating a greater degree of control into the language by means of functions instead of equations, and (ii) avoiding comparing two functional terms, so any equation can relate at least one constructor-rooted term. Thus DL Prolog unification prevents applying higher-order unification which is a NP-complete problem [3].

As discussed by Kahn, Uniform lacked of more theoretical foundations in equational unification to deal with the combinatorial search involved in the generation of solutions. Taking advantage of the reasonably efficient implementations of functional languages, the integration efforts were later oriented to approach relations as Boolean functions. Curry [4] is a functional logic programming language based on the evaluation by need of expressions (a mechanism called *narrowing* used in lazy functional programming) along with the possible instantiation of free variables occurring in the expressions (a mechanism called *residuation* used in constraint-logic programming). Nonetheless, imperative programming is no directly available but through a monadic IO system that enable sequence of input-output actions. Despite its growing success and acceptance, the use of evaluation by need makes difficult to reason about program behavior and memory requirements. This means that a programmer needs to understand the not so intuitive IO monadic system in order to write even the simplest imperative programs.

Concurrent programming, particularly the so-called algebraic theory of processes has provided a basis for its integration with the functional, imperative and object-oriented paradigms. The language PICT based on the π -calculus is an example of this approach [5]. In addition to the formidable task of implementing enough programming concepts as networks of interacting processes, the several layers that will emerge from such design makes no clear how efficient and of practical use would be

this integration. However, the concurrent programming approach has also led to more successful pragmatic language designs like Oz. Oz is termed as a multi-paradigm programming language supporting logic, functional, imperative, object-oriented, constraint, concurrent and distributed paradigms [6]. Oz is based on residuation, a simple yet powerful mechanism to solve concurrently constraints by processes that suspend their execution whenever a variable is undefined in a predicate and resume their execution soon after the variable becomes defined.

3 Unification.

The importance of unification and equational unification (E-unification) is due to the fact that it is widely used in automated theorem proving and related fields like logic programming. The term unification generally stands for syntactic equality on terms, whereas semantic E-unification generally stands for syntactic equality modulo an equational theory. Augmented unification is a restricted version of E-unification to be used as a model of execution [2]. For a brief account of unification and E-unification, consider the following DL Prolog equations to define variables. Hereinafter, variables are typed in italics:

1. $[u, b, c, d] = [a \mid vs]$
2. $(us, a, x, [x \mid xs]) = ([a, b], a, y, us)$
3. $f(us, [us]) = g([a, b], vs)$
4. $[x, y \mid xs] = rev([a, b])$
5. $[x, y \mid xs] = rev([a])$

As term constructors are injections (like the list and the tuple constructors), solving the equation $t = s$ amounts to finding an assignment of values to variables that makes terms t and s syntactically identical. Unification transforms the equation $t = s$ into a set of consistent equations formed by comparing the corresponding subterms of t and s whenever they have the same root constructors and the same number of subterms. Thus, the equation (1) is transformed into the set of equations $\{u = a, [b, c, d] = vs\}$, where variables u and vs are respectively bound to a and $[b, c, d]$. Similarly the equation (2) is transformed into the set of equations $\{us = [a, b], a = a, x = y, [x \mid xs] = us\}$. From this set, the first equation defines the value of us , the second can be discarded, the third is recorded (as x and y remain unknown so far), and the last leads to the equations $x = a$ and $xs = [b]$. After finding that $x = a$, variable y becomes also bound to a . In (3), the unification fails because the compared terms have different rooted constructors f and g , and there is not function definitions for either f or g (but not both) whose evaluation may result in a constructor-based term that could be used instead in the equation.

The last two examples deal with E-unification. In (4), sole unification cannot solve the equational problem because the constructor of the list is not equal to the function name rev that stands for the usual reverse function. As unification fails, E-unification is applied by searching for an equation or function defined for rev . If the search succeeds, the function is evaluated with the given parameters and its result is replaced where the function application occurs; otherwise, the E-unification fails and the equational problem has no solution. For this example, because there is a function definition with the usual meaning for rev , equation $rev([a, b]) = [b, a]$ holds

and equation (4) becomes $[x, y \mid xs] = [b, a]$, producing the solutions $x = b, y = a$ and $xs = []$. Finally, in (5) even E-unification fails because assuming that function `rev` is defined, the result of its application on `[a]` has only one element whereas the list $[x, y \mid xs]$ can only unify with lists of at least two elements.

4 A DL Prolog Program Example.

The following DL program shows the dynamic logic modality introduced in the pure Prolog language and model. The program simply checks that the well-known property of lists $\text{rev}(\text{app}(xs, ys)) = \text{app}(\text{rev}(ys), \text{rev}(xs))$ holds for the lists `[a, b, c, d]` and `[e, f, g, h]`.

```
[ new ys, zs: (
  xs = [a, b, c, d]; ys := [e, f, g, h];
  zs := rev(app(xs, ys)); us = zs; writeln(revapp : us);
  zs := app(rev(ys), rev(xs)); vs = zs; writeln(apprev : vs)
)
] (us = vs).
```

This DL Prolog program consists of two parts: (i) the modal connective that encloses in brackets the well-formed imperative program: `new ys, zs: (xs = [a, b, c, d], ...)`, and (ii) the postcondition $us = vs$ that states the properties of the values bound to the variables at the program termination. The property tested is valid in a theory of lists with the usual operations $\text{rev}(xs)$ and $\text{app}(xs, ys)$ that respectively reverses the order of the elements of list xs and appends the elements of list ys at the end of list xs . In the program, the *logical variable* xs is defined with list `[a, b, c, d]`, whereas the *local imperative variable* ys is defined with list `[e, f, g, h]`. Logical variables can be defined by equality at most once, whereas local variables can arbitrarily be redefined as many times as needed by means of the destructive assignment of imperative programming. In any case, the occurrence of either a logical or local variable in an expression denotes the value bound to it. Program instructions are sequentially executed as usual from left to right of the semicolon connective. The assignment $zs := \text{rev}(\text{app}(xs, ys))$ defines the local variable zs with a list that results from reversing the concatenation of lists xs and ys . This list is used to define the logical variable us that is next written in the terminal output preceded by label `revapp`. The assignment $zs := \text{app}(\text{rev}(ys), \text{rev}(xs))$ redefines the local variable zs with the concatenation of the reversed lists of ys and xs . This list is used to define the logical variable vs that is next written preceded by label `apprev`. Beyond this point in the program text, the scope of the binder `new` reaches its end and the local variables ys and zs cease their existence. Nonetheless, logical variables us and vs remain unaltered at the program postcondition $us = vs$. The partial correctness of the reverse program ensures that the postcondition is always satisfied upon its termination whenever the input parameter satisfies a suitable precondition. Hence there is no need to test the postcondition as much as the partial correctness of the reverse program can be proved.

4.1 The reverse program

The DL Prolog program for the reverse function $rev(xs) = zs$ is presented next in a stylized format that helps to show its syntactic structure:

$$\left[\text{new } u, us, vs : \left(\begin{array}{l} (us, vs) := ([], xs); \\ ([u \mid vs] := vs; us := [u \mid us])^*; \\ zs = us \end{array} \right) \right] rev(xs) = zs :- list(xs).$$

The above program text is equivalent to its usual linear presentation $[\text{new } u, us, vs : ((us, vs) := ([], xs); ([u \mid vs] := vs; us := [u \mid us])^*; zs = us?)] rev(xs) = zs :- list(xs)$ that can freely be formatted according to the programmer preferences. The above program is in fact the partial correctness property of the program to compute the reverse function. Like Prolog, the reverse program in DL Prolog is a Horn clause consisting of a simple conclusion predicate called the *head of the clause* and an antecedent formed by a conjunction of predicates called the *body of the clause*. Unlike Prolog, the head of the clause in DL Prolog is a *dynamic logic assertion* consisting of the reverse program enclosed in brackets followed by the program postcondition $rev(xs) = zs$. A DL Prolog program like reverse *fails*, either if it does not terminate, or if upon its completion, it cannot satisfy its postcondition.

The imperative programming constructs used in the DL Prolog assertions comprise the most common structured programming connectives for sequential, conditional and iterative composition. The program $(A_1; A_2)$ stands for the sequential composition of the program (fragments) A_1 and A_2 . The non-deterministic execution $(A_1 :: A_2)$ stands for the non-deterministic selection of one of two program fragments A_1 and A_2 : if both programs terminate, the non-deterministic composition terminates by arbitrarily choosing one; if both programs fail, the non-deterministic composition fails too; otherwise, if one program fails, the accompanying unailing program is chosen. This behavior is required to introduce the usual conditional composition of programs $(C?; A)$, where a possibly failing program fragment C , called *the guard*, prevents the execution to continue with program A if C is false. The usual if-then-else construct $(\text{if } C \text{ then } A_1 \text{ else } A_2)$ is defined by the non-deterministic selection $(C?; A_1 :: \neg C?; A_2)$ between two programs A_1 and A_2 guarded with mutually exclusive conditions C and $\neg C$. The unary suffix connective (A^*) stands for the iterative execution of A , where the repeated execution of A continues until it fails. Furthermore, the language is block structured, where local variables are introduced in the block by binder new. Local variables correspond to the variables of usual imperative languages that can be redefined by means of a newly assignment operator. The multiple-assignment operator extends the single assignment by assigning to a list of variables the values of a list of terms of the same length. Logical variables are mainly used to name values because they can be defined at most once, whereas local variables are used to store partial computations.

Though the previous program connectives have the usual meaning, the assignment has a remarkably unusual meaning as it is shown by analyzing the three assignments of the reverse program. In the assignment $(us, vs) := ([], xs)$, local varia-

bles us and vs are initially defined by the empty list $[]$ and the list xs to be reversed, respectively. In the assignment $[u \mid vs] := vs$, the list bound to vs is decomposed in a first element and in a list with all the elements of vs but the first. If vs contains at least one element, the decomposition is possible, and the assignment succeeds by binding the first element of vs to u and by removing the first element of the list bound to vs . However, if vs is the empty list, the decomposition of vs is not possible. Hence the assignment fails and so does the sequential composition in that it appears, causing the termination of the iteration. Finally, in the third assignment $us := [u \mid us]$, the variable us is redefined by the list $[u \mid us]$. The term $[u \mid us]$ denotes a list whose first element is u and the rest is us . Thus, after the assignment, the list us is as before augmented with the value of u placed at its first element. Note that among the three assignments, only $[u \mid vs] := vs$ may fail due to the constraint that both lists at each side of the assignment must agree in their number of elements. In the context of the iterative connective in which this assignment occurs, this failing behavior is needed to ensure termination. At the end of the program fragment, the elements accumulated in list us define the value of the output variable zs .

4.2 The reverse postcondition

For the reverse program, the postcondition $rev(xs) = zs$ asserts that the elements of zs , if any, are reversed order with respect to those of xs whenever the precondition $list(xs)$ holds. The postcondition of the reverse program establishes the formal relation between the input and the output variables, comprising three elementary conditions: (i) xs and zs must be both lists, (ii) the length of xs is equal to the length of zs , and (iii) when xs contains at least one element, the element of xs at position i is equal to the element of zs at position $length(xs) - i$, for any index i between 0 and $length(xs) - 1$. Fortunately, the partial correctness property of the program rev ensures that is not necessary to check that the postcondition holds at the end of each execution.

The modal assertion of DL Prolog extends logic programming with functions defined by imperative programs in a novel manner. Next the formal description of the language syntax and semantics is briefly presented, emphasizing the central role of E-unification in equations and assignments.

5. Formal Description.

The signature Σ, Ξ of DL Prolog comprises all its symbols grouped in two disjoint sets: the set $\Sigma = \{a, b, c, \dots, f, g, \dots\}$ of constructor and function names and the set $\Xi = \{u, v, x, y, \dots\}$ of variable names. Constants are constructors with no structure. The set $T(\Sigma, \Xi)$ of terms with variables is the minimal set containing constants, variables and that is closed under the composition of constructors and functions from Σ with a (possibly empty) sequence of terms. The set $T(\Sigma)$ of ground terms consists of terms with no variables. The following grammar rules describe the syntactic structure of the main categories of the language built upon the signature Σ, Ξ :

Terms $T(\Sigma, \Xi)$:

$$t ::= x \mid c \mid f(t_1, \dots, t_n)$$

Predicates $P(\Sigma, \Xi)$:

$$P ::= \perp \mid \top \mid t_1 = t_2 \mid p(t_1, \dots, t_n)$$

Goals $G(\Sigma, \Xi)$:

$$G ::= P \mid G_1, G_2$$

(Horn) Clauses $H(\Sigma, \Xi)$:

$$H ::= P \mid P :- G$$

Programs $A(\Sigma, \Xi)$:

$$A ::= t_1 = t_2 \mid t_1 := t_2 \mid C? \mid A_1 ; A_2 \mid A_1 :: A_2 \mid A^* \mid (A) \mid \text{new } x_1, \dots, x_n : A$$

Modal programs $M(\Sigma, \Xi)$:

$$M ::= P \mid [A] M$$

Modal clauses $C(\Sigma, \Xi)$:

$$C ::= M \mid M :- G$$

where the entities of each category are distinguished by indexing. The symbols \perp and \top stand for the truth values false and true, respectively. Horn clauses and modal clauses are assumed to be universally closed. The set of variable names $\text{fv}(F)$ occurring free in a (modal) clause F occurring within the scope of a binder are those not captured by the binder, i.e. $\text{fv}(\text{new}(x) F) = \text{fv}(F) - \{x\}$ for the binder new and similarly for the universal and existential quantifiers.

Substitutions model program states. A *program state* is the set of bindings between variables and their bound values. Substitutions are partial functions $S(\Sigma, \Xi) = \Xi \rightarrow T(\Sigma, \Xi)$ from variables to terms. A *substitution* $\sigma : S(\Sigma, \Xi) = \{x_1 \rightarrow t_1, \dots, x_n \rightarrow t_n\}$ binds a set $\{x_1, \dots, x_n\}$ of different variables, called the domain of the substitution $\text{dom}(\sigma)$, with a set $\{t_1, \dots, t_n\}$ of terms whenever x_i does not occur in t_i for $i = 1, \dots, n$. A substitution is admissible if and only if none of the variables of t_i becomes bound by a binder after its substitution for all $i = 1, \dots, n$. A *ground substitution* $\sigma : \Xi \rightarrow T(\Sigma)$ is a substitution on ground terms. A substitution σ is naturally extended as a map $T(\Sigma, \Xi) \rightarrow T(\Sigma, \Xi)$ from terms to terms and also for the other syntactic categories of the language. The application $t\sigma$ of a substitution σ on a term t (or the instance $t\sigma$ of t under σ) is a term obtained from the simultaneous replacement of x_i by t_i for $i = 1, \dots, n$, such that:

- $c \sigma = c$, for all $c \in \Sigma$
- $x \sigma = x$, for all $x \in \Xi$ and x not in $\text{dom}(\sigma)$
- $x_i \sigma = t_i$, for all x in $\text{dom}(\sigma)$
- $f(x_1, \dots, x_n) \sigma = f(x_1 \sigma, \dots, x_n \sigma)$
- $[A]M \sigma = [\sigma][A]M$

where notation $[\sigma]$ stands for the conjunction of the equations $x_1 = t_1, \dots, x_n = t_n$, also written as $(x_1, \dots, x_n) = (t_1, \dots, t_n)$. These equations are written in *solved form* whenever they are satisfied by the ground substitution $\sigma = \{x_1 \rightarrow t_1, \dots, x_n \rightarrow t_n\}$. Note that unlike the application of a substitution to terms, the application to the program modality $[A]M \sigma$ is defined as the initial binding of the program variables $[\sigma][A]M$, i.e. not for the usual simultaneous replacement of all the variables of σ through the entire program A . The composition $\sigma_1 \sigma_2$ of substitutions $\sigma_1 = \{x_1 \rightarrow t_1, \dots, x_n \rightarrow t_n\}$ and $\sigma_2 = \{y_1 \rightarrow s_1, \dots, y_m \rightarrow s_m\}$ is defined as:

$$\begin{aligned} \sigma_1 \sigma_2 &= \{x_i \mapsto t_i \sigma_2 \mid x_i \in \text{dom}(\sigma_1) \wedge t_i \sigma_2 \notin \text{dom}(\sigma_1), i = 1, \dots, n\} \\ &\cup \{y_i \mapsto s_i \mid y_i \in \text{dom}(\sigma_2) \wedge y_i \notin \text{dom}(\sigma_1), i = 1, \dots, m\} \end{aligned}$$

The unification of two terms either finds a unifier, i.e. a substitution that solves the equational problem or terminates in failure. Unification is a step by step transformation process between sets of equations until no further transformations can be applied. The unification procedure is defined through the following rules.

- $(\{t = t\} \cup P, S) \rightarrow (P, S)$
- $(\{f(t_1, \dots, t_n) = f(s_1, \dots, s_n)\} \cup P, S) \rightarrow (\{t_1 = s_1, \dots, t_n = s_n\} \cup P, S)$
- $(\{f(t_1, \dots, t_n) = g(s_1, \dots, s_m)\} \cup P, S) \rightarrow (\{\}, \{\perp\})$
- $(\{x = t, x = s\} \cup P, S) \rightarrow (\{x = s, t = s\} \cup P, S)$ if $x = t \succeq t = s$ for some appropriate well-founded lexicographic ordering \succeq
- $(\{x = t\} \cup P, S) \rightarrow (\{\}, \{\perp\})$ if $x \in \text{fv}(t)$ and $x \neq t$
- $(\{x = t\} \cup P, S) \rightarrow (P\{x \mapsto t\}, S\{x \mapsto t\} \cup \{x = t\})$ if $t \notin \Xi, x \notin \text{fv}(t)$

By applying the unification rules to the equational problem $t = s$, the pair $(\{t = s\}, \{\})$ is transformed into the pair $(\{\}, S)$, meaning that the set S of equations in solved form is the solution to the problem; otherwise, the pair $(\{t = s\}, \{\})$ is transformed into the pair $(\{\}, \{\perp\})$, meaning that the equational problem has no solution.

The semantic description of DL Prolog is established according to the *structural operational semantics* [7] by defining program descriptions and transitions between program descriptions. Assuming the termination of the program, a *program description*, or simply a *description*, can be either instantaneous or final. An *instantaneous description* $I(\Sigma, \Xi) : M(\Sigma, \Xi) \times S(\Sigma, \Xi)$ is a pair $([A]M, \sigma)$ relating a program modality $[A]M$ and a program state σ , meaning that the program A starts its execution in the state σ , reaching a state that is assumed to satisfy M whenever A terminates. Note that M may be either another program modality or a predicate. For the former case, a new instantaneous description can be established, whereas for the latter case, a final description is reached. A *final description* $F(\Sigma, \Xi) : S(\Sigma, \Xi) \vee \{\perp\}$ corresponds to the state reached when the program terminates or fails. The final description of a program that terminates successfully is a state $\sigma : S(\Sigma, \Xi) = \{x_1 \mapsto t_1, \dots, x_n \mapsto t_n\}$, containing the bindings of the variables x_i with their final values t_i for $i = 1, \dots, n$. Instead, the final description of a program that terminates in failure is represented by \perp . A *transition* is a relation over pairs of program configurations $\rightarrow : I(\Sigma, \Xi) \times (I(\Sigma, \Xi) \vee F(\Sigma, \Xi))$. A *execution of a program* corresponds to the reflexive and transitive closure of the transition relation.

The execution of a terminating program is a finite sequence of instantaneous descriptions ended by a final description. For the successful program execution:

$$(M_0, \sigma_0), (M_1, \sigma_1), \dots, (M_i, \sigma_i), (M_{i+1}, \sigma_{i+1}), \dots, (M_n, \sigma_n), \sigma_n, \text{ for some } n \geq 0,$$

the program starts with the initial modality $M_0 = [A]P$ along with the variables initialized according to σ_0 and the program terminates reaching the state σ_n , satisfying the partial correctness property. Thus, $[A]P\sigma_0$ implies $P\sigma_n$. This property follows by induction on the length n of the sequence of instantaneous descriptions, where $M_i \sigma_i$ implies $M_{i+1} \sigma_{i+1}$ for all $i = 1, \dots, n$. For the unsuccessful program execution:

$$(M_0, \sigma_0), (M_1, \sigma_1), \dots, (M_i, \sigma_i), (M_{i+1}, \sigma_{i+1}), \dots, (M_n, \sigma_n), \perp\perp, \text{ for some } n \geq 0,$$

the program starts like before, though it now terminates in failure, meaning that nothing can be asserted about the final program state.

The following inference rules define the SOS semantics of the imperative fragment of DL Prolog:

1.
$$\frac{(\{f(t_1, \dots, t_n)\sigma = f(s_1, \dots, s_n)\sigma\}, \{\}) \rightarrow (\{\}, \{x_1 = r_1, \dots, x_m = r_m\})}{([\![f(t_1, \dots, t_n) = f(s_1, \dots, s_n)]\!]M, \sigma) \triangleright (M, \sigma \{x_1 \mapsto r_1, \dots, x_m \mapsto r_m\})}, \text{ with } 0 \leq n$$
2.
$$([\![f(t_1, \dots, t_n) = g(s_1, \dots, s_m)]\!]M, \sigma) \triangleright \perp\perp \text{ if } f \text{ and } g \text{ are different and there are not defined functions for either } f \text{ or } g, \text{ with } 0 \leq m, n$$
3.
$$\frac{(\{g(s_1\sigma, \dots, s_m\sigma) = s, t\sigma = s\}, \{\}) \rightarrow (\{\}, \{x_1 = r_1, \dots, x_n = r_n\})}{([\![t = g(s_1, \dots, s_m)]\!]M, \sigma) \triangleright (M, \sigma \{x_1 = r_1, \dots, x_n = r_n\})} \text{ if there is defined a function } g \text{ such that } g(s_1\sigma, \dots, s_m\sigma) = s, \text{ with } 0 \leq m, n$$
4.
$$([\![t := s]\!]M, \sigma_1\sigma_2) \triangleright ([\![t = s\sigma_1\sigma_2]\!]M, \sigma_1) \text{ if } s\sigma_1\sigma_2 \text{ is a ground term and } \sigma_1 = \{x \mapsto r \mid x \notin fv(t)\}, \sigma_2 = \{x \mapsto r \mid x \in fv(t)\}$$
5.
$$([\![t := s]\!]M, \sigma) \triangleright \perp\perp \text{ if } s\sigma \text{ is not a ground term or the terms } t \text{ and } s\sigma \text{ do not unify}$$
6.
$$([\![\top ?]\!]M, \sigma) \triangleright (M, \sigma)$$
7.
$$([\![\perp ?]\!]M, \sigma) \triangleright \perp\perp$$
8.
$$([\![A_1 ; A_2]\!]M, \sigma) \triangleright ([\![A_1]\!] [\![A_2]\!]M, \sigma)$$
9.
$$([\![A_1 :: A_2]\!]M, \sigma) \triangleright ([\![A_1]\!]M, \sigma)$$
10.
$$([\![A_1 :: A_2]\!]M, \sigma) \triangleright ([\![A_2]\!]M, \sigma)$$
11.
$$([\![A^*]\!]M, \sigma) \triangleright ([\![\top ? :: A; A^*]\!]M, \sigma)$$
12.
$$([\![\]\!]p, \sigma) \triangleright \sigma$$

Rules from (1) to (3) describe the equational definition. In (1), the unification of two constructor-based terms with identical constructor is converted into the equational problem $f(t_1, \dots, t_n)\sigma = f(s_1, \dots, s_n)\sigma$. If the problem has a solution, the

solution is composed with the current state. In (2), the equation $f(t_1, \dots, t_n) = g(s_1, \dots, s_n)$ has no solution if f and g are different constructors or at most one of them have no definition as a function whose evaluation does not unify with the other. In (3), the equation $t = g(s_1, \dots, s_n)$ can be solved if there is a functional definition of g and both the application $g(s_1 \sigma, \dots, s_n \sigma)$ and the term $t\sigma$ are equals to a term s . Rules (4) and (5) describe the assignment instruction. In (4), the assignment instruction succeeds if the instance of the term at the right-hand side under $\sigma_1\sigma_2$ is a ground term and all the variables occurring in the term at the left-hand side have become unbound from their values. In (5), the assignment fails either if $s\sigma$ is neither a ground term nor $s\sigma$ unifies with t . In (6), the guard true, named skip, always succeeds, whereas in (7) the guard false, named fail, always fails. In (8), the sequential execution of $A_1 ; A_2$ corresponds to the execution of A_1 followed by the execution of A_2 ; otherwise $A_1 ; A_2$ fails if either A_1 or A_2 fails. In (9) and (10), the non-deterministic execution $A_1 :: A_2$ of A_1 and A_2 corresponds to the selection of the un-failing execution of any of them. In the case that both succeeded, one of them is chosen, whereas if both fail, the instruction fails too. In (11), the iterative execution of A is obtained by non-deterministically choosing between skip and unwinding the iteration with the sequence $A ; A^*$. Finally, in (12) the program modality terminates when no instructions remain to execute. In this case, if the program executed is partially correct with postcondition p , then $p\sigma$ holds.

6 Conclusions and Future Work.

In this paper the use of a restricted form of E-unification has been proposed for the imperative programming fragment of the DL Prolog language. E-unification extends unification of terms to include equational theories encoded as functions, gaining some control at the cost of limiting its expressive power. Nonetheless this reduced expressivity brings in more efficiency as required for most practical applications. E-unification is used to define logical variables by term unification and to redefine imperative variables by the assignment instruction.

Further research directions include: the proof of the soundness and completeness results of the imperative fragment, the design and implementation of DL Prolog through an extension to the WAM, and its comparison with other programming languages like standard Prolog in terms of their expressiveness and efficiency.

References

1. DeGroot, D., Lindstrom, G. (eds.): Logic Programming: Functions, Relations and Equations. Prentice-Hall, 1986.
2. Kahn, K.M.: Uniform : A Language Base upon Unification which unifies much of Lisp, Prolog and Act 1. In [1], 411-440.
3. Baader, F., Nipkow. T.: Term Rewriting and All That. Cambridge, 1998.
4. Antoy, S., Hanus, M.: Functional Logic Programming. Communications of the ACM, 53(4):74-85, 2010.
5. Sewell, P., Wojciechowski, P.T., Unyapoth, A.: Nomadic PICT: Programming Languages, Communication Infrastructure Overlays, and Semantics of Mobile Computation. ACM Transactions of Computer Languages, 32(4):12:1-12:63, 2010.

6. Van Roy, P.: Multiparadigm programming in Mozart/Oz. Second International Conference MOZ 2004, 2005.
7. Reynolds, J.C.: Theories of Programming Languages. Cambridge, 1998.
8. Olmedo-Aguirre, J.O., Morales-Luna, G.: A Dynamic-Logic-based Modal Prolog. In proceedings MICAI 2012. To appear.

Clasificador Asociativo de Patrones Iterativo (CAPI)

Sergio Valadez¹, Raúl Santiago¹ y Humberto Sossa²

¹ División de Estudios de Posgrado e Investigación-Instituto Tecnológico de León,
Av. Tecnológico s/n, León, Gto., México
svaladezg@gmail.com, rsantiago66@gmail.com

² Centro de Investigación en Computación-IPN,
Av. Juan de Dios Bátiz s/n, México, D.F., 07738. México
hsossa@cic.ipn.mx

Paper received on 21/08/12, Accepted on 07/09/12.

Resumen. En este trabajo se presenta una hibridación del llamado Clasificador Asociativo de Patrones (CAP) y la Memoria Asociativa Bidireccional (BAM). La nueva propuesta (CAPI) permite mejores desempeños de clasificación. No requiere el ajuste de los parámetros manualmente. Es capaz de clasificar correctamente un patrón en muy pocas iteraciones. Además, bajo ciertas condiciones, puede clasificar correctamente todos los patrones de entrenamiento.

Palabras Clave: Clasificación de patrones, Iteración, Memorias Asociativas.

1 Introducción.

Las *clases* son estados asociadas con conceptos [1]. El conjunto de todas las clases es conocido como el *espacio de interpretación* [2]. Los rasgos mediante los cuales se caracteriza a los objetos son conocidos como el *espacio de representación*. Entonces, el problema de la clasificación de patrones es encontrar un algoritmo que permita dividir el espacio de interpretación en diferentes regiones y así separar los patrones conocidos y clasificar los patrones desconocidos [3, 4, 5, 6]. Una manera de hacer esto es mediante el uso de memorias asociativas (MAS), las cuales permiten la clasificación de patrones mediante la asociación de los mismos con una clase [7]. Una memoria asociativa (MA) es una red neuronal artificial (RNA) de una sola capa la cuál mapea un conjunto de patrones de entrada \mathbf{x}^k hacia un conjunto de patrones de salida \mathbf{y}^k de tal manera que cada patrón \mathbf{x}^k es asociado con un patrón \mathbf{y}^k [8]. Una memoria M puede representarse mediante el esquema: $\mathbf{x}^k \rightarrow M \rightarrow \mathbf{y}^k$. En este caso M es una matriz de correlación de p asociaciones [9], cuyo *conjunto fundamental* se representa como:

$$S = \{(\mathbf{x}^k, \mathbf{y}^k) | k = 1, 2, \dots, p\} \quad (1)$$

En el *proceso de construcción* de la memoria M , cada par $(\mathbf{x}^k, \mathbf{y}^k) \in S$ es presentado a la MA. En el *proceso de recuperación* de la memoria M , un patrón de entrada \mathbf{x}^ω se presenta a la entrada de la memoria. Si $m = n$ entonces la memoria es *autoasociativa* y de otro modo es *heteroasociativa* [10].

En [11, 12, 13] se propone un algoritmo llamado *Clasificador Asociativo de Patrones* (CAP) basado en memorias asociativas. Combina la regla de construcción del *Asociador Lineal* (AL) de Anderson [14], Kohonen [10] y Nakano [15] y la regla de recuperación de la *Lernmatrix* (LM) [16, 17]. La nueva propuesta permite la operación con números reales, eliminando la desventaja de la LM de operar solo en el conjunto de los números binarios. También elimina la restricción de ortogonalidad del conjunto de entrenamiento del AL [18] así como la restricción de que el número de los patrones del conjunto de entrenamiento deba ser pequeño con respecto a la dimensión de los patrones de entrada [19, 20]. Además, se mantiene estable en su rendimiento de clasificación al ser entrenado con al menos el 10% del total de patrones de una base de datos dada [21].

La Memoria Asociativa Bidireccional (BAM) es una memoria heteroasociativa de dos capas que procesa información hacia adelante y hacia atrás para producir una búsqueda de los patrones almacenados [22, 23]. Cuando las neuronas de la memoria son activadas, la red evoluciona a un estado estable de dos patrones, un patrón en la salida de cada capa. La evolución de la BAM se detiene cuando ésta llega a un mínimo local de energía. La BAM se diferencia del AL y de la LM por que procesa información a través del tiempo de manera bidireccional, aunque también son utilizadas para almacenar pares de asociación. El funcionamiento de la BAM está basado principalmente sobre el funcionamiento de la memoria de Hopfield [24].

En el presente trabajo se propone una hibridación del CAP con la BAM. Esta hibridación permite obtener mejores resultados de recuperación que con el CAP solo. En forma iterada, los parámetros son ajustados automáticamente. A pesar del proceso iterativo, nuestra propuesta es capaz de clasificar correctamente un patrón en muy pocas iteraciones. Además, bajo ciertas condiciones, el nuevo clasificador puede clasificar correctamente todos los patrones de entrenamiento.

2 Estado del arte.

Son muy pocos los trabajos sobre clasificación de patrones mediante memorias asociativas y procesos de clasificación dinámicos. Esto debe en gran medida a que la meta de las memorias asociativas es el reconocimiento de patrones y no el de la clasificación [25]. Por ejemplo, en [26] se propone el uso de las memorias asociativas alfa-beta [27] combinadas con la BAM (sin el proceso iterativo). Esta combinación permite clasificar pacientes que han presentado una reaparición de cáncer. En [28] la BAM es utilizada para clasificar pacientes con cáncer, al ser probada para diferentes umbrales de la función de activación y después de realizar un proceso de normalización y discretización de los patrones. En [29] la BAM es utilizada para clasificar imágenes binarias de caracteres manuscritos y de diodos encendidos.

3 Diseño del Clasificador Asociativo de Patrones Iterativo.

Dado un conjunto de patrones asociados:

$$S = \{(\mathbf{x}^k, \mathbf{y}^k) | k = 1, 2, \dots, p\} \quad (2)$$

Donde p es la cardinalidad del conjunto fundamental de asociaciones, $\mathbf{x}^k \in \mathbb{R}^n$ es el conjunto de patrones de entrada y $\mathbf{y}^k \in \{-1, 1\}^m$ es el conjunto de clases; n es la dimensión de \mathbf{x}^k y m es la dimensión de \mathbf{y}^k . La clase $c \in \{1, 2, \dots, m\}$ a la que pertenece cada patrón de entrada \mathbf{x}^k se define mediante:

$$y_j^k = \begin{cases} 1 & \text{para } j = c \\ -1 & \text{para } j = 1, 2, \dots, c-1, c+1, \dots, m \end{cases} \forall k \in \{1, 2, \dots, p\} \quad (3)$$

3.1 Mecanismo de hibridación.

Debido a que el CAP funciona sobre los números reales y la BAM sobre los bipolares, se plantea un mecanismo de bipolarización de patrones reales a bipolares con el objetivo de realizar la hibridación. El proceso es como sigue:

1. Se calcula un vector de puntos de corte que servirá para bipolarizar los centroides de cada clase:

$$\theta = \frac{1}{p} \sum_{k=1}^p \mathbf{x}^k \quad (4)$$

2. Se calcula el ξ -ésimo centroide de cada clase para $\xi \in \{1, 2, \dots, m\}$:

$$\bar{\mathbf{x}}^\xi = \frac{1}{p_\xi} \sum_{j=1}^{p_\xi} \mathbf{x}^j \quad (5)$$

Donde $\bar{\mathbf{x}}^\xi$ es el ξ -ésimo centroide, p_ξ es el número de patrones de la ξ -ésima clase y $\mathbf{x}^j \forall j \in \{1, 2, \dots, p_\xi\}$ es el conjunto de patrones que forman la ξ -ésima clase.

3. Los centroides $\bar{\mathbf{x}}^\xi$ se bipolarizan mediante:

$$x_i^b = \begin{cases} -1 & \text{si } \bar{x}_i^\xi < \theta_i \\ 1 & \text{si } \bar{x}_i^\xi \geq \theta_i \end{cases} \forall i \in \{1, 2, \dots, n\} \quad (6)$$

Donde x_i^b es el i -ésimo elemento bipolarizado del centroide, \bar{x}_i^ξ es el i -ésimo componente del centroide y θ_i es el i -ésimo componente del vector de puntos de corte. De esta manera se obtiene un nuevo conjunto de patrones asociados:

$$S = \{(\mathbf{x}^b, \mathbf{y}^b) | b = 1, 2, \dots, p\} \quad (7)$$

donde p es la cardinalidad del conjunto fundamental de asociaciones, $\mathbf{x}^b \in \{-1, 1\}^n$ es el conjunto de centroides bipolarizados y $\mathbf{y}^k \in \{-1, 1\}^m$ es el conjunto de clases; n es la dimensión de \mathbf{x}^k y m es la dimensión de \mathbf{y}^k .

3.2 Fase de construcción.

La *fase de construcción* de la matriz de entrenamiento M se realiza de la misma manera que en el CAP, pero ahora sobre el conjunto fundamental bipolarizado S como sigue:

$$M = \sum_{b=1}^m \mathbf{y}^b [\mathbf{x}^b]^t \quad (8)$$

3.3 Fase de recuperación.

Sea $\mathbf{x}^\omega \in \mathbb{R}^n$ un patrón de entrada de dimensión n . El patrón de salida $\mathbf{y}^\omega \in \{-1, 1\}^m$ recuperado a partir de \mathbf{x}^ω se determina mediante los siguientes pasos:

1. El patrón de entrada \mathbf{x}^ω se bipolariza mediante la siguiente ecuación para obtener así un nuevo patrón bipolar \mathbf{x}^b :

$$x_i^b = \begin{cases} -1 & \text{si } x_i^\omega < \theta_i \\ 1 & \text{si } x_i^\omega \geq \theta_i \end{cases} \quad \forall i \in \{1, 2, \dots, n\} \quad (9)$$

Donde x_i^b es el i -ésimo elemento bipolarizado del patrón, x_i^ω es el i -ésimo componente del patrón y θ_i es el i -ésimo componente del vector de puntos de corte.

2. Se define un patrón de clase $\mathbf{y}^\omega = \mathbf{0}$.
3. Se define $\mathbf{x}^b(t) = \mathbf{x}^b$ y $\mathbf{y}^b(t) = \mathbf{y}^\omega$ para $t=0$.
4. El patrón $\mathbf{x}^b(t)$ se aplica en la entrada de la primer capa mediante:

$$y_i^b(t+1) = \begin{cases} 1 & > 0 \\ y_i^b(t) & \text{si } \sum_{j=1}^n m_{ij} x_j^b(t) = 0 \\ -1 & < 0 \end{cases} \quad (10)$$

Donde $y_i^b(t+1)$ es el i -ésimo componente del patrón de salida en el instante de tiempo siguiente, $y_i^b(t)$ es el i -ésimo componente del patrón de salida en el instante de tiempo actual, m_{ij} es el ij -ésimo elemento de la matriz de pesos M y $x_j^b(t)$ es el j -ésimo componente del patrón de entrada en el instante de tiempo actual.

5. Suponiendo que las actualizaciones en la ecuacion (10) son síncronas, y el vector \mathbf{y} se propaga hacia la segunda capa, este se procesa usando la matriz transpuesta M^t mediante:

$$x_j^b(t+1) = \begin{cases} 1 & > 0 \\ x_j^b(t) & \text{si } \sum_{i=1}^m m_{ij}y_i^b(t+1) = 0 \\ -1 & < 0 \end{cases} \quad (11)$$

Donde $x_j^b(t+1)$ es el j -ésimo componente del patrón de salida en el instante de tiempo siguiente, $x_j^b(t)$ es el j -ésimo componente del patrón de salida en el instante de tiempo actual, m_{ij} es el ij -ésimo elemento de la matriz de pesos M y $y_i^b(t+1)$ es el i -ésimo componente del patrón de entrada calculado en el paso 4.

La secuencia de recuperación se repite a través de los pasos 4 y 5 hasta que no hay más actualizaciones de los patrones de salida \mathbf{x} y \mathbf{y} .

3.4 Modos de actualización.

Para el caso de una actualización síncrona, todas las neuronas del vector de salida de las ecuaciones (10) y (11) son actualizadas simultáneamente. Para el caso de una actualización asíncrona solo una neurona del vector de salida es actualizada de manera aleatoria. Por lo tanto solo una componente del vector de salida es calculada, según la salida correspondiente.

4 Evaluación del Clasificador Asociativo de Patrones Iterativo.

En esta sección se presentan los resultados obtenidos al comparar los dos algoritmos: el CAP y el CAPI. Este último en los modos de actualización síncrona y asíncrona. Para identificar en que modo de actualización se está probando el algoritmo propuesto, se define como Clasificador Asociativo de Patrones Iterativo Síncrono-CAPIS cuando se utiliza una actualización síncrona y Clasificador Asociativo de Patrones Iterativo Asíncrono-CAPIA cuando se utiliza una actualización asíncrona.

4.1 Descripción de los datos de prueba.

Cada una de las bases de datos que se utilizaron en la experimentación se obtuvo del repositorio de aprendizaje de máquina de la Universidad de California Irvine (UCI) [30]. En la Tabla 1 se presenta un resumen de cada una de ellas: *Wisconsin Breast Cancer Database*, *Pima Indians Diabetes Database*, *Haberman's Survival Data*, *Heart Disease Dataset*, *Hepatitis Domain*, *Johns Hopkins University Ionosphere database*, *Chess End-Game - King+Rook versus King+Pawn on a7*, *BUPA liver disorders* y *Sonar, Mines vs. Rocks*.

Tabla 1. Resumen de las bases de datos utilizadas para la experimentación.

No.	Base de datos	Instancias			Atributos	Tipo de los atributos
		Total	Clase 1	Clase 2		
1	Breast	699	458	241	9	Numérico
2	Diabetes	768	500	268	8	Numérico
3	Haberman	306	225	81	3	Numérico y nominal
4	Heart	270	120	150	13	Numérico
5	Hepatitis	155	32	123	19	Numérico y nominal
6	Ionosphere	351	126	225	34	Numérico
7	KingRook	3196	1669	1527	36	Nominal
8	Liver	345	145	200	7	Numérico
9	Sonar	208	111	97	60	Numérico

Los valores nominales de las bases de datos Hepatitis y KingRook fueron pre-procesados. En la base de datos Hepatitis, los valores “yes” fueron asignados numéricamente como 1 y los valores “no” como 2, mientras que “female” como 1 y “male” como 2. En la base de datos KingRook, los valores “f” se reemplazaron por 1, los valores “t” por 2, los valores “l” por 1, los valores “g” por 2, los valores “b” por 1, los valores “n” por 2 y los valores “w” por 3.

4.2 Método de validación de resultados.

Para la obtención de los resultados se utilizó el método k -Folds de validación cruzada para $k = 10$. Cada una de las bases de datos se dividió en diez partes de aproximadamente el mismo tamaño. En cada una de las diez partes, las clases fueron representadas en las mismas proporciones, es decir, se hizo una partición estratificada. Además, las instancias de cada una de las diez partes fueron seleccionadas de forma aleatoria. Cada parte se utilizó como conjunto de prueba exactamente una vez y la unión de las nueve particiones restantes se utilizó como conjunto de entrenamiento. Con esto se obtuvieron diez pares prueba-entrenamiento, los cuales fueron pasados a cada uno de los tres clasificadores. Este experimento se repitió diez veces.

4.3 Análisis del desempeño.

Los resultados de los experimentos realizados se muestran en la Tabla 2. Se pueden observar el rendimiento promedio con intervalo de confianza (μ_1 para el CAP, μ_2 para el CAPIS y μ_3 para el CAPIA,) y la desviación estándar (μ_1 para el CAP, μ_2 para el CAPIS y μ_3 para el CAPIA). En negrita aparecen los mejores rendimientos de clasificación.

Para saber si existe diferencia entre el rendimiento de un algoritmo y otro se utiliza la prueba de ranqueo de signos con suma de Wilcoxon [5]. Esta prueba considera que las distribuciones de los datos no siguen una distribución normal. Primeramente se plantea la hipótesis nula de que el primer algoritmo tiene menor o igual rendimiento que el segundo y una hipótesis alternativa de que el primero tiene mayor rendimiento que el segundo. La hipótesis nula se rechaza cuando $W^- < W^+$. En las Tablas 3-5 se muestra el desarrollo de esta prueba considerando los rendimientos promedio del CAP vs el CAPIS, del CAP vs el CAPIA y de CAPIS vs el CAPIA. En la Tabla 3 se puede observar que $W^- > W^+$ por lo tanto se acepta la hipótesis nula. Entonces el CAP tiene menor o igual rendimiento que el CAPIS. En la Tabla 4 se

puede observar que $W^- < W^+$ por lo tanto se rechaza la hipótesis nula. Entonces el CAP tiene mayor rendimiento que el CAPIA. En la Tabla 5 se puede observar que $W^- > W^+$ por lo tanto se rechaza la hipótesis nula. Entonces el CAPIS tiene mayor rendimiento que el CAPIA.

4.4 Iteraciones promedio.

En la Tabla 7 se muestran las iteraciones promedio obtenidas para el CAPIS y CAPIA al recuperar (clasificar) un patrón dado, considerando las mismas pruebas con la que se obtuvieron los resultados de la Tabla 1.

Tabla 2. Rendimientos de clasificación promedio del CAP, CAPIS y CAPIA.

No.	Base de datos	Clasificadores						Promedio
		CAP		CAPIS		CAPIA		
		μ_1	σ_1	μ_2	σ_2	μ_3	σ_3	
1	Breast	97.15±0.02	0.03	93.25±0.04	0.06	86.23±0.3	0.49	92.21
2	Diabetes	61.78±0.08	0.13	66.91±0.13	0.22	59.09±0.33	0.53	62.59
3	Haberman	66.60±0.49	0.80	59.04±0.52	0.83	63.10±0.78	1.25	62.91
4	Heart	63.54±0.11	0.18	81.75±0.19	0.31	73.08±0.32	0.52	72.79
5	Hepatitis	54.07±0.55	0.88	80.01±0.3	0.48	78.67±0.59	0.94	70.92
6	Ionosphere	68.86±0.08	0.12	80.13±0.24	0.39	77.42±0.34	0.55	75.47
7	KingRook	82.62±0.02	0.04	66.02±0.17	0.27	63.52±0.14	0.22	70.72
8	Liver	54.62±0.16	0.26	51.89±0.18	0.29	51.83±0.31	0.50	52.78
9	Sonar	67.94±0.78	1.25	71.54±0.65	1.05	65.22±0.44	0.71	68.23
	Promedio	68.58		72.28		68.68		

Tabla 3. Prueba de ranqueo de signos con suma de Wilcoxon para el CAP y el CAPIS.

No.	Base de datos	CAP	CAPIS	CAP-CAPIS	CAP-CAPIS	Rango	W^-	W^+
1	Breast	97.15	93.25	3.90	3.90	3	0	3
2	Diabetes	61.78	66.91	-5.13	5.13	4	4	0
3	Haberman	66.60	59.04	7.56	7.56	5	0	5
4	Heart	63.54	81.75	-18.21	18.21	8	8	0
5	Hepatitis	54.07	80.01	-25.94	25.94	9	9	0
6	Ionosphere	68.86	80.13	-11.27	11.27	6	6	0
7	KingRook	82.62	66.02	16.60	16.60	7	0	7
8	Liver	54.62	51.89	2.74	2.74	1	0	1
9	Sonar	67.94	71.54	-3.60	3.60	2	2	0
						Suma	29	16

Tabla 4. Prueba de ranqueo de signos con suma de Wilcoxon para el CAP y el CAPIA.

No.	Base de datos	CAP	CAPIA	CAP-CAPIA	CAP-CAPIA	Rango	W^-	W^+
1	Breast	97.15	86.23	10.92	10.92	7	0	7
2	Diabetes	61.78	59.09	2.69	2.69	1	0	1
3	Haberman	66.60	63.10	3.51	3.51	4	0	4
4	Heart	63.54	73.08	-9.54	9.54	6	6	0
5	Hepatitis	54.07	78.67	-24.60	24.60	9	9	0
6	Ionosphere	68.86	77.42	-8.56	8.56	5	5	0
7	KingRook	82.62	63.52	19.10	19.10	8	0	8
8	Liver	54.62	51.83	2.79	2.79	3	0	3
9	Sonar	67.94	65.22	2.72	2.72	2	0	2
						Suma	20	25

Tabla 5. Prueba de ranqueo de signos con suma de Wilcoxon para el CAPIS y el CAPIA.

No.	Base de datos	CAPIS	CAPIA	CAPIS-CAPIA	CAPIS-CAPIA	Rango	W^-	W^+
1	Breast	93.25	86.23	7.02	7.02	7	0	7
2	Diabetes	66.91	59.09	7.83	7.83	8	0	8
3	Haberman	59.04	63.10	-4.06	4.06	5	5	0
4	Heart	81.75	73.08	8.67	8.67	9	0	9
5	Hepatitis	80.01	78.67	1.34	1.34	2	0	2
6	Ionosphere	80.13	77.42	2.70	2.70	4	0	4
7	KingRook	66.02	63.52	2.50	2.50	3	0	3
8	Liver	51.89	51.83	0.05	0.05	1	0	1
9	Sonar	71.54	65.22	6.32	6.32	6	0	6
Suma							5	40

Tabla 6. Iteraciones promedio del CAPIS y CAPIA al recuperar un patrón dado.

No.	Base de datos	CAPIS		CAPIA	
		μ	σ	μ	σ
1	Breast	2.00±0.00	0.00	2.61±0.00	0.01
2	Diabetes	1.83±0.00	0.00	2.46±0.01	0.01
3	Haberman	2.00±0.00	0.00	2.62±0.01	0.02
4	Heart	1.99±0.00	0.00	2.86±0.01	0.02
5	Hepatitis	2.00±0.00	0.00	2.91±0.04	0.06
6	Ionosphere	2.00±0.00	0.00	3.10±0.01	0.02
7	KingRook	1.90±0.00	0.00	3.00±0.01	0.01
8	Liver	1.63±0.01	0.01	2.11±0.02	0.03
9	Sonar	1.94±0.00	0.00	3.04±0.02	0.03
Promedio		1.92		2.75	

5 Conclusiones.

El CAPI funciona de manera autónoma y automática. Trabaja en dos fases: de construcción y de recuperación. En la fase de construcción se obtienen los vectores centroides de cada clase y se bipolarizan mediante el método propuesto de discretización de valores continuos. Posteriormente estos centroides son utilizados para construir la memoria del clasificador. En la fase de recuperación, el patrón presentado a la memoria, previamente construida, también se bipolariza. Esta fase funciona en modos de actualización síncrono (CAPIS) y asíncrono (CAPIA). El CAPI es entrenado con los centroides bipolarizados de cada clase de manera supervisada. Es decir, el mapeo que se hace entre el espacio de representación y el espacio de interpretación, se hace de forma inductiva.

El CAP por ser un sistema de recuperación estático realiza una recuperación en una sola iteración. El CAPI, en ambos modos de recuperación, permite obtener una solución en más de una iteración. Para el CAPIS es de 1.92 y para el CAPIA de 2.75 iteraciones promedio.

Las evaluaciones indican que el CAPIS presenta mejor desempeño sobre el CAP y que el CAP sobre el CAPIA. Por lo tanto, el proceso de recuperación iterativo es capaz de incrementar el rendimiento de clasificación.

Como trabajo futuro se sugiere 1) Probar con algún otro método de discretización de valores continuos, 2) Probar la construcción de la memoria del CAPI con el total de patrones de entrenamiento, no solamente centroides. Implementar una extensión al CAPI para procesos multiclase, 3) Evaluar el CAPI con un número mayor de bases de datos, 4) Comparar el CAPI con otros algoritmos clasificadores, entre otros.

Reconocimientos

S. Valadez agradece al CONACYT por la beca recibida, a la DGEST por el apoyo recibido (4574.12-P), al Insitituto Tecnológico de León y al CIC-IPN para realizar los estudios de posgrado. R. Santiago agradece a la DGEST, y al Insitituto Tecnológico de León. H. Sossa agradece al CONACYT y a la SIP-IPN por el apoyo económico en el marco de los proyectos 155014 y 20121311. Agradece también al ICYTDF por el apoyo brindado en el marco del proeycto: Control inteligente de robots humanoides y robots aéreos y aplicaciones.

Referencias

1. J. P. Marqués de Sá, *Pattern Recognition, Concepts, Methods, and Applications*, Springer-Verlag, 2002.
2. A. R. Webb, *Statistical Pattern Recognition*, 2nd Edition, 2 ed., {John Wiley & Sons}, 2002.
3. T. M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.
4. K. Fukunaga, *Introduction to statistical pattern recognition*, 2 ed., Academic Press Professional, Inc., 1990
5. E. Alpaydin, *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*, 2 ed., The MIT Press, 2010.
6. M. Friedman and A. Kandel, *Introduction to Pattern Recognition : Statistical, Structural, Neural and Fuzzy Logic Approaches (Series in Machine Perception and Artificial Intelligence)*, World Scientific Publishing Company, 1999.
7. J. M. Zurada, *Introduction to Artificial Neural Systems*, West Publishing Company, 1992.
8. G. Palm, F. Schwenker, F. T. Sommer and A. Strey, "Neural Associative Memories," *Biological Cybernetics*, vol. 36, pp. 36-19, 1993
9. B. Prasad, P. E. S, Y. Sagar and P. S. Rama, "A Study on Associative Neural Memories," *International Journal of Advanced Computer Sciences and Applications*, vol. 1, no. 6, pp. 124-133, 2011.
10. T. Kohonen, "Correlation Matrix Memories," *IEEE Transactions on Computers*, vol. 21, pp. 353-359, 1972.
11. R. Santiago, C. Yañez and J. L. Diaz-de-Leon, "Clasificador hibrido de patrones basado en la Lernmarix de Steinbuch y el Linear Associator de Anderson-Kohonen," *Research on Computing Science. Reconocimiento de patrones, avances y perspectivas*. pp. 449-460, 2002.
12. R. Santiago-Montero, "Clasificador hibrido de patrones basado en la Lernmatrix de Steinbuch y el Linear Associator de Anderson-Kohonen," 2003.
13. R. Santiago-Montero, C. Yañez-Marquez and J. L. Diaz-de-Leon, "Clasificador Asociativo de Patrones: Avances Teoricos," *Laboratorio de reconocimiento de patrones. Centro de Investigacion en Computacion, IPN, Mexico*, 2002.
14. J. A. Anderson, "A simple neural network generating an interactive memory," *Mathematical Biosciences*, vol. 14, no. 3-4, pp. 197-220, 1972.
15. K. Nakano, "Associatron-A model for associative memory," *IEEE Transactions on Systems, Man and Cybernetics*, Vols. SMC-2,3, pp. 380-388, 1972.
16. K. Steinbuch, "Die Lernmatrix," *Kybernetic*, vol. 1, pp. 36-45, 1961.
17. K. Steinbuch and H. Frank, "Nichtdigitale Lernmatrizen als Perzeptoren," *Kybernetics*, vol. 1, no. 3, pp. 117-124, 1963
18. M. H. Hassoun, *Fundamentals of Artificial Neural Networks*, 1st ed., MIT Press, 1995.

19. E. Rosenfeld and J. A. Anderson, *Neurocomputing : foundations of research* / edited by James A. Anderson and Edward Rosenfeld, Cambridge, Mass : MIT Press, 1988.
20. M. H. Hassoun, *Associative Neural Memories: Theory and Implementation*, M. H. Hassoun, Ed., New York: Oxford University Press, 1993
21. J. A. Soria-Alcaraz, R. Santiago-Montero and M. Carpio, "One criterion for the selection of the cardinality of learning set used by the Associative Pattern Classifier," *Proceedings of the 2010 IEEE Electronics, Robotics and Automotive Mechanics Conference*, pp. 80-84, 2010
22. B. Kosko, "Adaptive bidirectional associative memories," *Appl. Opt.*, vol. 26, no. 23, pp. 4947-4960, 1987
23. B. Kosko, "Bidirectional Associative Memories," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 18, no. 1, pp. 49-60, 1988.
24. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554-2558, 1982
25. R. Duda, P. Hart and D. Stork, *Pattern classification*, 2nd Edition, Wiley Interscience, New York, 2001
26. M. E. Acevedo, M. A. Acevedo and F. Felipe, "Classification of Cancer Recurrence with Alpha-Beta BAM," *Mathematical Problems in Engineering*, p. 14, 2009.
27. C. Yañez-Márquez, "Memorias asociativas basadas en relaciones de orden y operaciones binarias," *Computación y Sistemas. CIC, IPN.*, vol. 6, no. 4, pp. 300-311, 2003.
28. Z. Zhang, H. Zhuo, S. Liu and P. d. B., "Classification of cancer patients based on elemental contents of serums using bidirectional associative memory networks," *Analytica Chimica Acta*, vol. 436, pp. 281-291, 2001.
29. F. J. López-Aligué, I. Acevedo-Sotocal, I. Álvarez-Troncoso, C. García-Orellana and H. González-Velasco, "A Neural Associative Pattern Classifier," *Advances in Artificial Intelligence - IBERAMIA 2002*, 8th Ibero-American Conference on AI, pp. 430-439, 2002.
30. A. Frank and A. Asuncion, {UCI} Machine Learning Repository, 2010

Análisis de preguntas de cuestionarios con Redes Neuronales

José de Jesús Zamarripa-Topete¹, Luz Evelia López-Chico¹ y Víctor Rafael Velázquez-Mejía¹

¹Universidad Autónoma de Baja California
Ensenada, B.C., México

jesuszamarripa@uabc.edu.mx, evelia@uabc.edu.mx, vvmejia@uabc.edu.mx

Paper received on 23/07/12, Accepted on 12/09/12.

Resumen. Se relacionan áreas conocimiento: 1. Diseño y Evaluación de Proyectos con cuestionario de Estudio Factibilidad; 2. Confiabilidad cuestionarios con alfa Cronbach; 3. Redes Neuronales para determinar preguntas a reelaborar. Alumnos en materias: Diseño y Evaluación Proyectos Electrónicos y Emprendedores, realizan cuestionario, pilotean y miden confiabilidad con alfa Cronbach, problema los menores de alfa 0.7. La Solución se integra: Módulo Estadístico y diez módulos Redes Neuronales. Estructura Redes Neuronales: capa entrada 10 neuronas, tres capas ocultas 50 neuronas, capa salida una neurona. Analizaron cuestionarios y Redes Neuronales establecieron preguntas con problemas, alumnos reelaboraron y en segundo piloteo lograron alfa mayor 0.7

Palabras Clave: Redes Neuronales, alfa de Cronbach.

1 Introducción

En este trabajo convergen tres áreas de conocimiento: 1. Diseño y Evaluación de Proyectos, en la Etapa de Inicio con el Estudio de Factibilidad del Proyecto y la elaboración de cuestionarios para determinar la opinión de los futuros usuarios del Proyecto; 2. El análisis estadístico de la confiabilidad de los cuestionarios con el alfa de Cronbach; 3. Las Redes Neuronales para establecer las preguntas a reelaborar en los cuestionarios que no alcanzan el alfa de Cronbach establecida.

La aportación principal del presente trabajo es la aplicación de las Redes Neuronales al analizar la información de la confiabilidad de los cuestionarios y establecer las preguntas que se deben hacer nuevamente, para mejorar el cuestionario y obtener un valor de alfa de Cronbach aceptable. La aportación secundaria es la propuesta de aplicar el alfa de Cronbach en el Diseño y Evaluación de Proyectos.

Esta aplicación se utilizó para apoyar a los estudiantes de Ingeniería en Electrónica del 9no semestre en las materias de Diseño y Evaluación de Proyectos Electrónicos y Emprendedores.

1.1 Diseño y Evaluación de Proyectos

Los alumnos proponen un Proyecto a desarrollar durante el ciclo escolar. Los Proyectos son: 1. Administrador inalámbrico de aparatos domésticos de bajo costo; 2. Tablero didáctico que determina el tipo de acentuación de palabras; 3. Cargador solar de celulares; 4. Control para un sistema doméstico de reciclaje de agua; 5. Llenador automático de garrafones de agua potable; 6. Casa inteligente.

En el Diseño y Evaluación de Proyectos, en su primera etapa se incluye el Estudio de Factibilidad para determinar la viabilidad del Proyecto. Las partes básicas del Estudio de Factibilidad son: 1. Análisis Técnico Operativo; 2. Análisis Económico Financiero; 3. Análisis de Normatividad Asociada; 5. Análisis de Mercado; 4. Análisis Social [1].

Parte del Análisis Social es la opinión del usuario con respecto al Proyecto: 1. Rechazo; 2. Poco Rechazo; 3. Indiferencia; 4. Poca Aceptación; 5. Aceptación.

Esto se realiza directamente con el usuario (fuente primaria), se emplean entrevista y cuestionarios, de diversos tipos [1].

Acotamiento del problema: 1. Los cuestionarios los realizan los alumnos. 2. La finalidad no es elaborar cuestionarios validados. 3. El tiempo disponible para realizar el cuestionario es de solo un par de semanas en el semestre escolar, por las fases de las materias que son: a) elegir el proyecto; b) determinar su factibilidad; c) diseñarlo; d) construirlo; e) probarlo; f) documentar el proceso; g) realizar una exposición del proyecto en funcionamiento [2].

1.2 Confiabilidad de cuestionarios con alfa de Cronbach

El cuestionario es el medio con el que se mide determinado comportamiento o conducta. La confiabilidad de un cuestionario es el grado en que sus resultados son consistentes y coherentes con el atributo a medir [6].

El índice de confiabilidad es el alfa de Cronbach que mide la consistencia interna del cuestionario. Sus valores son de 0 a 1, los valores mínimos de aceptación son 0.6 a 0.7, los valores entre 0.9 a 1 implican que existen redundancias [4]. Para este trabajo los valores de aceptación son de 0.7 a 0.9.

El alfa de Cronbach se emplea ampliamente en la Psicología, Ciencias Humanas, Ciencias de la Educación, etc. Pero muy poco en Diseño y Evaluación de Proyectos.

En este trabajo se calculará el alfa de Cronbach de los cuestionarios elaborados por los alumnos de sus respectivos Proyectos.

1.3 Redes Neuronales

La neurona es la representación en programación de su contraparte biológica, con respecto a sus entradas, procesamiento y activación [3].

La neurona presentada en la Figura 1 (*neurona_i*), se constituye de las entradas (x_1, x_2, \dots, x_j), los pesos ($w_{i1}, w_{i2}, \dots, w_{ij}$) que se calculan con el entrenamiento, las entradas se multiplican por su respectivo peso ($x_1w_{i1}, x_2w_{i2}, \dots, x_jw_{ij}$), en el cuerpo de la neurona se suman las multiplicaciones ($\sum_j x_jw_{ij}$), la salida (y_i) es el resultado de la

función de activación, para este estudio es la función Sigmoidea ($1/[1+e^{-x}]$). Las Redes Neuronales son arreglos de neuronas interconectadas, que se acomodan por capas: capa de entrada, una o varias capas ocultas y una capa de salida. Se utilizan las redes neuronales para identificar patrones o características en la información [3].

En el presente trabajo se emplearán las Redes Neuronales para analizar la información de la confiabilidad de los cuestionarios elaborados por los estudiantes.

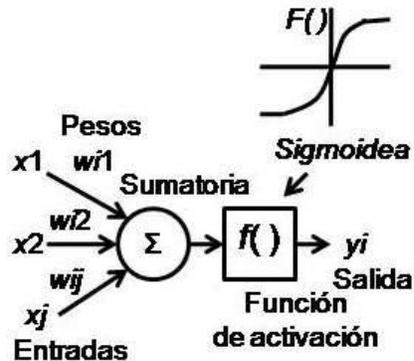


Figura 1. Neurona.

2. Desarrollo

2.1 Elaboración del cuestionario

Son seis equipos de alumnos que realizaron cada uno un cuestionario de su respectivo Proyecto. Por sus características los Proyectos pueden ser: 1. Mejorar un producto o servicio, que actualmente se ofrece; 2. Es nuevo y nadie lo ofrece [1].

Se define el perfil de los potenciales usuarios de los Proyectos:

1. Proyecto: Administrador inalámbrico de aparatos domésticos de bajo costo; Usuarios: con dispositivos WiFi.
2. Proyecto: Tablero didáctico que determina el tipo de acentuación de palabras; Usuarios: maestros de educación básica.
3. Proyecto: Cargador solar de celulares; Usuarios: con teléfono celular;
4. Proyecto: Control para un sistema doméstico de reciclaje de agua; Usuarios: con problema de suministro de agua.
5. Proyecto: Llenador automático de garrafones de agua potable; Usuarios: tiendas de autoservicio que vendan agua.
6. Proyecto: Casa inteligente; Usuarios: constructores o personas que deseen instrumentar viviendas.

A los alumnos se les entrega una plantilla para el desarrollo de preguntas del cuestionario (Tabla 1). Los alumnos elaboran diez preguntas con respuesta dicotómica de "sí" o "no", que cubren (en su caso) los indicadores, ejemplo de preguntas: ¿En el último año ha comprado un Smartphone?, ¿Compraría un cargador solar para su Smartphone? De esta forma se genera la versión 1.0 de su cuestionario.

Tabla 1. Plantilla para el desarrollo de preguntas del cuestionario.

Dimensión:	Indicadores:	Actividad:
Problema.	- Identificación del problema. - Importancia del problema.	Elaborar diez preguntas dicotómicas que cubran los indicadores de su Proyecto.
Uso de tecnología.	- Aceptación de la tecnología. - Uso de la tecnología.	
Gasto en tecnología.	- Gasto en el pasado en tecnología. - Intención de inversión en tecnología.	
Proveedor de producto o servicio (en su caso).	- Satisfacción con el proveedor actual. - Disposición a cambiar de proveedor.	
Aceptación del Proyecto.	- Intención de adquisición del Proyecto. - Intención de uso del Proyecto.	

2.2 Prueba Piloto de los cuestionarios versión 1.0

En el Diseño y Evaluación de Proyectos, el muestreo probabilístico para la aplicación de cuestionarios es poco usado [1]. En este trabajo se emplea el Muestreo de Conveniencia de Sitio, el cual reúne a los usuarios que cumplen con el perfil predefinido. Para algunos Proyectos se eligen tiendas que ofrezcan los últimos modelos tecnológicos relacionados al Proyecto.

Al ser una actividad académica y por los tiempos escolares (solo dos semanas), se proponen a diez personas por Proyecto para aplicarles la prueba piloto. Con respecto a la representatividad y evitar el sesgo, se solicita a los alumnos que cada persona elegida sea de diferente fraccionamiento, escuela o zona de la ciudad.

2.3 Codificación por cuestionario

Las personas se codificaron como "Persona1" a "Persona10" y las preguntas de "X1" a "X10". A las respuestas se les asigna para el "sí" valor de "1" y para "no" el "0".

El ejemplo de un cuestionario del proyecto 4 (control para un sistema doméstico de reciclaje de agua) versión 1.0 contestado se vacía en el formato de la Tabla 2.

Tabla 2. Formato Cuestionario de Proyecto 4 contestado.

Personas:	Preguntas:									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Persona1	1	0	0	0	0	1	0	0	1	0
Persona2	1	1	0	1	1	1	1	0	0	0
Persona3	0	0	0	0	0	0	0	0	1	0
Persona4	1	1	1	1	1	0	0	1	1	0
Persona5	0	0	0	0	0	0	1	1	1	1
Persona6	1	1	1	0	1	1	1	1	1	1
Persona7	0	0	1	1	1	0	0	0	0	1
Persona8	0	0	1	1	1	1	1	1	1	1
Persona9	0	1	0	1	1	0	0	0	0	1
Persona10	1	1	0	0	1	1	1	1	1	1

2.4 Confiabilidad del cuestionario

Para determinar la confiabilidad del cuestionario se emplea el alfa de Cronbach (α) [5]:

Paso 1. Se obtiene la correlación entre las preguntas (r_{ij}) con la Ecuación 1 (r de Pearson). Requiere: la pregunta número i (X_i); la pregunta número j (X_j); el número de preguntas (n).

$$r_{ij} = \frac{\sum X_i X_j - \frac{(\sum X_i)(\sum X_j)}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n} \right] \left[\sum X_j^2 - \frac{(\sum X_j)^2}{n} \right]}} \quad (1)$$

Paso 2. Calcular la correlación promedio (\bar{r}) con la Ecuación 2. Los términos son: la correlación entre las preguntas (r_{ij}); el número de correlaciones (N).

$$\bar{r} = \frac{\sum r_{ij}}{N} \quad (2)$$

Paso 3. Obtener el alfa de Cronbach (α) de la Ecuación 3. Los elementos son: número de preguntas (n); correlación promedio (\bar{r}).

$$\alpha = \frac{n(\bar{r})}{1 + \bar{r}(n-1)} \quad (3)$$

Los cuestionarios de los Proyectos obtuvieron los valores de alfa presentados en la Tabla 3. Para determinar que un cuestionario es confiable el valor de alfa debe ser entre 0.7 a 0.9, solo el "Cuestionario Proyecto 2" cumple (0.714) y puede ser aplicado a su público meta. El resto de los cuestionarios no cumplen y deben ser reelaborados [4]. El siguiente paso es determinar las preguntas a corregir.

Tabla 3. Alfa de los cuestionarios.

	Alfa de Cronbach (α):	$0.7 < \alpha < 0.9$
Cuestionario Proyecto 1	0.634	No
Cuestionario Proyecto 2	0.714	Si
Cuestionario Proyecto 3	0.230	No
Cuestionario Proyecto 4	0.661	No
Cuestionario Proyecto 5	0.595	No
Cuestionario Proyecto 6	0.290	No

2.5 Establecer las preguntas con problemas

Para el cuestionario en su versión 1.0 pilotado que no cumple con el alfa de Cronbach, se utiliza la Ecuación 1 para obtener la matriz de correlación inter-elemento, en la Tabla 4 se presenta el Cuestionario Proyecto 4. Esta matriz se procesa con redes neuronales para establecer las preguntas que tienen problemas.

Tabla 4. Matriz de correlación inter-elemento " r_{ij} " Cuestionario Proyecto 4.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
X1	r_{11} 1.00	0.60	0.00	-0.20	r_{15} 0.22	0.60	0.20	0.20	0.22	r_{110} -0.41
X2	0.60	1.00	0.00	0.20	0.66	0.20	0.20	0.20	-0.22	0.00
X3	0.00	0.00	1.00	0.41	0.54	0.00	0.00	0.41	0.09	0.25
X4	-0.20	0.20	0.41	1.00	0.66	-0.20	-0.20	-0.20	-0.66	0.00
X5	r_{51} 0.22	0.66	0.54	0.66	r_{55} 1.00	0.22	0.22	0.22	-0.43	r_{510} 0.37
X6	0.60	0.20	0.00	-0.20	0.22	1.00	0.60	0.20	0.22	0.00
X7	0.20	0.20	0.00	-0.20	0.22	0.60	1.00	0.60	0.22	0.41
X8	0.20	0.20	0.41	-0.20	0.22	0.20	0.60	1.00	0.66	0.41
X9	0.22	-0.22	0.09	-0.66	-0.43	0.22	0.22	0.66	1.00	-0.09
X10	r_{101} -0.41	0.00	0.25	0.00	r_{105} 0.36	0.00	0.41	0.41	-0.09	r_{1010} 1.00

2.6 Integración de la solución

La solución presentada en la Figura 2, se integra de un Módulo Estadístico y diez Redes Neuronales.

La información de la Tabla 2 entra en el Módulo Estadístico y este entrega la matriz de correlación inter-elemento de la Tabla 4, las correlaciones son de -1 a 1, los valores negativos presentan las inconsistencias de la pregunta con respecto al cuestionario y los positivos la consistencia de la pregunta con el cuestionario.

Son diez Redes Neuronales por facilidad de su entrenamiento [8], cada Red Neuronal tiene la siguiente estructura (Figura 3): diez neuronas en la capa de entrada (1 a 10); 30 neuronas en la primera capa oculta (11 a 40); diez neuronas en la segunda capa oculta (41 a 50); 11 neuronas en la tercera capa oculta (51 a 61); una neurona en la capa de salida (62).

La información que entra a cada red neuronal corresponde a diez elementos de la matriz de correlación inter-elemento: $r_{11} - r_{110}$; $r_{21} - r_{210}$; ...; $r_{91} - r_{910}$; $r_{101} - r_{1010}$. La salida de la Red Neuronal ($OutX1 \dots OutX10$) entrega el grado de problema (inconsistencia de la pregunta con respecto al cuestionario) de la pregunta (X1, X2, ..., X9, X10) en el intervalo de 0.0 (sin problema) a 0.9 (problema grave) [8], con el resultado mayor a 0.3 (0.4 a 0.9) se debe reelaborar la pregunta. La solución se programó en C++ [7].

2.7 Entrenamiento de las Redes Neuronales

Las Redes Neuronales se entrenaron con aprendizaje supervisado de retropropagación, tolerancia 0.005, término de momentum 0.95 y razón de aprendizaje 0.5 [7].

Se realiza un entrenamiento segmentado [8]. El entrenamiento segmento 1: neurona 1 (capa entrada) y las neuronas 11 a 13 (capa oculta 1). Se tiene una entrada con 201 posibilidades (-1.0, -0.99...-0.01, 0, 0.01...0.99, 1.0) y 2 salidas (1, 0). Por lo tanto se tienen 201 vectores de entrenamiento. Se repite con las restantes (2 a 40).

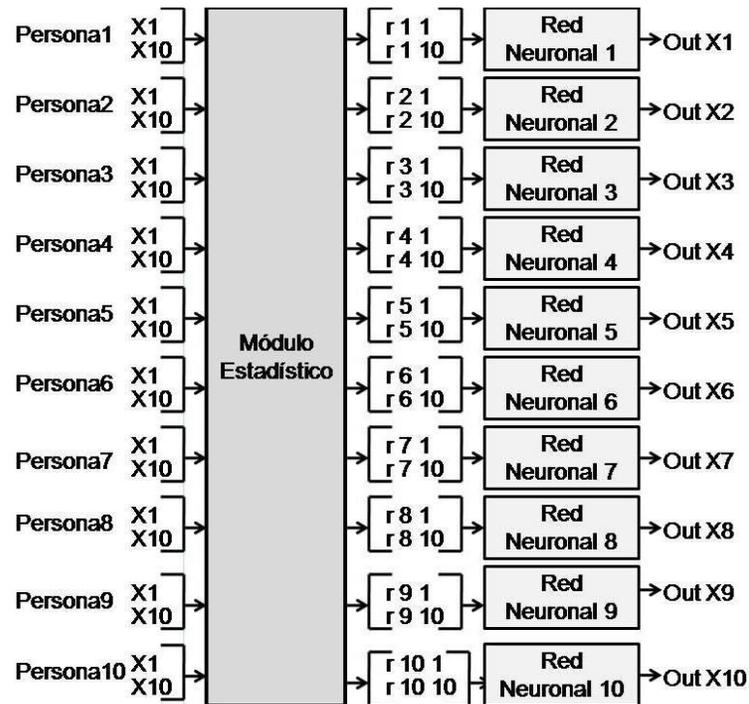


Figura 2. Integración de la solución.

El entrenamiento segmento 2: neuronas 41 a 50 (capa oculta 2), las neuronas 51 a 61 (capa oculta 3) y la neurona 62 (capa salida). Este segmento en cada entrada puede tener dos posibles valores (0 o 1) y la neurona de salida diez valores (0 a 0.9). Los vectores de entrenamiento requerido son: entradas (10) elevadas a las posibilidades (2), $10^2=100$ vectores de entrenamiento.

Se integran las Redes Neuronales entrenadas de forma segmentada, se prueban con vectores de entrada y salida conocidos para validar su entrenamiento.

3 Resultados

Con varias Redes Neuronales y su entrenamiento segmentado, el aprendizaje de las mismas se alcanzó en poco tiempo. Las Redes Neuronales con muchos elemen-

tos o con un número grande de vectores de entrenamiento, su aprendizaje puede demorar varios días o no ocurrir.

Los Cuestionarios de Proyecto que no cumplieron con el alfa de Cronbach (Tabla 3), se procesaron y se obtuvo la información de la Tabla 5. De los resultados se recomienda a los estudiantes que en su respectivo cuestionario reelaboren las siguientes preguntas con valores mayores de 0.3 (0.4 a 0.9): *Cuestionario Proyecto 1* la pregunta X4; *Cuestionario Proyecto 3* las preguntas X2, X3, X7; *Cuestionario Proyecto 4* la pregunta X4, X9; *Cuestionario Proyecto 5* las preguntas X4, X5; *Cuestionario Proyecto 6* las preguntas X2, X3, X6, X9, X10.

Los alumnos corrigieron las preguntas indicadas para generar la versión 2.0 de su cuestionario y realizaron una segunda prueba piloto con otros usuarios. Los cuestionarios contestados del segundo pilotaje se les calculo el alfa de Cronbach (Tabla 6), la totalidad cumplen con el valor de alfa y proceden a aplicarlos a sus respectivos públicos meta. Al utilizar la solución (programa) se sistematiza el análisis a los cuestionarios con problemas y permite a los alumnos trabajar en el tiempo asignado (dos semanas).

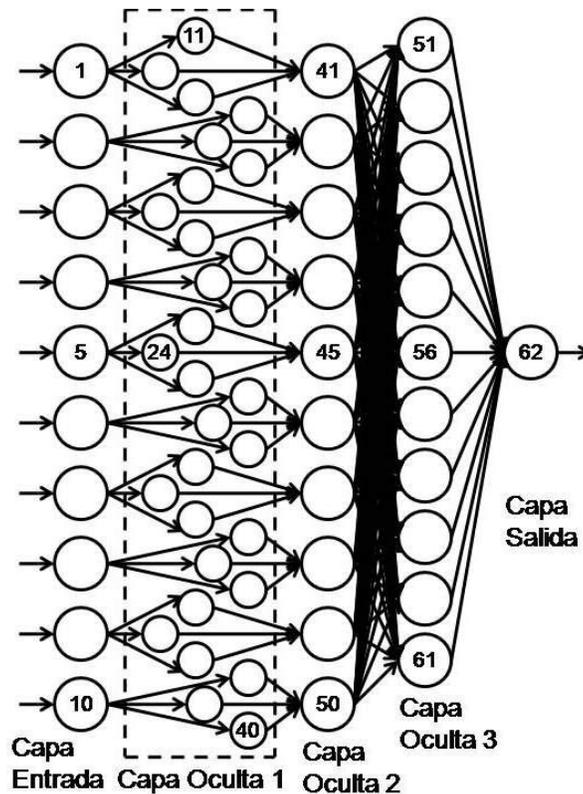


Figura. 3. Red Neuronal.

Tabla 5. Grado de problema de las preguntas.

	Pregunta:									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Cuestionario Proyecto 1	0.2	0.2	0.1	0.4	0.2	0.0	0.0	0.0	0.3	0.2
Cuestionario Proyecto 3	0.2	0.7	0.6	0.3	0.3	0.1	0.6	0.3	0.2	0.3
Cuestionario Proyecto 4	0.2	0.1	0.0	0.5	0.1	0.1	0.1	0.1	0.4	0.2
Cuestionario Proyecto 5	0.2	0.0	0.1	0.6	0.5	0.2	0.1	0.2	0.2	0.1
Cuestionario Proyecto 6	0.3	0.4	0.5	0.1	0.3	0.5	0.1	0.1	0.6	0.5

Tabla 6. Alfa de Cronbach segundo piloteo.

	Alfa de Cronbach (α):	$0.7 < \alpha < 0.9$
Cuestionario Proyecto 1	0.722	Si
Cuestionario Proyecto 3	0.719	Si
Cuestionario Proyecto 4	0.751	Si
Cuestionario Proyecto 5	0.733	Si
Cuestionario Proyecto 6	0.729	Si

4 Conclusiones

Al tener la solución implementada se sistematiza el análisis de los cuestionarios con problemas. Con las Redes Neuronales se puede establecer cuáles son las preguntas que deben ser reelaboradas. Con esto se apoya a los alumnos para que realicen mejores cuestionarios con respecto a sus Proyectos y así su Estudio de Factibilidad se fortalece. De esta forma los alumnos cumplen en esta parte del desarrollo de proyectos en tiempo y forma. Al ser una solución en programa se puede instalar en un servidor y analizar los cuestionarios recibidos por correos electrónicos o contestados en línea. Al trabajar la solución con varias Redes Neuronales (diez) y estas con varias capas ocultas (tres) se facilitó el proceso de entrenamiento segmentado de las mismas. Las Redes Neuronales pueden ser una herramienta de apoyo para analizar información en las diversas fases del desarrollo de los Proyectos.

Referencias

1. Baca, G.: Evaluación de proyectos. 6ª Edición. Mc Graw Hill. ISBN 978-607-15-0260-5, (2010)

2. Cleand, D., King, W.: Manual para la administración de proyectos. Grupo Editorial Patria. ISBN 968-26-1151-2, (2007)
3. Martín del Brío, B., Sanz A.: Redes neuronales y sistemas borrosos. 3ª Edición. Alfaomega Ra-Ma. ISBN: 978-970-15-1250-0, (2007)
4. Oviedo, H. C., Campos, A.: Aproximación al uso del coeficiente alfa de Cronbach. Revista colombiana de psiquiatría. Vol XXXIV, número 004. ISSN 0034-7450, pp. 572-580, (2005)
5. Pagano, R.: Estadística para las ciencias del comportamiento. 7ª Edición. Thomson, ISBN 0-534-61767-0, (2006)
6. Pérez, H. E.: Estadística para las Ciencias Sociales, del Compostamiento y de la Salud. Cengage Learning. ISBN 978-970-686-491-1, (2007)
7. Rogers, J.: Object-Oriented Neural Networks in C++. Academic Press. ISBN 0125931158, (1997)
8. Sharkey, A. J.: Improving the expert networks of a modular multi-net system for pattern recognition. Springer. ISBN 978-3-540-38625-4, (2006)
9. Van Der Zwaag, B. J., Slump, C. H., Spaanenburg, L.: Process Identification Through Modular Neural Networks and Rule Extraction. World Scientific. ISBN 981-238-066-3, (2002)

Heurísticas para resolver el problema de la Supersecuencia Común más Corta

Marcela Rivera-Martínez¹ Luis-René Marcial-Castillo²
Lourdes Sandoval-Solís³

Benemérita Universidad Autónoma de Puebla, Puebla, Pue., México
¹mrvimar@hotmail.com, ²lmc@cs.buap.mx, ³sandoval@cs.buap.mx
Paper received on 29/06/12, Accepted on 08/09/12.

Resumen. El problema de la Supersecuencia Común más Corta, es un problema NP y tiene aplicaciones en compresión de datos, ingeniería mecánica, biología molecular y planeación entre otros. Para resolver el problema de la Supersecuencia Común más Corta, en este trabajo, se describen algunas heurísticas, las cuales están basadas en la heurística clásica Majority Merge, sin embargo, difieren en el criterio que se sigue para elegir el símbolo que formará parte de la Supersecuencia. Se realizaron varios experimentos para comparar los resultados de las heurísticas con Majority Merge.

Palabras Clave: Supersecuencia Común más Corta, Heurísticas, Majority Merge.

1 Introducción.

El problema de la Supersecuencia Común más Corta (Shortest Common Supersequence, SCS) está dado de la siguiente forma: Dado un lenguaje L sobre un alfabeto Σ , se desea encontrar una de las cadenas w de menor longitud que sea una Supersecuencia de L . Se dice que S es una Supersecuencia de una cadena T , si S se puede obtener desde T mediante la inserción de cero o más símbolos de la misma cadena. Ejemplo: Dado $\Sigma = \{a, b\}$ y $L = \{ab, ba\}$. Las siguientes cadenas son supersecuencias de L : “abba”, “aabba”, “aba”, “bab”. No son las únicas, existen muchas más, se dice que todas ellas son supersecuencias porque cada una de ellas contiene a todas las cadenas de L (aún cuando no las contengan de manera contigua). Ahora, las supersecuencias comunes más cortas son dos y son: $S = aba$ y $S = bab$. Ambas son supersecuencias de L y son las más cortas porque ambas son de longitud 3, se denota: $|S| = 3$, es decir, la supersecuencia puede no ser única.

El problema de la supersecuencia común más corta es un problema NP-duro bajo ciertas restricciones concernientes al tamaño del alfabeto y la longitud de las cadenas [6, 9]. El problema de la SCS tiene aplicaciones en compresión de datos, ingeniería mecánica, planeación y biología molecular entre otros [8].

Existen diferentes formas de resolver el problema de la SCS, entre ellas tenemos:

Algoritmos secuencia por secuencia. Este tipo de algoritmos construyen la supersecuencia tomando pares de cadenas, la forma de elegir esos pares de cadenas varía dependiendo del algoritmo, entre ellos tenemos el algoritmo de programación dinámica [4], el algoritmo de torneo, etc.

Algoritmos de reducción y/o expansión. Inician con alguna supersecuencia común del lenguaje L y van eliminando (y/o expandiendo) dicha supersecuencia, entre ellos están el algoritmo de reducción-expansión [2,7], el algoritmo de Hubbell-Morris-Winkler [5], etc.

Algoritmos con metaheurísticas. La mayoría de ellos hibridan con heurísticas clásicas como Majority Merge (MM) [3], sin embargo, ésta heurística clásica no siempre funciona de manera adecuada, por lo cual se buscan nuevas heurísticas que ayuden a superar los problemas que presenta la heurística MM.

En éste trabajo se proponen heurísticas basadas en Majority Merge, pero que difieren de ésta en el criterio para elegir el nuevo elemento a adicionar a la Supersecuencia, lo cual permitirá encontrar mejores soluciones de la SCS y poder aplicarlas a problemas reales como la hibridación de oligonucleótidos [1, 2].

2 Heurísticas.

Muchos problemas de optimización no pueden ser resueltos utilizando métodos exactos, ya sea, porque tienen un alto grado combinatorio o por la dificultad de generar un modelo basado en programación matemática que represente exactamente una situación real. Para este tipo de problemas, se han venido generando desde la década de los sesenta métodos conocidos como heurísticas, capaces de encontrar soluciones de buena calidad, aproximadas a la solución óptima y en un periodo de tiempo razonable. De acuerdo con el estándar ANSI/IEEE 100-1984, “la heurística trata de métodos o algoritmos exploratorios, durante la solución de problemas, en los cuales las soluciones se descubren por la evaluación del progreso logrado en la búsqueda de un resultado final.”

A continuación se describen 4 heurísticas que difieren en la forma de elegir el símbolo a adicionar a la Supersecuencia.

2.1 Heurística Majority Merge (MM).

La heurística MM es una de las heurísticas clásicas más utilizadas para resolver el problema de la SCS, está clasificada dentro de los algoritmos símbolo por símbolo. Dado un lenguaje L y un alfabeto Σ , el símbolo σ que se agrega es aquel que estando al frente de cada cadena L sea el de mayor frecuencia, y una vez agregado elimina ese símbolo del frente de cada cadena donde se encuentre, éste proceso se termina cuando el lenguaje L contiene únicamente la cadena vacía.

Algoritmo MM.

Entrada: Lenguaje (L) y Alfabeto (Σ).

$S \leftarrow$ cadena vacía.

REPETIR

Calcular $\Sigma' = \{a \in \Sigma \mid a = \max(\text{frec}(a))\}$, donde $\text{frec}(a)$ es la frecuencia con que aparece el símbolo a al inicio de cada cadena en L .

Elegir $\sigma \in \Sigma'$.

Actualizar L , eliminando σ del frente.

$S \leftarrow S\sigma$.

HASTA que L sea la cadena vacía.

Salida: S y su longitud.

Ejemplo 1: Sean $\Sigma = \{a, b, c\}$ y $L = \{acacb, cbaac, bbacc\}$.

La forma en que la heurística Majority Merge va construyendo la supersecuencia es la siguiente: Los caracteres al frente de cada cadena son: a, c y b . Se calcula la frecuencia de cada uno de ellos y se obtiene el máximo: $\max\{1, 1, 1\} = 1$. Por lo cual $\Sigma' = \{a, b, c\}$, de ahí se elige el símbolo a adicionar a la Supersecuencia, por ejemplo a , es decir, $S = a$, entonces, el símbolo a es eliminado del frente de las cadenas de L donde aparece, resultando el nuevo lenguaje $L = \{cacb, cbaac, bbacc\}$. Este proceso se repite hasta que L sea la cadena vacía.

La heurística MM, tiene problemas en encontrar la Supersecuencia de menor longitud en cadenas de longitud aleatoria, debido a que el criterio que utiliza para adicionar un símbolo a la Supersecuencia hace que aquellas cadenas con mayor longitud sean las que vayan quedando en el lenguaje L aumentando de ésta forma la longitud de S .

2.2 Heurística H1.

Debido a la problemática que presenta la heurística MM, se asignan pesos a los caracteres de cada cadena, lo cual esta dado por la longitud de la cadena menos la posición del símbolo dentro de la cadena, es decir, para un caracter a_i de una cadena $A = a_1a_2\dots a_n$, el peso $w(a_i)$ es $n-i$.

Algoritmo H1.

Entrada: Lenguaje (L) y Alfabeto (Σ).

$S \leftarrow$ cadena vacía.

REPETIR

PARA cada $a \in \Sigma$ HACER

Calcular $w(a)$.

FIN_PARA

Calcular $\Sigma' = \{a \in \Sigma \mid wsum(a) = \max\{wsum(a) \text{ para cada } a \in \Sigma\}\}$, donde $wsum(a)$ es la suma de los pesos de las ocurrencias de a en el frente actual de cada cadena de L .

Elegir $\sigma \in \Sigma'$.

Actualizar L , eliminando σ del frente de cada cadena de L donde aparezca σ .

$S \leftarrow S\sigma$.

HASTA que L sea la cadena vacía.

Salida: S y su longitud.

Considerando el Ejemplo 1, veamos la forma de construir la supersecuencia aplicando la heurística H1: primero se calcula el peso para cada símbolo del alfabe-

to, obteniendo: $w(a) = 4$, $w(b) = 4$ y $w(c) = 4$, ahora se calcula la suma de los pesos de cada símbolo del alfabeto y se obtiene: $wsum(a) = 4$, $wsum(b) = 4$ y $wsum(c) = 4$. Se calcula el máximo de la suma de los pesos y se forma Sigma Prima, obteniendo: $\Sigma' = \{a, b, c\}$ y de éste último conjunto, se elige el símbolo a adicionar a la Supersecuencia, digamos el símbolo c , ahora, el símbolo c se elimina de las cadenas de L donde aparezca al frente, obteniendo el nuevo lenguaje $L = \{acacb, baac, bbacc\}$. Este proceso se repite hasta que sólo quede la cadena vacía en el lenguaje L .

2.3 Heurística H2.

Otra heurística diferente, llamada H2, considera no sólo el caracter que está al frente de cada cadena, sino considera además del primero, el segundo caracter y elige el símbolo a cuyo número de ocurrencias al frente de cada cadena más el número de ocurrencias del segundo símbolo más frecuente, sea máximo. Para ello se utiliza la siguiente notación:

$L-a$ está formado por aquellas cadenas del lenguaje L que tienen al frente el caracter a y que se ha eliminado de las cadenas donde aparece como caracter inicial; $occ(a, L-a)$ denota la cantidad de a 's al frente de cada cadena del Lenguaje ($L-a$).

Algoritmo H2.

Entrada: Lenguaje (L) y Alfabeto (Σ).

$S \leftarrow$ cadena vacía.

REPETIR

 PARA cada $a \in \Sigma$ HACER

 Calcular $L-a$.

 Calcular $occ(a, L-a)$.

 FIN_PARA

 PARA cada $a \in \Sigma$ HACER

 Calcular $r_a = \max\{occ(a, L-a) \text{ para cada } a \in \Sigma\}$.

 FIN_PARA

 Calcular $\Sigma' = \{a_i \in \Sigma \mid i = \max\{occ(a, L) + r_a\}\}$, donde $occ(a, L)$ es la cantidad de a 's en el frente actual de cada cadena de L .

 Elegir $\sigma \in \Sigma'$.

 Actualizar L eliminando σ del frente de cada cadena de L donde aparezca σ .

$S \leftarrow S\sigma$.

HASTA que L sea la cadena vacía.

Salida: S y su longitud.

La forma que la heurística H2 forma la supersecuencia para el ejemplo 1 es la siguiente: Para cada símbolo del alfabeto se calcula un nuevo lenguaje (denotado $L-a$) y se obtiene: $L-a = \{cacb, cbaac, bbacc\}$, $L-b = \{acacb, cbaac, bacc\}$ y $L-c = \{acacb, baac, bbacc\}$, ahora se calcula $occ(L-a)$ de lo cual se obtiene: $occ(a, L-a) = 0$, $occ(b, L-a) = 1$, $occ(c, L-a) = 2$, $occ(a, L-b) = 1$, $occ(b, L-b) = 1$, $occ(c, L-b) = 1$, $occ(a, L-c) = 1$, $occ(b, L-c) = 2$ y $occ(c, L-c) = 0$, se calcula $r_a = \max\{0, 1, 2\} = 2$, $r_b = \max\{1\} = 1$, $r_c = \max\{0, 1, 2\} = 2$. Ahora se calcula para cada caracter del alfabeto $occ(a, L)$ y se obtiene: $occ(a, L) = 1$, $occ(b, L) = 1$ y $occ(c, L) = 1$, se procede a

calcular $\max\{occ(a, L) + r_a\}$ para cada símbolo a del alfabeto y se obtiene: $\max\{(1+2), (1+1), (1+2)\} = 3$, con ello se forma $\Sigma' = \{a, c\}$, de éste nuevo conjunto se elige el símbolo a adicionar a la Supersecuencia, digamos a , por lo cual se tiene $S = a$ y se elimina el símbolo a de las cadenas de L donde aparezca al frente, obteniendo el nuevo lenguaje $L = \{cacb, cbaac, bbacc\}$. Este proceso continúa hasta que el lenguaje L contenga únicamente la cadena vacía.

2.4 Heurística H3.

La heurística llamada H3, calcula para cada símbolo a al frente de cada cadena de L , la longitud de la supersecuencia mediante el algoritmo H1, al eliminar ese carácter a del frente ($L-a$), después de ésto H3 toma como símbolo para agregar a S aquel que tenga la mínima longitud, dado que generalmente hay empates H3 lo resuelve calculando para cada cadena en L que tenga al frente el carácter a (tal que ese carácter haya sido elegido como mínimo en el paso anterior) la supersecuencia de (L_a) usando H1 y se elige el símbolo cuya longitud devuelta por H1 sea máxima. La notación que se usa es la siguiente:

$L-a$ está formado por aquellas cadenas del lenguaje L que tienen al frente el carácter a y éste carácter se ha eliminado de las cadenas donde aparece como carácter inicial.

L_a está formada por aquellas cadenas de L que tienen al frente el carácter a .

Algoritmo H3.

Entrada: Lenguaje (L) y Alfabeto (Σ).

$S \leftarrow$ cadena vacía.

REPETIR

 PARA cada $a \in \Sigma$ HACER

 Calcular $L-a$.

 FIN_PARA

 Calcular $\Sigma' = \{a \in \Sigma \text{ tal que } a = \min\{|H1(L-a)|\} \text{ para cada } a \in \Sigma\}$.

 PARA cada $a \in \Sigma'$ HACER

 Calcular L_a .

 FIN_PARA

 Calcular $\Sigma^* = \{a \in \Sigma' \text{ tal que } a = \max\{|H1(L_a)|\} \text{ para cada } a \in \Sigma'\}$.

 Elegir $\sigma \in \Sigma^*$.

 Actualizar L eliminando σ del frente de cada cadena de L donde aparezca σ .

$S \leftarrow S\sigma$.

HASTA que L sea la cadena vacía.

Salida: S y su longitud.

La heurística H3 forma la supersecuencia para el ejemplo 1 de la siguiente manera: Para cada símbolo del alfabeto se calcula $L-a$ y se obtiene la longitud de la Supersecuencia Común más Corta de acuerdo a H1, así se obtiene: $L-a = \{cacb, cbaac, bbacc\}$ y $|S| = 8$, $L-b = \{acacb, cbaac, bacc\}$ y $|S| = 8$, y $L-c = \{acacb, baac, bbacc\}$ y $|S| = 9$, después se obtiene el mínimo de las longitudes de las supersecuencias obtenidas por H1: $\min\{8, 8, 9\} = 8$ y con ello se forma $\Sigma' = \{a, b\}$. Ahora

para cada elemento de Σ' se calcula L_a y posteriormente se aplica H1 al lenguaje L_a obtenido, para el ejemplo esto quedaría como: $L_a = \{acacb\}$ con $|S| = 5$ y $L_b = \{bbacc\}$ con $|S| = 5$, ahora calculamos el valor máximo de las longitudes de las supersecuencias obtenidas $\max\{5,5\} = 5$ y con ello formamos $\Sigma^* = \{a, b\}$, de éste último conjunto se elige el símbolo que formará parte de la Supersecuencia, digamos b ($S = b$) y se procede a eliminar b del frente de cada cadena de L donde aparezca, obteniendo el nuevo lenguaje $L = \{acacb, cbaac, bacc\}$, el proceso se repite hasta que la cadena vacía sea la única cadena del lenguaje L .

2.5 Otras variantes.

Se proponen además algunas variantes para los algoritmos anteriores (MM, H1, H2 y H3). Una vez que se ha determinado la forma en la cual se elige el símbolo σ para que forme parte de la Supersecuencia S , puede ocurrir con mucha frecuencia (en el caso de cadenas de diferente longitud) que el símbolo σ no sea único, es decir, que existen empates, en cuyo caso las formas en que se determinará la forma de romper los empates serán las siguientes:

1. El símbolo que se concatenará para formar la supersecuencia S , se elige en orden lexicográfico.
2. La elección del símbolo se realiza de forma aleatoria(A).
3. El símbolo a elegir es aquel cuya longitud de la cadena donde se encuentre sea la máxima, en caso de que las cadenas coincidan en longitud, se toma en forma aleatoria(AS).

3 Experimentos.

Para el problema de la SCS no existen archivos benchmark, por lo cual los problemas se generaron de manera aleatoria, la longitud de la supersecuencia común más corta para cada problema se desconoce, sin embargo, se desea con estos experimentos observar y comparar los resultados de las heurísticas expuestas anteriormente para el caso promedio, y observar el comportamiento de la dispersión de los resultados en el rango del máximo y el mínimo valor encontrado para la longitud de la Supersecuencia, es decir, interesa observar cuál de todos los algoritmos encuentra la mejor solución con la menor dispersión y cuál algoritmo es el mejor en encontrar la mejor solución en el caso promedio.

Para cada experimento se reportan los siguientes resultados: el número 1 denota el promedio de la longitud de la SCS, el número 2 se refiere al valor máximo de la longitud de la SCS encontrado en todas las corridas y finalmente el número 3, reporta el valor mínimo de la longitud de la SCS hallado a lo largo de las corridas. Los algoritmos que se reportan son: Majority Merge (MM), Majority Merge aleatorio (MMA), Majority Merge de cadena máxima (MMAS), H1, H1 aleatorio (H1A), H1 de cadena máxima (H1AS), H2, H2 aleatorio (H2A), H2 de cadena máxima (H2AS), H3 usando H1 (H3-H1), H3 usando H1 aleatorio (H3-H1A), H3 usando H1 de cadena máxima (H3-H1AS), H3 aleatorio usando H1 (H3A-H1), H3 aleatorio usando H1 aleatorio (H3A-H1A), H3 aleatorio usando H1 de cadena máxima (H3-H1AS), H3

de cadena máxima usando H1 (H3AS-H1), H3 de cadena máxima usando H1 aleatorio (H3AS-H1A), y H3 de cadena máxima usando H1 de cadena máxima (H3AS-H1AS).

Para cada prueba generada aleatoriamente, se corrió 30 veces el mismo problema para reportar en la Tabla 1: en la columna 1 el valor promedio, en la columna 2 el máximo valor y en la columna 3 el mínimo valor, de la longitud de la Supersecuencia Común, el menor promedio encontrado por alguno de los algoritmos se remarca con un doble subrayado, y el valor más pequeño de la longitud durante las corridas encontrado durante la prueba se resalta con un estilo de “negritas”.

El primer experimento consta de 6 problemas que tienen la característica de que para cada lenguaje la longitud de cada una de las cadenas que forman el lenguaje es constante, cada problema está especificado de la siguiente manera:

- L1:** $|L| = 10, |\Sigma| = 3$, cada cadena de longitud 7.
- L2:** $|L| = 15, |\Sigma| = 3$, cada cadena de longitud 10.
- L3:** $|L| = 15, |\Sigma| = 4$, cada cadena de longitud 10.
- L4:** $|L| = 20, |\Sigma| = 5$, cada cadena de longitud 10.
- L5:** $|L| = 30, |\Sigma| = 5$, cada cadena de longitud 15.
- L6:** $|L| = 30, |\Sigma| = 5$, cada cadena de longitud 20.

Tabla 1. Resultados de 30 corridas.

ALG	L1			L2			L3			L4			L5			L6		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
MM	17.0	17	17	23.0	23	23	30.0	30	30	37.0	37	37	50.0	50	50	72.0	72	72
MMA	17.0	17	17	23.8	25	23	28.8	31	27	35.4	38	33	53.5	57	50	68.7	72	66
MMAS	17.0	17	17	23.0	23	23	28.4	29	28	35.7	36	35	52.0	52	52	66.5	67	66
H1	17.0	17	17	23.0	23	23	29.0	29	29	36.0	36	36	51.0	51	51	68.0	68	68
H1A	17.7	23	17	24.2	29	23	32.0	42	28	39.0	56	34	51.8	55	51	71.7	87	66
H1AS	17.0	17	17	23.0	23	23	28.7	29	28	36.2	37	35	51.0	51	51	66.7	68	65
H2	16.0	16	16	24.0	24	24	29.0	29	29	34.0	34	34	52.0	52	52	66.0	66	66
H2A	18.6	24	16	23.6	28	22	30.3	39	28	35.0	47	33	52.7	62	49	69.9	74	66
H2AS	18.2	23	16	23.4	26	22	30.1	36	28	34.8	40	33	52.3	62	50	68.5	72	66
H3-H1	16.0	16	16	22.0	22	22	28.0	28	28	32.0	32	32	49.0	49	49	64.0	64	64
H3-H1A	19.5	25	17	25.0	29	23	37.0	45	31	45.7	60	33	67.4	84	51	78.9	92	64
H3-H1AS	16.0	16	16	22.0	22	22	27.8	28	27	32.6	35	31	49.4	51	48	64.6	69	63
H3A-H1	16.0	16	16	22.0	22	22	28.0	28	28	31.4	32	31	49.0	49	49	64.0	64	64
H3A-H1A	19.9	28	17	25.1	29	23	37.2	44	32	46.4	64	35	68.0	90	56	80.4	101	66
H3A-H1AS	16.0	16	16	22.0	22	22	27.6	28	27	32.5	35	31	49.5	51	48	64.6	68	63
H3AS-H1	16.0	16	16	22.0	22	22	28.0	28	28	31.0	31	31	49.0	49	49	64.0	64	64
H3AS-H1A	19.3	22	16	25.5	32	22	36.4	42	32	45.2	54	37	65.6	82	52	81.1	92	64
H3AS-H1AS	16.0	16	16	22.0	22	22	27.7	28	27	32.5	37	31	49.6	51	49	65.0	67	63

Se puede observar en la tabla anterior que el algoritmo *H3AS-H1* encuentra la mayoría de veces (5 de 6), en promedio las mejores Supersecuencias (excepto para **L3**), en segundo lugar los algoritmos que encuentran supersecuencias más cortas son *H3-H1* y *H3A-H1* (4 de 6 veces). Se puede observar también que para **L2** y **L4**, se llega a encontrar una Supersecuencia de longitud igual a la encontrada en el promedio por *H3AS-H1*, por algoritmos diferentes a los que encontraron la menor supersecuencia en el promedio de 30 corridas; para **L5** y **L6** se encuentran, incluso, longitudes de la supersecuencia menores que los encontrados por *H3AS-H1* en el promedio

de 30 corridas. Considerando la Supersecuencia de menor longitud como aquella que se reporta como mínima al calcular el promedio de las 30 corridas, el algoritmo que encuentra con mayor frecuencia las de menor longitud es *H3-HIAS*, es decir, *H3* usando *H1* donde la forma de romper empates es de cadena máxima. Finalmente también se observa en la Tabla 1 que en general, los algoritmos que rompen los empates de manera aleatoria (A) tienen un rango amplio entre la longitud mínima y la longitud máxima encontrada para una Supersecuencia, mientras que los algoritmos que rompen el empate de acuerdo a la máxima longitud de la cadena (AS) se dispersan menos, es decir, la diferencia encontrada entre el valor máximo y el valor mínimo de la longitud de la Supersecuencia es poca (entre cero y once). También se realizó 100 veces el mismo experimento ejecutando cada algoritmo para los seis problemas descritos y se observó que los resultados no varían cuando el número de veces que se ejecutan los algoritmos aumenta.

El segundo experimento consta de lenguajes en los cuales la heurística MM tiene un mal comportamiento y la Supersecuencia común más corta es conocida [1]. Son tres lenguajes descritos de la siguiente manera:

L1: 9 cadenas a^{40} , 4 cadenas ba^{39} , 2 cadenas bba^{38} y 1 cadena $bbba^{37}$.
L2: 9 cadenas a^{40} , 4 cadenas $b^{13}a^{27}$, 2 cadenas $b^{26}a^{14}$ y 1 cadena $b^{39}a$.
L3: 8 cadenas $a^{20}b^{20}$, 8 cadenas $b^{20}c^{20}$.

Para **L1**, la solución óptima conocida es una cadena de longitud 43; la longitud de la cadena óptima para **L2** es de 79 y para **L3** la longitud de la Supersecuencia óptima es de 60.

Se observa en los resultados de la Tabla 2, que para **L1**, el algoritmo *Majority Merge* en sus tres versiones obtiene las peores longitudes de la Supersecuencia, para el caso promedio y para el valor mínimo de todas las corridas; los algoritmos *H3* en sus diferentes versiones, salvo aquellos combinados con *H1A*, son los que obtienen los valores óptimos en el caso promedio y en el valor mínimo de todas las corridas.

Para **L2**, nuevamente los algoritmos *H3* en todas sus versiones, excepto los que se combinan con *H1A*, son los que obtienen los valores óptimos conocidos tanto para el caso promedio como para el valor mínimo de todas las corridas.

Se observa que para el lenguaje **L3**, tanto *Majority Merge* como *H2* obtienen el valor óptimo de la Supersecuencia.

Por lo que resumiendo, la familia de *H3*, excepto la que se combina con *H1A*, obtiene el óptimo conocido en dos ocasiones de tres problemas propuestos. Además, se observa que la dispersión de la familia *H3*, excepto la que se combina con *H1A*, es cero para los problemas que encuentra el óptimo conocido.

También se realizó el mismo experimento 100 veces y se observó que los resultados no varían cuando el número de veces que se ejecutan los algoritmos para resolver el mismo problema aumenta.

4 Conclusiones.

De acuerdo a los resultados obtenidos al realizar los experimentos, con problemas generados aleatoriamente, se puede concluir lo siguiente:

En general, la mejor heurística es H3AS-H1, ya que encontró la menor supersecuencia común en la mayoría de los problemas presentados. Cuando existen empates en el caracter a adicionar a la Supersecuencia, la variante que toma en orden lexicográfico dicho caracter, encuentra mejores resultados que las de forma aleatoria o las guiadas por la longitud de la cadena, en caso de que se reporte el promedio de las longitudes.

Si la longitud de la Supersecuencia a reportar es aquella que tenga el mínimo valor a lo largo de todas las corridas efectuadas, en caso de empate, los algoritmos que utilizan el criterio de cadena máxima para tomar la decisión sobre el caracter a adicionar a la Supersecuencia son los mejores.

Se observa menor dispersión en los algoritmos, en caso de empate, lo deciden tomando en consideración la longitud de la cadena.

Para los lenguajes especiales con mal comportamiento, los mejores algoritmos son los que corresponden a la familia H3, excepto los que se combinan con H3-H1A.

No es significativo el aumento en la cantidad de corridas para cada problema.

Tabla 2. Resultados de lenguajes especiales: 30 corridas.

ALG	L1			L2			L3		
	1	2	3	1	2	3	1	2	3
MM	157.00	157	157	121.00	121	121	<u>60.00</u>	60	60
MMA	157.00	157	157	121.00	121	121	77.53	80	68
MMAS	157.00	157	157	121.00	121	121	79.30	80	79
H1	92.00	92	92	91.00	91	91	79.00	79	79
H1A	92.93	97	92	92.23	97	91	80.06	83	79
H1AS	92.00	92	92	91.00	91	91	79.56	80	79
H2	116.00	116	116	121.00	121	121	<u>60.00</u>	60	60
H2A	81.66	86	80	121.00	121	121	79.36	86	72
H2AS	81.36	90	80	121.00	121	121	79.53	88	67
H3-H1	<u>43.00</u>	43	43	<u>79.00</u>	79	79	78.00	78	78
H3-H1A	50.90	55	46	80.16	85	79	80.73	90	76
H3-H1AS	<u>43.00</u>	43	43	<u>79.00</u>	79	79	77.93	79	76
H3A-H1	<u>43.00</u>	43	43	<u>79.00</u>	79	79	78.66	79	78
H3A-H1A	50.50	56	45	80.46	83	79	79.96	86	73
H3A-H1AS	<u>43.00</u>	43	43	<u>79.00</u>	79	79	78.30	79	77
H3AS-H1	<u>43.00</u>	43	43	<u>79.00</u>	79	79	78.20	79	78
H3AS-H1A	50.13	60	44	80.06	85	79	79.96	90	73
H3AS-H1AS	<u>43.00</u>	43	43	<u>79.00</u>	79	79	78.36	79	76

Agradecimientos.

Agradecemos el apoyo financiero de la Vicerrectoría de Investigación y Estudios de Posgrado de la Benemérita Universidad Autónoma de Puebla a través del proyecto 0333 Aplicación de las Meta-heurísticas Inspiradas en la Naturaleza.

Referencias

1. Branke J., Middendorf M., Schneider F.: Improved heuristic and a genetic algorithm for finding short supersequences. *OR Spektrum*. Vol. 20, 39–45. (1998).
2. Kang N. and Hon Wai L.: Towards a better solution to the shortest common supersequence problem: the deposition and reduction algorithm. *BMC Bioinformatics*. 7(Suppl 4):S12. (2006).
3. Foulser, D. E., Li M., Yang Q.: Theory and algorithms for plan merging, *Artificial Intelligence*. vol. 5, 143-181. (1992).
4. Fraser C. B., Irving R. W.: Approximation algorithms for the shortest common supersequence. *Nordic Journal of Computing*. Vol. 2, 303–325. (1995).
5. Hubbell E. A., Morris M. S., and Winkler J. L.: Computeraided engineering system for design of sequence arrays and lithographic masks. US Patent 5571639. (1996).
6. Middendorf M.: More on the complexity of common superstring and supersequence problems. *Theoretical Computer Science*. Vol. 125, 205–228. (1994).
7. Ning K., Leong H. W.: Towards a better solution to the shortest common supersequence problem: the deposition and reduction algorithm. *Symposium of Computations in Bioinformatics and Bioscience with the International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*. Vol. 1, 84-90 (2006).
8. Rahmann S.: The shortest common supersequence problem in a microarray production setting Approximation algorithms for the shortest common supersequence. *Bioinformatics*, Vol. 19, 156–161. (2003).
9. Timkovsky V. G.: Complexity of common subsequence and supersequence problems and related problems. *Cybernetics and Systems Analysis*. Vol. 25, 565–580. (1990).

Identificación de Locutor usando Vectores Acústicos basados en Cuantiles

José-Martín Olguín-Espinoza¹, Pedro Mayorga-Ortiz², Luis Vizcarra³.

^{1,3}Universidad Autónoma de Baja California, Mexicali, B.C., México

¹molguin@uabc.edu.mx, ³luiivi@uabc.edu.mx

²Instituto Tecnológico de Mexicali, Mexicali, B.C., México

²pedromayorga@hotmail.com

Paper received on 16/07/12, Accepted on 29/08/12.

Resumen. En los sistemas de reconocimiento de locutor, es importante la representación de la señal de voz en términos de vectores acústicos, ya que éstos la caracterizan. Aquí se propone el uso de vectores cuantílicos para sistemas de identificación de locutor. Para evaluar esta propuesta, se efectuaron experimentos creando modelos mezclados gaussianos de distintos tamaños usando vectores cuantílicos de distinta dimensión. Los experimentos fueron realizados con dos bases de datos en español: CEM y AHUMADA. Los resultados fueron alentadores ya que mostraron un muy buen desempeño con modelos de 10 densidades gaussianas.

Palabras Clave: Reconocimiento de Locutor, Vectores Cuantílicos.

1 Introducción

Los sistemas de reconocimiento automático de locutor modelan las características articulatorias y fisiológicas del tracto vocal, con el fin de extraer y realzar sus parámetros característicos. Entre las técnicas de extracción de parámetros, destacan *Mel Frequency Cepstral Coefficients* (MFCC), *Linear Filter Cepstral Coefficients* (LFCC), *Perceptual Linear Predictive Analysis* (PLP) y *Relative Spectral PLP* (RASTA PLP) [1, 2]. Cada técnica busca modelar el tracto vocal, o la forma en la que el oído humano funciona o como el tracto vocal genera la voz.

Una técnica escasamente encontrada en la literatura de reconocimiento automático de locutor es una que trata de aprovechar información concerniente a la capacidad pulmonar a partir de medidas acústicas. Dicho enfoque ha sido aplicado para fortalecer diagnósticos médicos, midiendo las tasas de flujo y volumen de la respiración en términos de cuantiles [3, 4].

En el presente trabajo se propone la construcción de vectores acústicos obtenidos a partir del análisis de los cuantiles en el dominio de la frecuencia como una alternativa a las técnicas ya citadas.

2 Vectores Acústicos para Voz

El análisis de la señal de voz comúnmente se realiza sobre segmentos elementales cuasi-estacionarios llamados *ventanas de análisis* (o tramas); en este proceso se toma una señal de voz y se particiona en ventanas de cierto tamaño (típicamente en el orden de los milisegundos), las cuales están traslapadas entre sí. A cada ventana se le aplica análisis espectral para generar un vector acústico correspondiente.

Una aproximación exitosa es la deconvolución Cepstral [5], la cual permite aislar las frecuencias fundamentales de la voz de aquellas que son generadas por el conducto vocal. Esto se obtiene aplicando la transformada discreta cosenoidal (*cosine discrete transform*, CDT) a los vectores espectrales previamente calculados mediante la transformada rápida de Fourier (FFT).

Una extensión de los principios cepstrales y su paso a un espacio frecuencial no lineal relacionado con la audición humana y muy exitoso son los Mel-Frequency Cepstral Coefficients (MFCC) [1, 5], como se muestra en la Fig. 1. Una variante de MFCC es *Linear Frequency Cepstral Coefficients* (LFCC), donde los filtros son repartidos uniformemente sobre una escala lineal de frecuencias [5, 6].

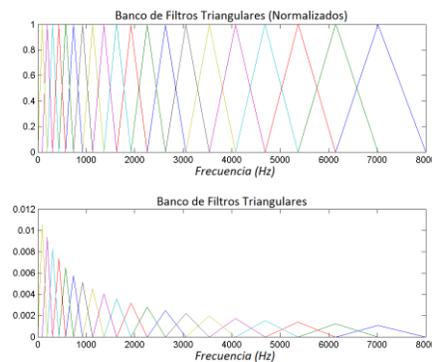


Figura 1. Distribución de las frecuencias en la escala Mel

Una aproximación muy difundida es la codificación lineal predictiva LPC (*Linear Predictive Coding*). Este método se basa en la hipótesis de que la voz puede ser modelada por un proceso lineal predictivo [7, 8].

Existen otros métodos, como PLP, que modifican el espectro de potencia de la voz antes de obtener una aproximación por un modelo auto-regresivo [2]. Además, hay otras variantes de esta metodología que son más adaptadas a un canal de comunicación tales como Relative Spectral PLP (RASTA PLP) [2, 7].

3 Antecedentes en Vectores basados en Cuantiles

Considerando evidencias en otros trabajos [3, 4], se propone analizar señales de voz, pero con un tratamiento basado en vectores cuantílicos, con el propósito de distinguir elementos que permitan caracterizar y reconocer a los locutores.

3.1 Cuantiles

En teoría estadística, las medidas de tendencia no-central permiten conocer aspectos particulares de una distribución (como en el caso de vectores acústicos). Dentro de estas medidas, unas de las más importantes son los *cuantiles*. En estas variables los datos son ordenados de forma creciente, dividiendo la función de distribución en partes, de tal forma que cada una contiene la misma cantidad de área.

Los cuantiles se basan en la función de distribución acumulativa (CDF). El cuantil q_p de una variable aleatoria está definido como el número q más pequeño tal que la función de distribución acumulativa es mayor o igual a algún valor p , donde p se encuentra entre $0 < p < 1$. Esto puede ser calculado para el caso de una función de distribución continua con su función de densidad $f(x)$ resolviendo la Ec. 1:

$$p = \int_{-\infty}^{q_p} f(x) dx \quad (1)$$

Como ejemplo, los *cuartiles* pueden representarse por la notación $q_{0.25}$, $q_{0.5}$, y $q_{0.75}$, respectivamente. En esencia, estos dividen la distribución en cuatro segmentos de igual probabilidad bajo la curva (Fig. 2). El segundo cuartil es conocido como la *mediana* y es muy utilizado en estadística, el cual satisface la Ec. 2:

$$0.5 = \int_{-\infty}^{q_{0.5}} f(x) dx \quad (2)$$

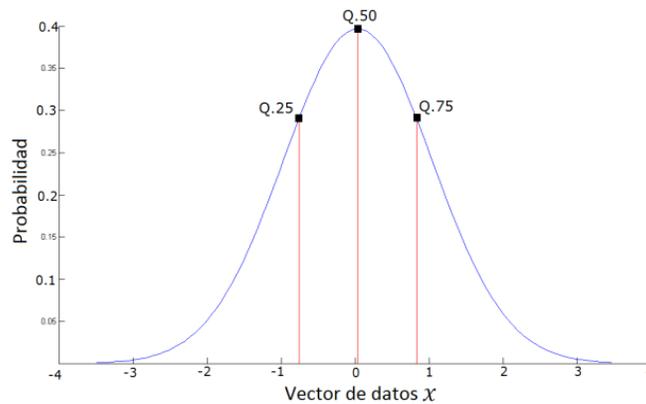


Figura 2. Representación gráfica del concepto de Cuartil

Una de las funciones de distribución de probabilidad (PDF) más conocidas es la distribución Gaussiana (o Normal). Donde la CDF continua está dada por la Ec. 3.

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad (3)$$

Donde el valor cuantílico de interés sería x . Cuando no es conveniente suponer una PDF específica, podemos aplicar la CDF, la cual se obtiene a partir de histogra-

mas de frecuencias de los vectores acústicos, sobre los cuales se pueden ajustar modelos mezclados gaussianos (GMM).

3.2 Vectores Acústicos basados en Cuantiles

Lo que se propone en el presente trabajo es, a partir de los datos de la señal de voz X obtener su FFT, normalizarla y tomarla como la función de distribución de frecuencias o una función de densidad de frecuencia. De esta forma se puede obtener un vector $Q = (q_1, q_2, \dots, q_n)$ sobre los valores frecuenciales, donde cada q_i es el valor de la frecuencia asociado al valor porcentual acumulado bajo la CDF específico del cuantil i . En otras palabras, Q representará un vector cuantílico para la señal X . De esta forma podemos hablar de vectores cuantílicos, cuando sus elementos representan los valores frecuenciales para el 25%, 50% y 75% del área bajo la curva de la FFT de X normalizada; o vectores octílicos en el caso de obtener estos valores para 12.5%, 25%, 37.5%, 50%, 62.5%, 75% y 87.5% del área.

El vector cuantílico se tomará como el vector acústico que representa un segmento, dígame estacionario, de la señal original X para efectos del modelado del locutor. Bajo esta idea los vectores se pueden obtener analizando la señal de voz de un locutor en dos modalidades: tiempo largo y tiempo corto.

El análisis en tiempo largo consiste en obtener un sólo vector cuantílico resultante de la señal de un registro completo de entrada (Fig. 3). Por otra parte, en el análisis en tiempo tiempo corto, se definen ventanas de cierta longitud de tiempo w (regularmente en el orden de los milisegundos y dentro del rango estacionario de la señal), de tal forma que por cada ventana se genera un vector. Además, se establece un valor de traslape $o < w$, de tal forma que el análisis se va realizando por tramas traslapadas, a fin de hacer la extracción de características en segmentos estacionarios. En análisis de tiempo corto se tendrán $T/(w-o)$ vectores para cada señal, donde T es el tiempo total de duración de la señal, w es el tamaño de la ventana y o representa el tiempo de traslape. Esta modalidad se esquematiza en la Fig. 4.

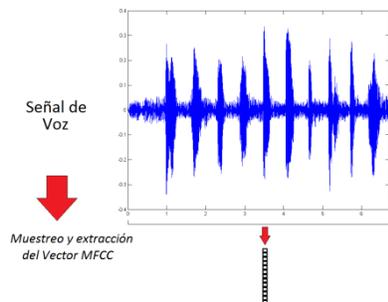


Figura 3. Vectores en tiempo largo.

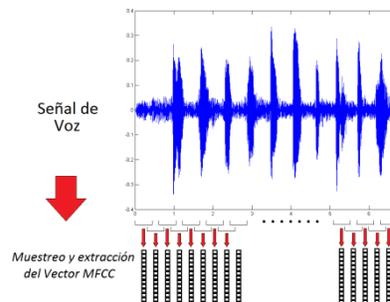


Figura 4. Vectores en tiempo corto.

4 Reconocimiento Automático de Locutor con GMM

El reconocimiento automático de locutor (RAL) es un término genérico que denota la identificación y verificación de locutor, siendo la primera el objetivo de este trabajo.

4.1 Identificación Automática del Locutor

La Identificación Automática del Locutor consiste en determinar de entre una población de locutores conocidos, la persona a la que pertenece cierta señal de voz dada como entrada. En la identificación se proponen dos modos: *Identificación en conjunto cerrado*, para el cual se asume que la señal de voz es pronunciada por un locutor conocido por el sistema. La salida del sistema de identificación en este modo será el identificador del locutor con la mayor similitud a la señal de voz de entrada. Por otro lado, en *Identificación en conjunto abierto*, para la cual cabe la posibilidad de que el locutor pueda no pertenecer al conjunto de locutores conocidos por el sistema, es decir que el locutor sea un impostor. En identificación en conjunto abierto, el sistema de identificación debe decidir la fiabilidad de su juicio aceptando o rechazando la identidad que encontró. Si el sistema la acepta debe además establecer el identificador del locutor al que pertenece la señal de voz de entrada.

4.2 Modelado de Locutor con Modelos Mezclados Gaussianos

Los GMM [9] consisten en una suma ponderada de M componentes de densidades, como lo muestra la Ec. 4:

$$p(\vec{x} | \lambda) = \sum_{i=1}^M m_i b_i(\vec{x}) \quad (4)$$

Donde \vec{x} es un vector aleatorio d -dimensional; $b_i(\vec{x})$ son las densidades, m_i representan las ponderaciones para cada mezcla o densidad Gaussiana, con la restricción de $\sum_{i=1}^M m_i = 1$ lo que es característico en una verdadera función de densidad de probabilidad. Cada componente de densidad es una función Gaussiana de dimensión D descrita por:

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_i)' \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right\} \quad (5)$$

Donde $\vec{\mu}_i$ corresponde al vector de medias y Σ_i es la matriz de covarianza. Un modelo GMM completo está parametrizado por m_i , $\vec{\mu}_i$ y Σ_i . En reconocimiento, cada locutor r está representado por un modelo GMM de la siguiente forma: $\lambda_r = \{m_i, \vec{\mu}_i, \Sigma_i\}$; $i = 1, \dots, M$ donde M representa el número de componentes gaussianas utilizadas para modelar al locutor r .

4.3 Identificación de Locutor con GMM y UBM

Los sistemas de identificación se dividen en dos etapas: entrenamiento y evaluación. En el entrenamiento se toman las señales de cada locutor para construir su modelo GMM correspondiente. Para realizar esta tarea, el algoritmo de Máxima Expectación (EM) es uno de los más exitosos [5, 10, 11, 12]. EM parte de un modelo λ y mediante iteraciones sucesivas trata de encontrar un λ tal que:

$$p(X | \bar{\lambda}) \geq p(X | \lambda) \quad (6)$$

El nuevo modelo resultante de cada iteración se toma como modelo inicial para la siguiente, y el proceso se repite hasta que se alcanza un umbral de convergencia basado en la maximización de la probabilidad. En cada iteración EM, se deben reestimar nuevos parámetros del modelo $\bar{\lambda}$. La elección del modelo inicial λ se puede hacer de varias maneras: Una es calcularlo aleatoriamente, la otra es utilizar lo que se conoce como el Universal Background Model (UBM). El UBM es un modelo que se construye a partir de las señales combinadas de todos los locutores [12], de esta forma se obtiene un modelo cuyos parámetros se usarán como valores iniciales en el algoritmo EM en el cálculo de cada modelo cliente.

La evaluación comprende seleccionar al modelo λ_i que resulte con la probabilidad más alta de una señal x desconocida, como se muestra en la Ec. 7:

$$\prod_{t=1}^T p(\vec{x}_t | \lambda_i) > \prod_{t=1}^T p(\vec{x}_t | \lambda_r); r = 1, \dots, I \quad (7)$$

5 Experimentos

Se construyó un Sistema de Identificación de Locutor, el cual fue evaluado con diferentes bases de datos de voz. Los vectores acústicos fueron construidos en base a cuartiles, octiles y deciles, eligiendo los que presentaran mejor nivel de reconocimiento.

5.1 Base de Datos de señales de Voz

El sistema de identificación fue validado con dos diferentes bases de datos: Corpus en Español Mexicano [13] y AHUMADA [14], las características de cada una se describen en los apartados siguientes.

5.1.1 Corpus en Español Mexicano (CEM)

Esta base de señales de voz contiene grabaciones de 33 personas (20 hombres y 13 mujeres), y fue desarrollada por la Universidad Autónoma de Baja California (UABC) [13]. Cada locutor grabó a lo largo de tres sesiones, con al menos 15 días de separación, tres frases siguiendo la distribución fonética del español mexicano y

un texto fijo; además durante cada sesión se grabaron elocuciones tanto con micrófono como con teléfono y utilizando un sistema de voz sobre IP (VoIP), totalizando 36 grabaciones por persona. Cada señal fue muestreada a 8Khz, 16 bits por dato, un canal (mono) y almacenados en formato WAV.

5.1.2 AHUMADA

Corpus en español Ibérico [14], consiste de registros de 25 locutores (hombres) los cuales grabaron dígitos, frases fonéticamente equilibradas, texto fijo, texto específico para cada locutor y además conversación espontánea. Cada tipo de elocución se grabó por medio de micrófono y teléfono. Cada elocución de las frases se repitió en tres sesiones al menos, con 15 días de separación entre ellas.

5.2 Obtención de Vectores Cuantílicos

En la obtención de los vectores se parte de un archivo WAV, seguida opcionalmente de un preprocesamiento de preénfasis y antitraslape [5]. Posteriormente, se aplica la transformada rápida de Fourier (FFT). Para cumplir con el principio básico para una distribución de probabilidad, se normaliza el área bajo la curva, garantizando que sea igual a 1. Enseguida, se buscan los valores para los cuantiles. Por ejemplo, para octiles se calculan los valores frecuenciales $f_{0.125}, \dots, f_{0.875}$, aplicando la Ec. 8. En los experimentos realizados se utilizaron cuantiles, octiles y deciles para comparar el porcentaje de identificación con diferentes configuraciones.

$$A = .125 = \int_{-\infty}^{f_{0.125}} F_N(f) df, \dots, A = .875 = \int_{-\infty}^{f_{0.875}} F_N(f) df \quad (8)$$

5.3 Construcción de los Modelos de Locutor

Una vez obtenidos los vectores cuantílicos para todos los locutores, se construyeron los modelos GMM de cada locutor con dos configuraciones. La primera utilizando un modelo aleatorio como modelo inicial para el algoritmo Expectación-Maximización (EM). Para la segunda configuración se creó un Universal Background Model (UBM) utilizando los datos de entrenamiento de todos los locutores combinados en un solo modelo GMM [12], posteriormente los modelos para cada locutor se crearon utilizando el UBM como modelo inicial para el algoritmo EM.

Los modelos GMM también fueron creados con configuraciones de 10 y 32 componentes gaussianas para las distintas combinaciones de tamaños de vectores (cuantiles, octiles y deciles).

6 Resultados y Discusión

Debido a que las bases de datos no tienen el mismo tamaño no fueron realizadas siguiendo el mismo protocolo, los experimentos tuvieron variaciones en cuanto al

número de archivos utilizados para entrenamiento y evaluación. En consecuencia, los experimentos serán explicados por separado para cada base de datos. Sin embargo, lo importante dentro de nuestra contribución está en mostrar la potencialidad de los vectores cuantílicos con distintas bases de datos las cuales fueron creadas bajo distintas condiciones de grabación e idioma.

Para el caso particular de la base de datos CEM se usaron los 20 locutores masculinos, tomando 40 segundos de entrenamiento, correspondiendo a las dos primeras sesiones de las frases 1,2 y 3, que son fonéticamente equilibradas. Para la evaluación se utilizaron 60 segundos correspondientes a la frase 4 (texto fijo) de la sesión 3. Debido a que nuestro sistema está orientado a la identificación de locutor, la medición de eficiencia fue efectuada en función de las clasificaciones correctas de locutor. La Tabla 1 muestra los resultados de eficiencia de reconocimiento considerando entrenamiento y evaluación de señales adquiridas por el mismo medio micrófono (M1), teléfono (T1) y VoIP (T3). La parte superior de la tabla indica el tipo de cuantil, el número de componentes gaussianas en los modelos GMM y los medios de adquisición/evaluación. El número de componentes de los vectores fueron 3, 7 y 9 para cuartiles, octiles y deciles. En este experimento se obtuvieron los mejores resultados con los vectores de mayor dimensión (deciles).

Tabla 1. Experimentos con CEM sin pre-procesamiento.

Cuantil	GMM	MI	T1	T3
Octil	10	35 %	32.00 %	36.66 %
Cuartil	10	25 %	8.33 %	31.66 %
Decil	10	47 %	20.00 %	40.00 %
Octil	32	20 %	17.00 %	23.00 %
Cuartil	32	17 %	8.33 %	26.66 %
Decil	32	33 %	12.00 %	35.00 %

Lo anterior pone de manifiesto que el vector cuantílico con más dimensiones captura con mayor precisión la estructura fina y/o densidad espectral importante del aparato fonatorio del locutor. Asimismo se observa que los resultados no mejoran cuando se aumenta el número de componentes gaussianas en los modelos GMM, denotándose mejores resultados con mezclados de 10 componentes gaussianas. Lo anterior no significa que modelos más pequeños mejoren los resultados, puesto que aquí no mostramos otras evaluaciones menos satisfactorias con modelos de menor tamaño. Lo que sí se puede observar es que modelos con más de 10 gaussianas pueden conducir a errores de otra naturaleza como sobreentrenamiento o redondeo para este sistema en particular. Otro aspecto interesante que nos permite emitir un juicio es que el mejor resultado fue con deciles y directamente con micrófono. Este hecho pone de manifiesto por un lado que el decil capturó mejor la estructura fina del locutor y que el usar directamente micrófono implicaba menos degradación en los modelos debida a ruido o a efectos de canal.

En base a estos resultados, se decidió construir modelos con vectores decílicos y tratando de mejorar los resultados se aplicó preprocesamiento que consiste en preénfasis en segmentos de tiempo corto de 30ms y traslape de 20ms. Además se construyó el modelo impostor aplicando UBM. En la Tabla 2 se puede observar una me-

oría evidente en la eficiencia de identificación de locutor. De nuevo los modelos con 10 GMM arrojaron resultados ligeramente mejores que los de 32 GMM. Lo anterior refuerza la idea de que modelos más complejos pueden conducir a errores de redondeo o problemas de sobre-entrenamiento y que los vectores decílicos son más consistentes con modelos de 10 gaussianas. Otro aspecto relevante es que tanto el preénfasis como el uso de UBM lograron paliar en gran medida los problemas derivados de los medios de transmisión o canal como se observa para el caso de la señal de teléfono (T1). Asimismo estos resultados refuerzan la hipótesis de que los cuantiles para señales orientadas a la identificación de locutor tiene potencial de aplicación.

Tabla 2. Experimentos con CEM y modelos GMM usando UBM

Cuantiles	GMM	MI	T1
Decil	10	75 %	80.00 %
Decil	32	75 %	75.00 %

Debido a que este trabajo presenta dos variantes con respecto a la mayoría de los trabajos en identificación de locutor, las cuales son, el uso de un corpus en español mexicano y la aplicación de vectores cuantílicos, a manera de comparación se realizaron experimentos con la bases de datos AHUMADA, la cual contiene grabaciones en español Ibérico. Se trató de establecer una partición entrenamiento/prueba con condiciones similares a las utilizadas en los experimentos con CEM. Para este caso se tomaron las grabaciones de 20 locutores, el entrenamiento se realizó con las 10 elocuciones fonéticamente equilibradas de las sesiones 1 y 2 (identificadas con el código C), para tener un total aproximado de 40 segundos de señal. Para la evaluación se usó la elocución fonéticamente equilibrada (identificada con el código D) de la sesión 3 de aproximadamente 60 segundos de duración.

En la Tabla 3 se muestran los resultados de estos experimentos, notándose la similitud con los mostrados en la Tabla 2. Los resultados aquí obtenidos con micrófono son iguales, lo cual demuestra consistencia en los métodos de grabación, sin embargo las grabaciones efectuadas por teléfono evidencian un mejor control en el caso de CEM. Lo cual es consistente con lo reportado por los autores de AHUMADA [14]. Un aspecto interesante a resaltar es que los modelos acústicos con 32 gaussianas lograron mejorar la eficiencia de reconocimiento con respecto a los de 10, es decir de alguna manera se adaptaron más a la distorsión debida al canal de transmisión. Pero lo más importante es que nuevamente los resultados refuerzan el potencial de los vectores cuantílicos.

Tabla 3. Experimentos con AHUMADA y modelos GMM usando UBM

Cuantiles	GMM	MI	T1
10	10	75 %	25.00 %
10	32	75 %	30.00 %

7 Conclusiones

Los vectores acústicos cuántlicos como una propuesta novedosa en el campo de vectores acústicos, cuya representación está sustentada en trabajos relacionados con medicina sobre la capacidad de flujo respiratorio, que para el caso de voz siguió un tratamiento de transformada de Fourier en tiempo corto para tomar en cuenta la cuasi-estacionariedad de la señal.

Dichos vectores proporcionan la oportunidad de relacionar la energía con los valores frecuenciales más importantes de la señal de voz. Hasta el momento se obtuvo una tasa de identificación correcta del 80% en análisis de tiempo corto para elocuciones obtenidas de CEM. Los vectores cuántlicos fueron aplicados sobre señales con un mínimo de preprocesamiento (preénfasis y antialiasing) lo cual deja la posibilidad de explorar variantes más elaboradas de esta técnica e intentar obtener mejores resultados.

Como trabajo futuro se plantea mejorar el sistema aplicando diversas técnicas tanto en el preprocesamiento, tales como suprimir información no correspondiente a voz (VAD) y técnicas de normalización como las reportadas en [5]; así como en el cálculo de modelos GMM (entrenamiento con MAP), todo orientado a mejorar la eficiencia de la tarea de identificación de locutor.

Referencias

1. Faundez-Zanuy, M.; Monte-Moreno, E.: State-of-the-art in speaker recognition, *Aerospace and Electronic Systems Magazine, IEEE*, vol.20, no.5, pp.7-12, May 2005 doi: 10.1109/MAES.2005.1432568.
2. Hermansky H. and Fousek P.: Multi-resolution RASTA filtering for TANDEM-based ASR, in *Proceedings of the European Conference on Speech Communication and Technology*, Lisbon, Portugal, 2005.
3. Mayorga P., Druzgalski C., González O.H., Zazueta A., Criollo M.A.: *Expanded Quantitative Models for Assessment of Respiratory Diseases and Monitoring*, PAHCE 2011 IEEE, Rio de Janeiro. March 2011; ISBN 978-1-4244-6291-9; DOI: 10.1109/PAHCE.2011.5871938
4. Mayorga P., Druzgalski C., González O.H.: *Quantile Vectors based Verification of Normal Lung Sounds*, PAHCE 2012 IEEE, Miami, USA. March 2012; ISBN 978-1-4244-6291-9; DOI:
5. Bimbot F., Bonastre J-F., Corinne Fredouille, Guillaume Gravier, Ivan Magrin-Chagnolleau, Sylvain Meignier, Teva Merlin, Javier Ortega-Garcia, Dijana Petrovska-Delacretaz, and Douglas A. Reynolds.: *A tutorial on text-independent speaker verification*, *EURASIP J. Appl. Signal Process.* 2004 (January 2004), 430-451. DOI=10.1155/S1110865704310024 <http://dx.doi.org/10.1155/S1110865704310024>
6. Istrate D.M.: *Detection et Reconnaissance des Sons pour la Surveillance Médicale*, Thèse pour obtenir le grade de docteur de l'INPG, spécialité Signal, Image, Parole, Télécoms, le 16 décembre 2003, 183 p.
7. Solé-Casals J., Zaiats V.: *Advances in Nonlinear Speech Processing*, *International Conference on Nonlinear Speech Processing*, Nolisip 2009, Vic, Spain, June 25-27, 2009, Springer, 2010, ISBN 364211508X, 9783642115080.
8. Milner B. and James A.: *Robust Speech Recognition Over Mobile and IP Networks in Burst-Like Packet Loss*, *IEEE Transactions On Audio, Speech, And Language Pro-*

- cessing, Vol. 14, No. 1, January 2006, ISSN: 1558-7916, DOI:10.1109/TSA.2005.852997.
9. Reynolds D. A.: An Overview of Automatic Speaker Recognition Technology, ICASSP 2002 (IEEE International Conference on Acoustics, Speech and Signal Processing) , Orlando, Florida, USA, May 13 - 17, 2002.
 10. Dempster A. P., Laird N. M. and Rubin D. B.: Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc. (B), vol. 39, no. 1, pp. 1–38, 1977
 11. Martinez W.L. and Martinez A.R.: Computational Statistics Handbook with Matlab, Second Edition, Chapman & Hall/CRC, 2008, ISBN 1-58488-566-1.
 12. Reynolds D.A., Quatieri T.F., and Dunn R.B.: Speaker Verification Using Adapted Gaussian Mixture Models Digital Signal Processing 10, 19–41 (2000) doi:10.1006/dspr.1999.0361
 13. Olgún J.M. and Mayorga P.: Corpus de Voz en Español Mexicano Para Experimentación en Reconocimiento Automático de Locutor, RESEARCH IN COMPUTING SCIENCE, CIC-IPN, 2010, Vol. 50, ISSN 18770-4069, México.
 14. Ortega-García J., González-Rodríguez J., Marrero-Aguilar V., Díaz-Gómez J., García-Jimenez M., Lucena-Molina J., Sánchez-Molero J.: AHUMADA: A large speech corpus in spanish for speaker Characterization and Identification, Speech Communication, Vol. 3, pp. 255-264, Junio 2000.

Identificación de personas por medio de la voz utilizando LPC's y reconocimiento por KNN (K-Nearest Neighbors)

Jesús-Gustavo Medrano-Romero¹, Yesenia Nohemí González-Meneses²,
José Federico Ramírez-Cruz³ y Blanca Estela Pedroza-Méndez⁴

^{1,2,3,4} Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México
¹didier9_10@hotmail.com, ²yeseniaglez@hotmail.com, ³federico_ramirez@yahoo.com
⁴thelismolina@hotmail.com

Paper received on 08/08/12, Accepted on 10/08/12

Resumen. El reconocimiento de voz es un área de la inteligencia computacional que sigue siendo explotada y muy utilizada en la actualidad por diferentes empresas e instituciones, logrando cada vez mejores resultados en cuanto al porcentaje de reconocimiento; dicho reconocimiento podemos dividirlo en dos grandes ramas, reconocimiento de habla (que intenta identificar lo que la persona dice) y reconocimiento de hablante (que intenta encontrar a la persona que habla). Este trabajo está enfocado directamente al reconocimiento de hablantes por lo que se presenta un modelo para el procesamiento de la señal obtenida de la voz y otro para el entrenamiento, al final se hace el reconocimiento de dicha señal con el objetivo de identificar al hablante. Esto se lleva a cabo a través de la aplicación de LPC's para el procesamiento de la señal y el reconocimiento se hace mediante la combinación de KNN para encontrar elementos parecidos, probabilidad para determinar la mejor solución y actualización de pesos para garantizar un mejor funcionamiento.

Palabras Clave: Reconocimiento de Voz, Identificación de hablante.

1 Introducción.

El proceso de reconocimiento automático del habla permite a las máquinas recibir mensajes hablados u orales. Tomando como entrada la señal acústica recogida por un micrófono, el proceso de reconocimiento automático del habla tiene como objetivo final decodificar el mensaje contenido en la señal acústica para realizar las acciones pertinentes. Para lograr este fin, una herramienta de cómputo necesita tener una gran cantidad de conocimientos sobre del sistema auditivo humano, sobre la estructura del lenguaje, la representación del significado de los mensajes y sobre todo el auto aprendizaje de la experiencia o el uso diario. Actualmente estamos lejos de lograr un sistema completo que pueda comprender cualquier mensaje oral en cualquier contexto tal y como lo podría hacer un ser humano. Sin embargo, la tecnología

actual sí permite realizar sistemas de reconocimiento del habla que pueden trabajar, con un error aceptable.

Existe gran cantidad de trabajos en donde se recopilan los avances que se han dado dentro del área del reconocimiento automático del habla, ya que no es un problema nuevo, un ejemplo de ello lo tenemos en [2], este trabajo es una recopilación de varios artículos cada uno dedicado al análisis de una rama específica dentro de esta área, componiendo así una colección de documentos dedicados a explicar los avances que se han tenido desde hace algunos años. En [4] se explican los fundamentos necesarios para iniciarse dentro del área del reconocimiento del habla, desde definiciones hasta aplicaciones funcionales; y así como se mencionan estos dos trabajos, existen muchos más que explican el porqué utilizar tecnologías inteligentes y como usarlas para solucionar el problema del reconocimiento del habla. En los siguientes apartados se describe como se ha abordado en este trabajo dicho problema, presentando primero la metodología general que se siguió, posteriormente las conclusiones y resultados obtenidos y al final se lista una serie de trabajos futuros.

2 Metodología.

Las principales etapas de la solución planteada se puede observar en la Fig. 1, cada uno de los paso implica un proceso en sí, en las siguientes secciones se presenta una breve explicación de cada uno de estos.

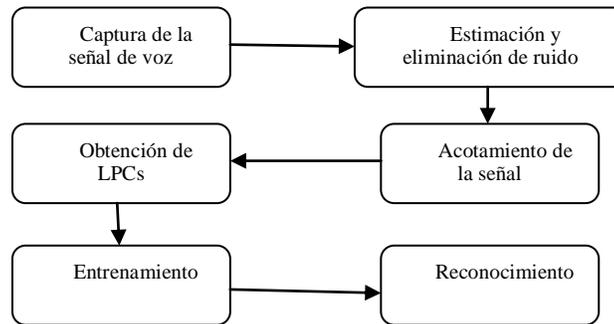


Figura 1. Metodología general.

2.1 Captura de la señal de voz.

Las muestras de voz con las que se estuvo trabajando durante toda la investigación se obtuvieron de dos maneras distintas:

- Base de datos proporcionada por el INAOE. Esta base de datos consta de comandos de voz cortos como “adelante” o “atrás”, así como números pronunciados por hombres y mujeres utilizados en otras áreas del reconocimiento de voz.
- Capturadas directamente con micrófono a través de una laptop Dell studio, igualmente consta de muestras de hombres y mujeres con la diferencia de que se prestó mayor atención en la frase “buenos días”

La frase “buenos días” fue elegida debido a que al contener las cinco vocales nos proporciona mayor información al obtener los componentes de estas señales, las imágenes que se muestran en todo el artículo pertenecen a este tipo de frases Fig. 2.

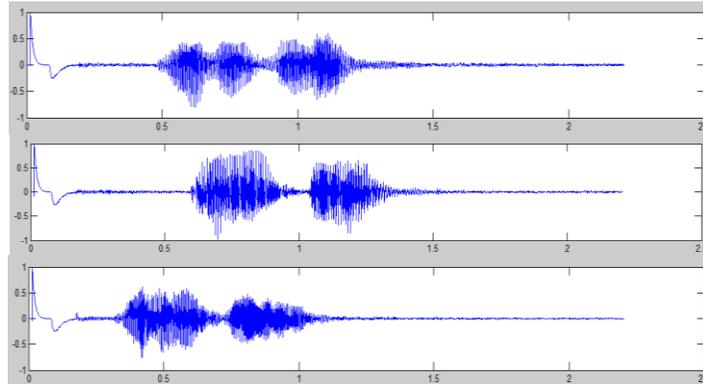


Figura 2. Ejemplo de señales que representan la frase “Buenos días”.

Cabe mencionar que todo el proceso de reconocimiento de voz se realizó mediante el software Matlab lo que facilitó la implementación de algunos algoritmos y sobre todo la graficación de las señales.

2.2 Estimación y eliminación de ruido.

El ruido es uno de los principales problemas con el que nos encontramos al analizar señales de voz, el ruido son básicamente datos irrelevantes que pueden afectar a los datos que si nos interesan y que difícilmente podemos evitar durante la captura de audio, así pues es necesario comprar dispositivos que filtren estos datos innecesarios o bien encontrar algún método que nos permita eliminar o minimizar el ruido después de haber sido capturado.

Intencionalmente todas las muestras con las que se trabajó tienen ruido moderado (también conocido como ruido blanco), esto con la finalidad de que reducir el ruido de la señal sea parte de todo el proceso de reconocimiento

Para eliminar el ruido blanco se analizaron principalmente dos trabajos [5] y [1] en donde se proponen métodos que atenúan este ruido de los cuales se utilizaron dos para realizar pruebas:

- Método de Wiener. Que básicamente hace una estimación del ruido calculado de los momentos de menor intensidad o silencios, este método se aplico aumentando gradualmente el ruido para evaluar su desempeño frente al ocultamiento de los datos relevantes.
- Método de Boll. Éste método busca siempre reducir a cero el ruido estimado dejando únicamente fragmentos de sonido puros.

Una vez que se analizaron ambos métodos se pudo determinar que la mejor solución para eliminar este tipo de ruido es el método de Wiener, ya que se presenta me-

nor pérdida de información aún en sectores donde podría haber datos importantes.

2.3 Acotamiento de la señal.

Antes de poder trabajar realmente sobre los datos de los que se compone cada señal es necesario encontrar el punto exacto donde inicia y termina realmente el sonido que es útil, así un primer paso es eliminar el salto que existe al inicio de cada señal, este salto se origina al encender el micrófono y por lo tanto se considera un ruido constante con lo que basta una resta al inicio de cada señal para eliminarlo, una vez realizado esto las señales nos quedan como se muestra en la Fig. 3.

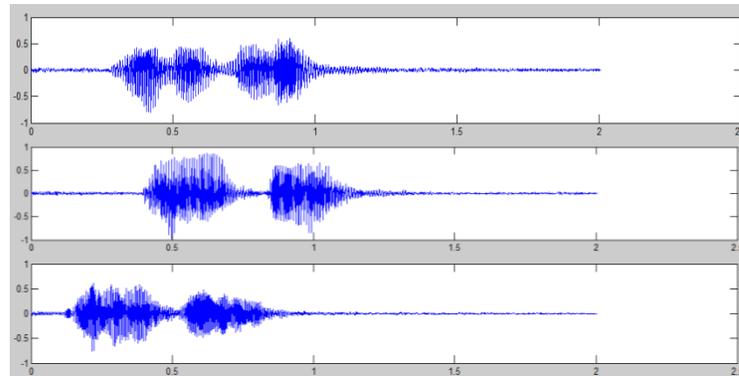


Figura 3. Señal con la frase “Buenos días” después del recorte del salto inicial.

Lo siguiente que hay que realizar es encontrar el punto en donde inicia y termina realmente cada señal de voz, eliminando todo lo demás ya que esto se consideran silencios que no aportan información útil. Para lograr esto se utiliza el cálculo de la energía de una señal, lo que da como resultado un espectro donde se puede notar con mayor facilidad los puntos donde inicia y termina la señal de voz mediante las siguientes formulas:

- La energía debe ser un vector que cumpla la condición de ser mayor a cero y menor a infinito, esto es:

$$0 < E < \infty$$

- La energía se calcula con:

$$E(n) = [(1 - \gamma) * E_{n-1}] + [\gamma * y_n^2]$$

Donde:

E=Es el vector de energía obtenido cuya primera posición vale 0

y=Es el vector de la señal original

γ =Es una constante definida como:

$$\gamma = 1000 / (fr * 16)$$

Una vez calculados estos valores se obtiene gráficamente lo que se muestra en la Fig. 4.

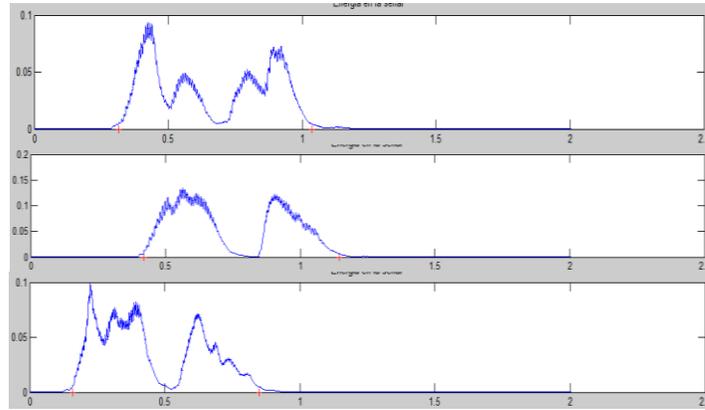


Figura 4. Espectro que muestra la energía de cada señal.

En base a estos datos es posible encontrar más fácilmente a partir de qué puntos se pueden considerar los datos de la señal como útiles o no ya que se logra ver más claramente que los valores innecesarios se aproximan mucho a cero y puesto que la energía de la señal son los valores respectivos a la señal original de voz se puede simplemente recortar los extremos de silencio y reflejarlos directamente en las señales de audio, lo que da como resultado lo que se observa en la Fig. 5.

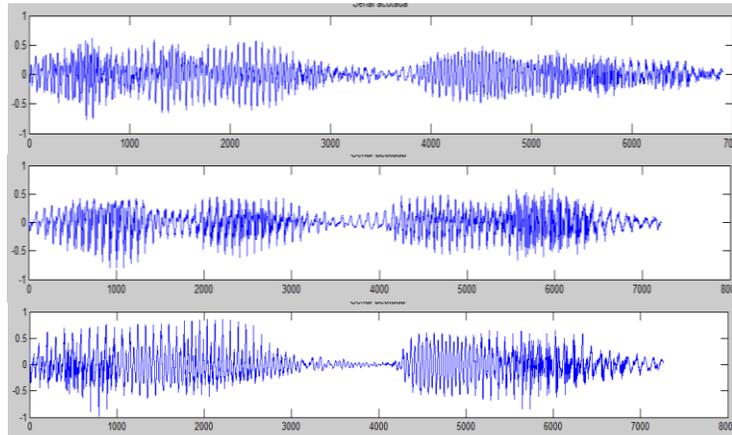


Figura 5. Señales después del corte de los extremos de silencio.

Ya podemos considerar que tenemos de las señales los datos que realmente nos aportan información pero como para cada señal los extremos que se eliminaron fueron cortados de diferentes puntos las señales que ahora tenemos son de diferente tamaño así pues realizamos una normalización en longitud para obtener señales de un mismo tamaño.

Las señales resultantes después de la normalización están compuestas de 6000 muestras cada una, con lo que ya es más sencillo realizar una buena comparación entre señales. Fig. 6.

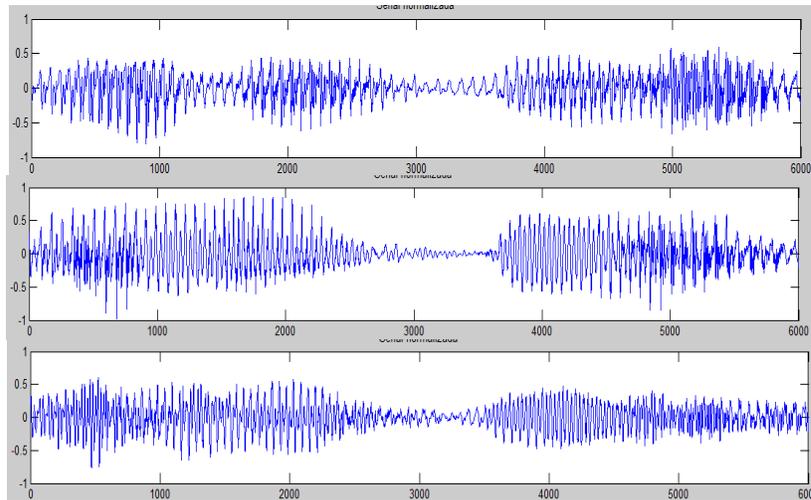


Figura. 6. Señales después de la normalización en longitud.

2.4 Obtención de LPC's.

Muchos trabajos de investigación parten de la idea de que para realizar un buen reconocimiento de voz se deben calcular primero coeficientes LPC de cada señal y posteriormente usando estos, calcular coeficientes cepstrales lo que garantiza un mejor reconocimiento [3], aquí se plantea que es posible realizar de igual manera un buen reconocimiento utilizando únicamente los coeficientes LPC lo que también ahorra tiempo de computo que bien puede eliminarse o utilizarse posteriormente en la fase de entrenamiento y/o reconocimiento.

Es posible calcular un número N de coeficientes LPC para cada señal de voz pero aquí vamos a dividir primero cada señal en varias partes y después calcularemos LPCs para cada parte de cada señal lo que nos da no solo un arreglo de coeficientes sino una matriz. El cálculo de los LPC's, en la implementación de una aplicación debe hacerse en tiempo real.

Para dividir cada señal se utilizar una técnica llamada enventanamiento (Fig. 7), la cual no solo divide la señal en varios fragmentos, sino que la multiplica a través de una función (en este caso es una gaussiana) de tal manera que resalte los elementos que se encuentren más al centro para cada ventana y también utilizaremos un solapamiento de ventanas (Fig. 8) para que los elementos que se encuentren sobre las orillas en una ventana puedan encontrarse en el centro de otras evitando así que perdamos información relevante que pueda estar presente en estos lugares.

Con estos dos métodos obtenemos para cada señal veintiocho ventanas de las cuales calcularemos 20 coeficientes para cada una, esto mediante el comando $A=lpc(x, P)$ en donde x es el número de coeficientes que queremos obtener y P la ventana que está siendo evaluada en ese momento.

Finalmente tenemos para cada señal una matriz de valores que podemos utilizar para entrenar y clasificar nuevas entradas.

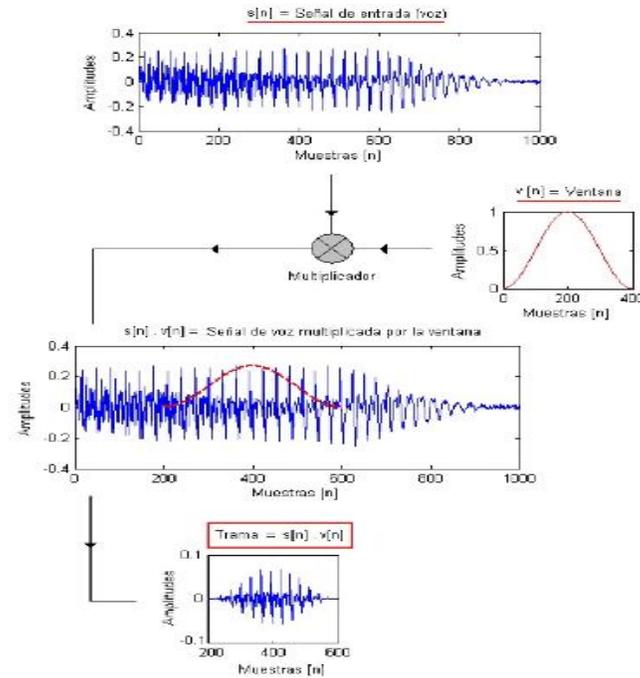


Figura 7. Diagrama representativo de enventanamiento.

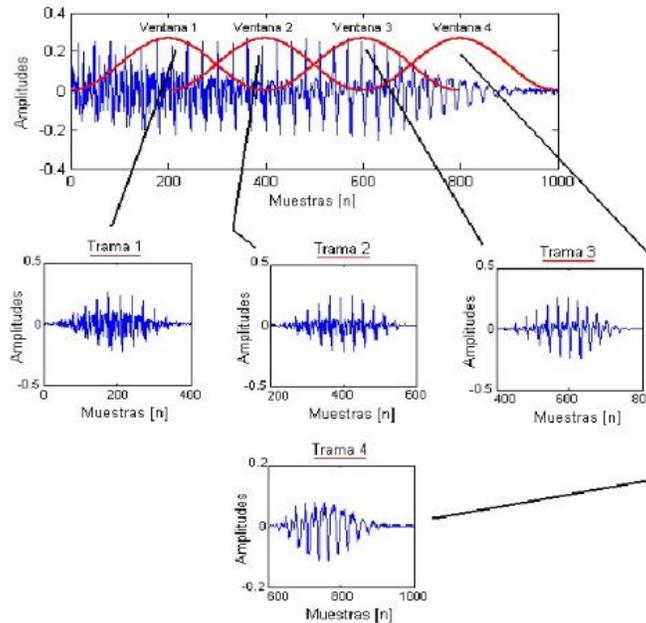


Figura 8. Diagrama representativo de solapamiento.

2.5 Entrenamiento y reconocimiento.

El reconocimiento y entrenamiento son dos etapas que van muy ligadas, es por eso que se utiliza un solo diagrama para explicar estas dos funciones (Fig. 9).

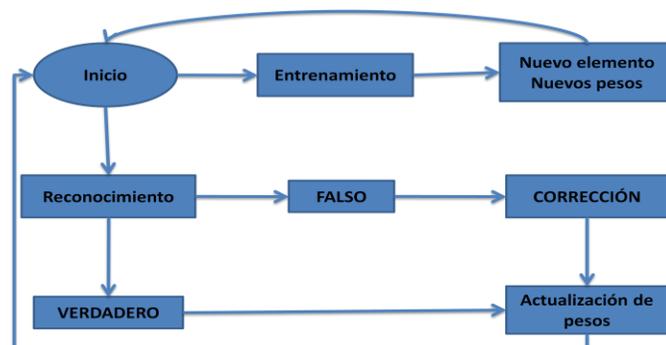


Figura 9. Diagrama general de entrenamiento y reconocimiento.

El entrenamiento se realiza únicamente para elementos nuevos y por lo tanto sus entradas son directas ya que de no ser así no podrá reconocer una nueva entrada que no se encuentre guardada anteriormente, o bien reconocerá a la más parecida cometiendo así un error.

El reconocimiento funciona de la siguiente manera:

- Primero la nueva entrada es descompuesta en una matriz de coeficientes tal y como se describió en secciones anteriores con el fin de tener elementos que comparar.
- La matriz de coeficientes de la nueva entrada es comparada contra las matrices guardadas mediante el algoritmo de KNN obteniendo los cinco elementos que más se parezcan.
- Mediante una probabilidad se elige uno de entre los cinco posibles elementos y este es el que arroja como salida.

Ahora bien una vez elegida la salida esta es evaluada y para ambos casos (Cierto o Falso) la matriz de coeficientes es considerada ahora como una matriz de pesos que se actualizarán tal y como lo harían los pesos de una red neuronal de tal modo que si es correcta la salida sea aún más fácil de identificar la próxima vez y en caso de que sea falsa para que pueda garantizar que la próxima vez acierte.

Resumiendo estos pasos tenemos un algoritmo que para reconocer utiliza: KNN para encontrar elementos parecidos, probabilidad para determinar la mejor solución y actualización de pesos para garantizar un mejor funcionamiento.

3 Conclusiones y Resultados.

Para poder realizar una buena comparación en cuanto a resultados no sólo se utilizó el método propuesto, también se hicieron pruebas con una red neuronal (ART) y con un mapa auto organizado (SOM).

A continuación se muestran los valores obtenidos para la palabra “nueve” del primer grupo de datos y para la frase “buenos días” del segundo conjunto.

TABLA DE RESULTADOS PARA LA PALABRA “NUEVE”

	SOM	ART	KNN
TIEMPO DE ENTRENAMIENTO	0.4 s	0.4 s	0.4 s
TIEMPO DE RECONOCIMIENTO	7 s	0.2 s	0.5 s
NO. DE ITERACIONES PARA RECONOCER CORRECTAMENTE	5	8	3
PORCENTAJE DE ACIERTO EN EL PRIMER INTENTO	60%	70%	85%
PORCENTAJE DE ACIERTO EN EL SEGUNDO INTENTO	70%	75%	92%

TABLA DE RESULTADOS PARA LA FRASE “BUENOS DIAS”

	SOM	ART	KNN
TIEMPO DE ENTRENAMIENTO	0.4 s	0.4 s	0.4 s
TIEMPO DE RECONOCIMIENTO	8 s	0.2s	0.5 s
NO. DE ITERACIONES PARA RECONOCER CORRECTAMENTE	3	5	2
PORCENTAJE DE ACIERTO EN EL PRIMER INTENTO	60%	60%	90%
PORCENTAJE DE ACIERTO EN EL SEGUNDO INTENTO	65%	70%	95%

Con estos resultados se puede observar que el primer método (SOM) tarda demasiado en reconocer una nueva señal, esto es debido a la separación que hace entre todos los datos de entrenamiento y más la competición que se realiza entre los elementos más cercanos al que se busca y aún así no realiza un correcto reconocimiento para ninguno de los dos casos.

En el caso de la ART se observa una disminución considerable en el tiempo de reconocimiento pero un aumento en el número de iteraciones necesarias para el mejor reconocimiento obtenido, que también es mejor en comparación con el obtenido en la prueba anterior.

Finalmente con el método propuesto con KNN se tiene un pequeño aumento en el tiempo de reconocimiento debido principalmente al cálculo de distancias pero que a fin de cuentas arroja un mejor resultado y en menos iteraciones que con cualquiera de los métodos anteriores.

Con todo esto se concluye que es posible realizar un reconocimiento eficiente utilizando únicamente los coeficientes LPCs, sin realizarles ningún otro procedimiento

posterior y que de los métodos propuestos el de KNN con actualización de pesos fue el mejor.

4 Trabajo futuro.

Puesto que la voz no es la única característica que identifica a una persona la metodología que aquí se presento puede ser utilizada para otro tipo de rasgo como podría ser el rostro, la retina, la huella digital, etc; aprovechando el mismo método de reconocimiento que aquí se presento. Por otro lado, siguiendo el área de reconocimiento de hablante es posible que los resultados puedan ser mejorados si las muestras que se adquieren son diferentes para cada persona, logrando una mayor diferencia entre individuos si aparte de las características únicas para cada individuo se toman en cuenta también diferentes frases para cada uno como podría ser el nombre hablado de cada quien.

Referencias

1. G. O. Borrás, "Reductor de ruido mediante resta espectral en entorno Matlab". EUIT Telecomunicación. 2006.
2. K. Daoudi, "State of the art in speech and audio processing". 2004
3. L. M. Hernández, A. E. Simancas, M. M Nakano, M. H. Pérez. "Reconocimiento de hablantes con dependencia del texto basado en LPC-Cepstral y red neuronal de retropropagación". SEPI ESIME Culhuacan, Instituto Politécnico Nacional. 2003.
4. C. A. Reyes, "Conceptos sobre Reconocimiento Automático del Habla". 2005
5. D. R. Tomassi, L. Aronson, C. E. Martínez, D. H. Milone, M. E. Torres y H. L. Rufiner, "Evaluación de técnicas clásicas de reducción de ruido en señales de voz". XV Congreso argentino de bioingeniería. Facultad de Ingeniería Universidad Nacional de Entre Ríos. 2004.

Redes bayesianas aplicadas a un modelo de Dinámica de Fluidos Computacional para las condiciones climáticas al interior de un invernadero

Guillermo Alfonso De la Torre-Gea^{1,2} y Enrique Rico-García²

Universidad Tecnológica de Corregidora, Querétaro, México.

gtorre.utcorregidora@gmail.com

Universidad Autónoma de Querétaro, Querétaro, México.

ricog@uaq.mx

Paper received on 14/06/12, Accepted on 03/09/12.

Resumen. En los últimos años se han realizado estudios para determinar las condiciones del clima al interior de invernaderos empleando la Dinámica de Fluidos Computacional, con lo cual se ha podido aumentar el grado de exactitud y realismo. A pesar de que dichos modelos proporciona una solución numérica del balance energético que permite un estudio eficaz del clima dentro del invernadero, esta técnica toma en cuenta los valores de las variables independientes como incógnitas. Las Redes Bayesianas son técnicas heurísticas que puede ser de ayuda para describir las relaciones entre las variables que definen las condiciones de ventilación. El presente trabajo consistió en desarrollar un modelo de Red bayesiana a partir de datos de radiación solar, velocidad del flujo de aire y temperatura, obtenidos de un modelo de Dinámica de Fluidos Computacional al interior de un invernadero, que nos permitió obtener las probabilidades condicionales de cada una de las relaciones con dichas variables. La red bayesiana mostró que la velocidad del flujo de aire influye más que la radiación solar sobre la temperatura del invernadero.

Palabras clave: Probabilidad, invernaderos, predicción, minería de datos, K2.

1 Introducción.

La ventilación es considerada como uno de los principales procesos para determinar el microclima de un invernadero [2,21]. La ventilación natural es el resultado de la acción de dos fuerzas naturales: la dinámica debida al viento y la térmica o de flotación [4]. La ventilación dinámica se incrementa con el aumento en la velocidad del viento y superficie de ventanas; mientras que la ventilación por efecto de flotabilidad se incrementa con el aumento de las diferencias entre la temperatura del aire interno y externo al invernadero [13]. Una de las ventajas de utilizar la ventilación natural como método de enfriamiento es la reducción del costo de operación de un invernadero [1]. Esto determina que se convierta en uno de los métodos de control del clima más adecuado bajo las condiciones de producción hortícola de bajos recursos.

© J. C. Hernández Hernández, J. F. Ramírez Cruz,
A. Cortés Fernández, J. H. Sossa Azuela. (Eds.)
Advances in Intelligent and Information Technologies.
Research in Computing Science 60, 2012, pp. 87 -96.



El incremento de la temperatura externa en invernaderos con ventilación natural ocasiona que la temperatura interna del mismo supere significativamente su valor óptimo, haciendo esencial la entrada de aire para prevenir daños al cultivo [19]. La ventilación en los invernaderos, influye principalmente en las variables: temperatura, humedad relativa y concentración de CO_2 , lo cual afecta directamente el desarrollo y crecimiento de los cultivos [4]. Un deficiente movimiento de aire provoca zonas con diferencias de temperatura y humedad, lo cual impacta en el rendimiento del cultivo.

En los últimos cinco años se han realizado estudios para determinar las condiciones del clima al interior de invernaderos empleando la Dinámica de Fluidos Computacional (CFD), con lo cual se ha podido aumentar el grado de realismo incorporando modelos 3D, mallas anti insectos y la simulación del cultivo mediante su incorporación como medio poroso, entre otros. Los resultados obtenidos han podido contribuir al mejor entendimiento del fenómeno de la ventilación del invernadero [17].

A pesar de que los modelos CFD proporciona una solución numérica de un balance energético en un volumen controlado, que en comparación con otros métodos y tecnologías costosas permite un estudio eficaz del clima dentro del invernadero, las técnicas de CFD toman en cuenta los valores de las variables independientes como incógnitas principales en un número finito de lugares dentro del dominio, por lo que se hace necesario determinar si existe relación entre los valores de las variables calculadas mediante este método.

Determinar las relaciones entre las variables climáticas al interior de un invernadero con ventilación natural es difícil, debido a la naturaleza inherente estocástica del flujo de aire. Las Redes Bayesianas (RB) son técnicas numéricas de incertidumbre que hacen uso de la inferencia bayesiana como método heurístico que puede ser de ayuda para describir las relaciones entre las variables que definen las condiciones de ventilación.

El objetivo de este estudio es demostrar que las RB pueden ser empleadas para encontrar un modelo que mejor describa las relaciones entre la radiación solar, la temperatura y la velocidad del flujo de aire al interior de un invernadero, mediante datos obtenidos de un modelo CFD, con lo cual realizar el pronóstico de las condiciones internas, calculando las probabilidades condicionales de cada una de las relaciones con dichas variables.

2 Comportamiento del clima al interior del invernadero.

Al interior del invernadero se lleva a cabo una interrelación de variables que definen el microclima, las cuales forman un sistema de dependencias complejo, dichas variables son: Radiación solar, Temperatura, Humedad relativa y CO_2 [3].

2.1 Radiación solar.

La radiación natural producida por el sol es la más importante fuente de energía, la cual varía con la estación y hora del día. Los tres métodos de intercambio de ca-

lor, conducción, convección y radiación contribuyen a las propiedades térmicas del invernadero. La conducción ocurre a través de la cubierta y el piso de invernadero. La convección es generada por diferencias de temperatura entre el aire, el suelo, la cubierta, las plantas así como el exterior del invernadero. La convección es el mecanismo más difícil de cuantificar con precisión ya que los intercambios de calor por convección se basan siempre en relaciones empíricas [6].

En los países con clima cálido e inviernos benignos, el principal problema del uso del invernadero se presenta en el verano, cuando existe una insolación muy alta que provoca elevaciones excesivas en la temperatura, las cuales afectan el desarrollo de los cultivos [19]. La radiación solar influye en el desarrollo de las plantas, ya que es un factor imprescindible en el proceso de fotosíntesis, por lo que su métrica se establece con el término de Radiación Fotosintética Activa (PAR). En períodos de elevada intensidad luminosa acompañada de altas temperaturas produce un activo crecimiento vegetativo. Cuanto mayor sea la luminosidad dentro del invernadero, mayores deben ser las temperaturas y la cantidad de CO₂ para que la PAR sea máxima.

2.2 Temperatura.

La temperatura al interior del invernadero constituye la variable de mayor peso, ya que interviene de manera directa con el crecimiento y desarrollo de las plantas. Aunque cada cultivo tiene un rango de temperatura óptimo para su desarrollo, podemos definir que en su mayoría se encuentran entre 10 a 25°C; dicho rango se encuentra establecido de acuerdo a los siguientes parámetros: [16].

- a) Temperatura mínima letal, que indica daños irreversibles en los cultivos.
- b) Temperaturas máxima y mínima biológicas, dentro de las cuales los cultivos pueden desarrollar sus etapas o fases vegetativas.
- c) Temperaturas nocturna y diurna que indican el rango óptimo para el desarrollo de cada fase de los cultivos.

2.3 Estudios previos sobre la radiación solar y temperatura.

Se han realizado estudios utilizando modelos de radiación solar y transpiración basados en el balance de calor y humedad de los cultivos, para investigar la distribución de temperatura y humedad del aire así como las interacciones con los procesos fisiológicos del cultivo, además de la distribución del flujo de aire. El factor de protección solar juega un papel crucial en el mantenimiento estable de las condiciones térmicas en el interior, incluso si la temperatura exterior es mayor. La temperatura de la cubierta del invernadero es un parámetro esencial necesario para cualquier análisis de la transferencia de energía en el invernadero. De acuerdo con De la Torre-Gea [7], en la Figura 1 se muestra la proporción de variables estudiadas mediante modelos CFD en los últimos años.

El efecto de la radiación solar y térmica a menudo se toma en cuenta mediante el establecimiento de flujos de calor en los límites físicos del invernadero. La transfe-

rencia de radiación dentro del espacio del cultivo son de mayor interés, ya que determinan los dos principales procesos fisiológicos de los cultivos: la transpiración y la fotosíntesis. Por lo que es probable que reciba más atención en los próximos años.

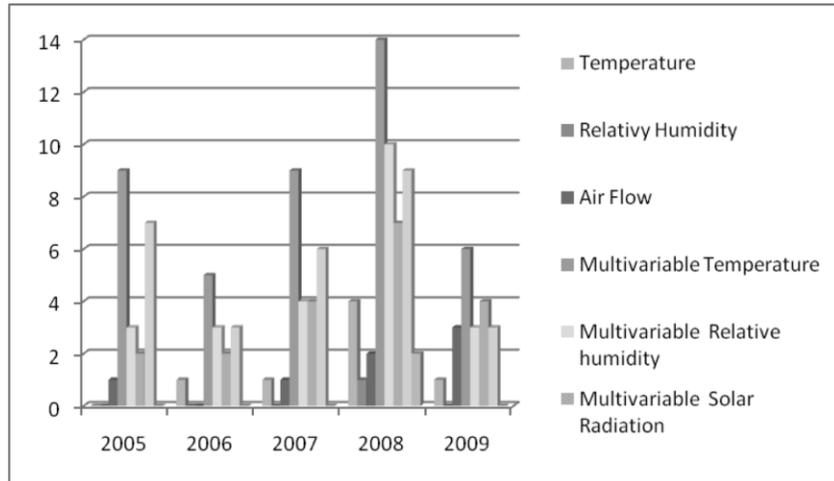


Figura 1. Proporción de las variables estudiadas en modelos CFD.

3 Teoría de las RB.

Las RB son representaciones del conocimiento desarrollados en el campo de la inteligencia artificial para el razonamiento aproximado [20, 18 y 5]. Una RB es un gráfico acíclico cuyos nodos corresponden a conceptos o variables, y cuyos enlaces definen las relaciones o funciones entre dichas variables [3]. Las variables se definen en un dominio discreto o cualitativo, y las relaciones funcionales describen las inferencias causales expresadas en términos de probabilidades condicionales, como se muestra en la ecuación 1.

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{padres}(x_i)) \quad (1)$$

Las RB pueden ser usadas para identificar las relaciones entre las variables anteriormente indeterminadas o para describir y cuantificar estas relaciones, incluso con un conjunto de datos incompletos [13] y [21]. Los algoritmos de solución de las RB permiten el cálculo de la distribución de probabilidad esperada de las variables de salida. El resultado de este cálculo es dependiente de la distribución de probabilidad de las variables de entrada. A nivel mundial, las RB puede ser percibida como una distribución de probabilidades conjunta de una colección de variables aleatorias discretas [9].

La probabilidad a priori $P(c_j)$ es la probabilidad de que una muestra x_i pertenezca a la clase C_j , definida sin ninguna información sobre sus valores característicos,

como se muestra en la Ecuación 2.

$$P(c_j / x_i) = P(x_i / c_j)P(c_j) / \sum P(x_i / c_k)P(c_k) \quad (2)$$

Las máquinas de aprendizaje, en la inteligencia artificial, está estrechamente relacionado con la minería de datos, métodos de clasificación o agrupamiento en estadística, razonamiento inductivo y reconocimiento de patrones. Métodos estadísticos de aprendizaje automático se pueden aplicar al marco de la estadística bayesiana, sin embargo, la máquina de aprendizaje pueden emplear una variedad de técnicas de clasificación para producir otros modelos de RB [19] y [22]. El objetivo del aprendizaje mediante RB es encontrar el arreglo que mejor describa a los datos observados. El número de posibles estructuras de grafos acíclicos directos para la búsqueda es exponencial al número de variables en el dominio, el cual se encuentra definido por la Ecuación 3:

$$f(n) = \sum_{i=1}^n (-1)^{i+1} C^n 2_i^{(n-i)} f(n-i) \quad (3)$$

El algoritmo K2 constituye el método más representativo entre las aproximaciones de “búsqueda y resultado”. El algoritmo comienza asignando a cada variable sin padres. A continuación, agrega de manera incremental los padres a la variable actual que aumenta su puntuación en la estructura resultante. Cuando cualquier adición de una madre soltera no puede aumentar la cuenta, deja de agregar padres a la variable. Desde un pedido de las variables conocido de antemano, el espacio de búsqueda bajo esta restricción es mucho menor que el espacio tomando la estructura entera, y no hay necesidad de comprobar los ciclos en el proceso de aprendizaje. Si el orden de las variables es desconocida, se puede realizar la búsqueda en los ordenamientos [10].

4 Materiales y métodos.

4.1 Delimitación y caracterización de la zona de estudio.

El proyecto se desarrolló en el invernadero experimental Ie-1 (Figura 2) ubicado en el Campus Amazcala, perteneciente a la Facultad de Ingeniería de la Universidad Autónoma de Querétaro, este se encuentra ubicado en la comunidad que lleva el mismo nombre, el cual pertenece al municipio del Marqués cuyas coordenadas geográficas son: latitud norte 20° 42', longitud oeste de 100° 16' y altitud de 1920 msnm.

Geomorfológicamente esta rodeados de sierras, mesetas y lomeríos que impiden el paso de los vientos húmedos del Golfo y la Mesa del Centro, que en sus laderas boreales retiene la humedad de los vientos que viajan de norte a sur. La disposición de esas áreas origina un índice de precipitación bajo en la zona y provoca una oscilación térmica de 7 a 34°C, que determina el carácter extremo de estos climas.

La temperatura media anual varía de 18° a 19°C, la precipitación total anual fluctúa entre 450 y 630 mm. Las áreas donde rige esta variante están situadas en los

alrededores de la ciudad de Querétaro, Villa Corregidora, El Marqués, Santa Rosa de Jáuregui, la localidad Paso de Cuadros y en la colindancia del municipio de Peñamiller con el estado de Guanajuato.

4.2 Infraestructura Disponible.

Invernadero tipo Gótico de 432 m² en 2 naves, cada nave de 9 x 24 m, 4.20 m a canaleta, 6.70 m a la cumbre (2.50 m de cumbre), sin ventanas cenitales sólo laterales de tipo enrollable, de 3 x 9 m a la cara frontal y posterior y de 3 x 16 m a los costados (Figura 2). La orientación es de norte-sur.

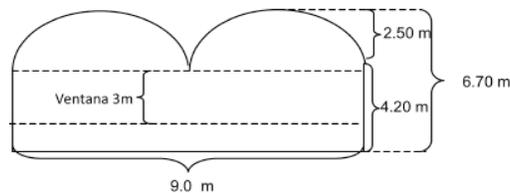


Figura 2. Dimensiones de invernadero experimental Ie-1.



Figura 3. Invernadero experimental Ie-1 ubicado en Amazcala, Querétaro.

4.3 Desarrollo del modelo CFD.

Se realizó el desarrollo y simulación numérica del modelo CFD con el software ANSYS v.12.1 el cual fue ejecutado sobre una máquina de 32 bits bajo el sistema operativo Windows Vista, con lo cual se resolvieron las ecuaciones de energía, continuidad y de momentum, bajo 3 valores diferentes de Radiación solar (500 w, 1000 w y 1350 w) y velocidad del flujo de aire (0 m/s, .45 m/s y .9 m/s).

La metodología para desarrollar el modelo CFD se realizó de acuerdo al modelo propuesto por Rico-García [19], del cual corresponden las siguientes etapas:

- 1) Discretización del flujo continuo: las variables de campo se aproximaron a un número finito de valores en puntos llamados nodos
- 2) Discretización de las ecuaciones de movimiento en función de los valores de los nodos

- 3) Solución del sistema de ecuaciones algebraicas y obtención de los valores de las variables en todos los nodos.

4.4 Análisis mediante Redes Bayesianas.

En análisis de las relaciones entre las variables se realizó mediante el sistema ELVIRA versión 0.162 en tres etapas descritas por De la torre-Gea [8]:

- a) Pre-procesamiento: Se llevó a cabo mediante el algoritmo de imputación “por promedios” para completar las series de datos parciales. Este algoritmo reemplaza los valores faltantes o desconocidos, por el promedio de los valores para cada variable. Este método no necesita parámetros y consiste en la discretización de los datos masivos mediante el algoritmo, empleando dos intervalos con la misma frecuencia.
- b) Procesamiento: Para los fines de este estudio, la mejor estructura de red bayesiana se obtiene empleando el algoritmo K2 learning con un número máximo de padres igual a 3, con estimación de máxima verosimilitud y sin restricciones.
- c) Post-procesamiento: Se realizó un análisis de dependencias para obtener la estructura topológica de la red, la cual representa a las variables y sus dependencias causales. Después de obtener la red de aprendizaje paramétrico, se realizó el cálculo de las probabilidades condicionales en las variables que muestren relación o dependencia.

5 Resultados y discusión.

5.1 Modelo CFD.

Para proporcionar validez al modelo CFD, los resultados obtenidos de la solución de las ecuaciones calculados mediante el software ANSYS, se compararon con los valores de las variables flujo de aire y temperatura medidos al interior del invernadero.

Una parte importante de esta etapa es la representación gráfica de las variables que gobiernan el flujo, para tener una visión rápida y amena de los resultados obtenidos, como se muestra en la Figura 4.

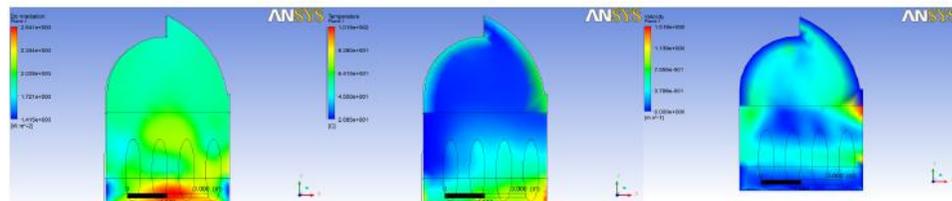


Figura 4. Modelo CFD de la Radiación solar, Temperatura y Velocidad del flujo de aire al interior de un invernadero.

5.2 Modelo de RB.

Se obtuvo un modelo de RB a partir de 4200 registros de datos de temperatura, radiación solar y velocidad del flujo de aire en intervalos de 500 W a 2300 W de radiación solar, 0 a 1.1 m/s de velocidad del flujo de aire y 15°C a 58°C de temperatura, cuyas inferencias se muestra en la Figura 5.

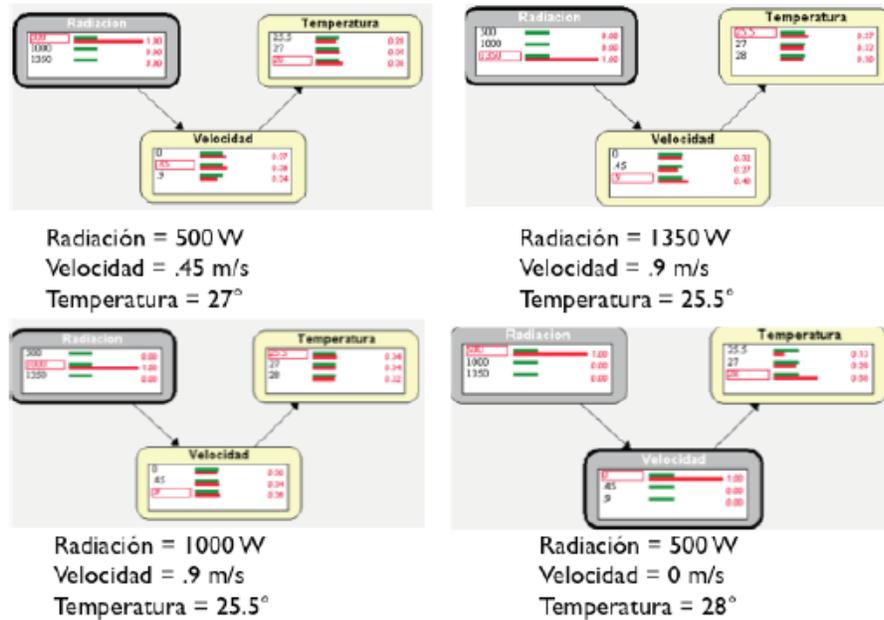


Figura 5. Cálculo de inferencias más importantes sobre el modelo de RB.

En el modelo de RB se muestran las relaciones entre las variables estudiadas, observándose una relación directamente proporcional entre la radiación solar y la velocidad del flujo de aire, así como una relación inversamente proporcional entre la velocidad del flujo de aire y la temperatura. No se observa ninguna relación entre radiación solar y temperatura sugiriendo que la temperatura al interior del invernadero depende sólo del flujo de aire.

Cuando la radiación solar es igual a 500 W, la velocidad del flujo de aire es igual a .45 m/s y la temperatura igual a 28°C con una probabilidad del 36%. Conforme la radiación solar aumenta hasta alcanzar 1350 W, la velocidad del flujo de aire aumenta hasta .9 m/s y la temperatura baja a 25.5°C con probabilidades del 40% y 37% respectivamente. Esto indica que las variaciones en la radiación solar influyen en la velocidad del flujo de aire y ésta a su vez a la temperatura, sin embargo los valores de probabilidad son bajos, por lo que aunque las relaciones entre variables están presentes, la radiación solar no es determinante.

Cuando la velocidad del flujo de aire es igual a cero, la radiación solar es igual a 500 W con una probabilidad del 100% y la temperatura es de 28°C con una probabilidad del 58%. Al incrementar la velocidad del flujo de aire hasta alcanzar .9 m/s, la radiación solar es igual a 1350 W y la temperatura desciende hasta los 15°C con similares valores de probabilidad. Esto indica que la velocidad del flujo de aire es la variable de mayor influencia al interior del invernadero.

Para los alcances de este trabajo se cubierto con el objetivo, sin embargo, es necesario realizar un estudio más detallado tomando en cuenta otros métodos de machine learning que permitan realizar una comparación, así como incorporar otras variables como humedad y concentración de CO₂.

6 Conclusiones.

Utilizando un modelo de RB es posible determinar las relaciones entre las variables radiación solar, temperatura, y velocidad de flujo de aire obtenidas a partir de un modelo CFD validado con mediciones del interior de un invernadero. Las redes bayesianas tienen la capacidad de mostrar las relaciones entre las variables de las condiciones climáticas al interior de un invernadero. Las distribuciones de probabilidad condicionales son una medida cuantitativa de las relaciones entre las variables y muestran el estado más probable de estas variables. El modelo de RB mostró que la variable de mayor influencia es la velocidad del flujo de aire, mientras que la radiación solar no mostró influencia significativa sobre la temperatura al interior del invernadero. Este método puede ser empleado en estudios posteriores para validar modelos de CFD.

Referencias

1. Baptista, F.J.; Bailey, B.J.; Meneses, J.F. 2001. Natural Ventilation Of Greenhouses: Comparison Of Measured And Predicted Ventilation Rates. *Agribuilding*, 136–151.
2. Bouchet, E.R.; Freyre, C.E.; Bouzo, C.A. 2003. Medición Y Predicción De La Tasa De Renovación Del Aire En Un Invernadero Con Ventilación Lateral. Ponencia Oral. Actas Del Xxvi Congreso Argentino De Horticultura, Paraná, Entre Ríos. 7 P.
3. Castilla Prados N. 2007. Invernaderos De Plástico. Tecnología Y Manejo. (Ed.) Mundi-Prensa, 2ª Ed.
4. Coelho, M.; Baptista, F.; Fitas Da Cruz, V.; Garcia, J.L. 2006. Comparison Of Four Natural Ventilation Systems In A Mediterranean Greenhouse. *Acta Hort.* 719:157-171.
5. Correa M, Bielza C, Paines-Teixeira J, Alique JR (2009). Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process. *Expert Syst. Appl.*, 36(3): 7270-7279.
6. Critten, D.L., Bailey, B.J. 2002. A Review Of Greenhouse Engineering Developments During The 1990s. *Agricultural And Forest Meteorology* 112: 1-22.
7. De la torre-Gea, G., Soto-Zarazúa, G.M., López-Crúz, I., Torres-Pacheco, I., Rico-García, E. (2011b). Computational fluid dynamics in greenhouses: A review. *African Journal of Biotechnology*, 10(77), 17651-17662.
8. De la torre-Gea G., Delfin-Santisteban, O. 2012. Análisis probabilístico mediante redes bayesianas para el cálculo de la probabilidad de lluvia en diferentes localidades. *Research in Computing Science*, 55, 249 – 258.

9. Gámez JA, Mateo JL, Puerta JM (2011). Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Min. Knowl. Discov.*, 22: 106-148.
10. Garrote L, Molina M, Mediero L (2007). Probabilistic Forecasts Using Bayesian Networks Calibrated with Deterministic Rainfall-Runoff Models. In Vasiliev et al. (eds.), *Extreme Hydrological Events: New Concepts for Security*, Springer, pp. 173-183.
11. Guoliang L (2009). Knowledge Discovery with Bayesian Networks. PhD dissertation, National University of Singapore, Singapore.
12. Hruschka E, Hruschka E, Ebecken NFF (2007). Bayesian networks for imputation in classification Problems. *J. Intell. Inform. Syst.*, 29: 231-252.
13. Katsoulas, N.; Bartzanas, T.; Boulard, T.; Mermier, M.; Kittas, C. 2006. Effect Of Vent Openings And Insect Screens On Greenhouse Ventilation, *Biosystems Engineering*, 93: 427-436.
14. Mediero OL (2007). Probabilistic forecast flood flows Through Bayesian Networks Applied to a Distributed Hydrological Model. PhD dissertation, Polytechnic University of Madrid, Madrid, Spain.
15. Naveed N, Choi MTS, Jaffar A (2011). Malignancy and abnormality detection of mammograms using DWT features and ensembling of classifiers. *Int. J. Phy. Sci.*, 6(8): 2107-2116.
16. Nisen A., Grafiadellis M., Jimenez R., Malfa G. La, Martinez García P.F., Monteiro A., Verlodt H., Villele O. de, Zabeltitz C.H. von, Denis I.Cl. y Baudoin W.O., 1988. Cultures protégées en climat méditerranéen. FAO. N° 90. Roma, 317 pp.
17. Norton, T., Sun, D.W. , Grant, J., Fallon, R., Dodd, V., 2007. Applications of Computational Fluid Dynamics (CFD) in the modeling and design of ventilation systems in the agricultural industry: a review. *Bioresource Technology* 98 2386–2414.
18. Reyes P (2010). Bayesian networks for setting genetic algorithm parameters used in problems of geometric constraint satisfaction. *Intell. Artificial.*, 45: 5-8.
19. Rico-García, E. 2008 Modelación Climática En Invernaderos: Ventilación Natural. Tesis Doctoral. Facultad De Ingeniería. Universidad Autónoma De Querétaro, Pp 74:97.
20. Subramaniam T, Jalab HA, Taqa AY (2010). Overview of textual antisipam filtering techniques. *Int. J. Phys. Sci.*, 5(12): 1869-1882.
21. Teitel, M.; Tanny, J.; Ben-Yakir, D.; Barak, M. 2005. Airflow Patterns Through Roof Openings Of A Naturally Ventilated Greenhouse And Their Effect On Insect Penetration, *Biosystems Engineering* 92:341-353.

Localización y Reconocimiento de Obstáculos de un Robot Móvil Mediante SLAM-R

Sergio Díaz-Zagal, Citlalih Gutiérrez-Estrada, José Rafael Cruz-Reyes
Rafael Lemus-González y Juan José Rivera-Alemán

Instituto Tecnológico de Toluca, Toluca, Edo de Méx., México
sdiaz@ittoluca.edu.mx, citlalihg@ittoluca.edu.mx

Paper received on 23/07/12, Accepted on 27/08/12.

Resumen. Para un robot móvil, la capacidad de localizarse en un entorno y mapearlo es fundamental, lo cual los hace realmente autónomos, capaces de adaptarse a las circunstancias y al medio que los rodea. La autonomía y la adaptación de los robots móviles a su entorno es el problema principal a resolver. Estos robots autónomos no solo se aplicarían en la industria, sino se emplearía para el apoyo de los seres humanos en sus diversas tareas. El presente artículo describe un sistema de simulación, para la localización y reconocimiento de obstáculos, utilizando un algoritmo de localización y mapeo simultáneos (SLAM) en un robot móvil con las características del Pioneer 3-AT, mediante técnicas de medición láser y un identificador de radiofrecuencia (RFID), para el reconocimiento de objetos. La integración del RFID y la técnica de navegación como resultado permitió disminuir el error de la posición de un robot móvil en un espacio interno, al mismo tiempo identificar lugares y objetos.

Palabras Clave: SLAM-R, RFID, Marilou, Telémetro Láser.

1 Introducción.

Una característica fundamental que debe tener un robot autónomo, utilizando la técnica del SLAM (Simultaneous Localization and Mapping), es la capacidad de explorar el entorno y construir un mapa del mismo entorno. En la técnica del SLAM los sensores más utilizados por muchos autores son los sensores SONAR [1] o LASER para 2D [2, 3] y 3D [4]. Por otra parte, también existe un gran interés por el uso de cámaras como sensores, técnica conocida como SLAM visual. Una de las ventajas que ofrece las cámaras, es la mayor información acerca del entorno, además que son menos costosas que los sensores previamente mencionados. Gran parte de la técnica SLAM visual, funcionan con mapas basados en características (landmark-based maps). Estos mapas representan la localización de un conjunto de puntos distintivos del entorno, respecto a un sistema de referencia global [5, 6].

En el estado del arte, los mapas “preconstruidos” e incluidos en los robots previamente, resuelven el problema de la localización de un robot móvil, actualmente estos mapas son elaborados a medida que el robot se va desplazando.

La pieza clave de la autonomía completa de un robot móvil, está en la búsqueda de la solución al problema del SLAM, lo que constituye el principal esfuerzo de los investigadores de robótica móvil en todo el mundo, sin embargo, actualmente se han conseguido avances importantes, pero se continúa buscando la mejor solución.

En la localización y la construcción de los mapas, el principal problema se deriva de la existencia del ruido y de las limitaciones en el rango de las características de los sensores. No obstante, a través de la herramienta computacional se ha tratado de corregir el problema del SLAM (y se continúa investigando) desde diversos puntos de vista.

Los algoritmos más utilizados en la robótica móvil actualmente en el campo del SLAM son: el SLAM-EKF (Filtro Extendido de Kalman), el basado en representaciones gráficas, aplica satisfactoriamente métodos de optimización no lineal (sparse nonlinear optimization methods) y el de técnicas de filtrado estadístico no paramétrico, conocidas como filtros de [7].

Una limitación importante que tiene el algoritmo SLAM-EKF, es el coste computacional, haciéndolo difícilmente escalable. Sin embargo, la construcción de sub-mapas son las versiones más prometedoras del SLAM-EKF, aunque los algoritmos resultantes se parecen demasiado a los basados en grafos.

El problema del SLAM completo, de naturaleza offline, es resuelto por los métodos basados en grafos. La estimación la realizan desde la observación, por ello el problema de SLAM puede ser modelado como un grafo no-denso de restricciones débiles, donde cada restricción corresponde a un movimiento o a la observación de un landmark. También hay versiones online que suelen eliminar del grafo las posiciones antiguas del robot.

Dentro del método que usa técnicas de filtrado estadístico no paramétrico, se encuentran las conocidas como filtro de partículas [7], dejan de lado la relación entre los objetos del mapa y el robot, representándolo como independiente. El algoritmo de FastSLAM computacionalmente presenta algunas ventajas, puede ser usado como un filtro, y su actualización requiere un tiempo lineal/logarítmico, mientras que el EKF necesita un tiempo cuadrático. Además, FastSLAM se salta el paso de asociación de datos (Porque no lo necesita), lo que lo convierte en un algoritmo ideal en casos en los que se desconoce cómo realizar asociaciones. Sin embargo, el número de partículas necesarias puede crecer en exceso, especialmente en robots que buscan realizar mapas de múltiples trayectorias cerradas.

Este artículo se centra en la implementación de un algoritmo de navegación en un robot móvil y su evaluación, a través de un estudio formal de simulación, que integra a la vez procesos de la Ingeniería del Software [8], en las etapas de obtención de requerimientos, programación y control de versiones; además de emplear el modelo de ciclo de vida conocido como Espiral, modelo que no se describió, debido a que no es el objetivo de este artículo, sin embargo, se utilizó para el diseño del prototipo del robot móvil, así mismo se empleó la ingeniería de software para darle la formalidad al trabajo, como el análisis de requerimientos, gestión y desarrollo del proyecto.

2 Diseño del prototipo virtual.

El modelado, el control así como la simulación, han sido desarrollados empleando el software de simulación Anycode Marilou Robotics 2010 [9], para el diseño de un robot móvil virtual, incluyendo un telémetro láser Hokuyo URG-04X [10] y un lector RFID ID-12 [11] de Innovations. Marilou Robotics una importante herramienta del software para el estudio, la creación y la evaluación de sistemas robóticos de alto desempeño.

Recordando que un modelo es una representación de un objeto o de un fenómeno [12], dentro del proyecto se busca modelar un robot móvil para lo cual es necesario primeramente obtener las dimensiones físicas del prototipo real tomadas de los requerimientos funcionales y no funcionales anteriormente mencionados, para poderlas plasmar de manera fidedigna con el software de diseño Marilou Robotics 2010 [13].

A continuación se detalla el modelo de un robot móvil, se deben integrar en el modelo los requerimientos funcionales, para toma de datos y control, que representen los sensores y actuadores que implementa el prototipo real, finalmente debido a que el tipo de simulación que se hace la cual es a través de una animación en 3D que toma en cuenta características de cinemática y dinámica, dentro del software de diseño, fue necesario establecer las restricciones físicas del modelo animado en el escenario 3D como son la gravedad, el peso y otras fuerzas que interactuaron sobre él. Este proceso se muestra en la Figura 1.

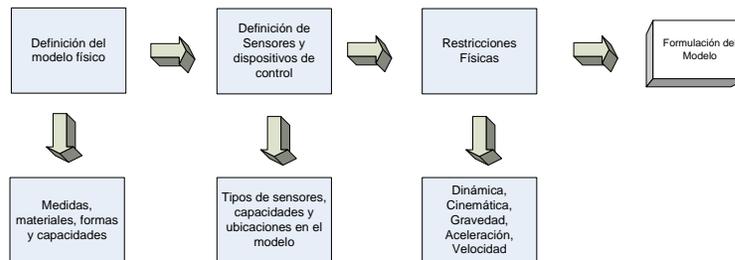


Figura 1. Proceso de Diseño del Modelo.

Con las definiciones y Especificación de Requerimientos, se logró el rediseño del algoritmo de navegación autónoma adaptado a escenarios con gran cantidad de obstáculos, se integró en el diseño el sensor laser y se configuró para realizar un barrido axial suficientemente amplio para implementar el algoritmo de mapeo (SLAM).

El prototipo final del robot móvil con las características del Pioneer 3-AT [14], se muestra en la Figura 2.

La evaluación del prototipo virtual presentó una similitud con el modelo real del 95% en características visuales y de rendimiento contra el modelo real propuesto. El algoritmo de mapeo se implementó con resultados óptimos en escenarios sin gran cantidad de obstáculos. Finalmente el algoritmo de navegación autónoma se adaptó mejor en escenarios con gran cantidad de obstáculos.

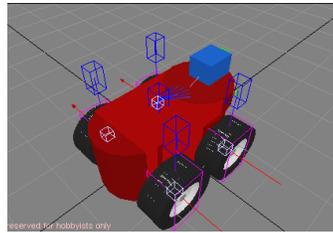


Figura 2. Prototipo Desarrollado.

2.1 Programación.

La existencia de robots que realicen de manera autónoma tareas de modo eficiente, depende fundamentalmente de su construcción mecánica y de su programación. Una vez construido el cuerpo mecánico del robot, conseguir que realice una tarea se convierte en la práctica en un problema de programación [15]. La generación de comportamiento en un robot consiste entonces en escribir el programa, que al ejecutarse en el robot, causa ese comportamiento cuando éste se encuentra en cierto entorno.

La autonomía y la “inteligencia” residen en el programa desarrollado. Por ejemplo, en los robots móviles el comportamiento principal se centra en su movimiento. Los programas que se ejecutan en el robot determinan cómo se mueve éste por su entorno, reaccionando ante obstáculos percibidos por los sensores, acercándose a algún destino, etc. Y para ello tienen que enviar continuamente las órdenes pertinentes a los motores. Este proceso se describe en la Figura 3.

El software de simulación Marilou Robotics 2010 [13], a través de su módulo MODA (Marilou Open Device Access), Permitió la interacción directa entre la lógica del programa y los sensores que éste controla, gracias a la integración directa entre los nombres de los dispositivos y su asociación con variables definidas dentro del programa que comunican la información entre sus dispositivos en tiempo real, durante la simulación.

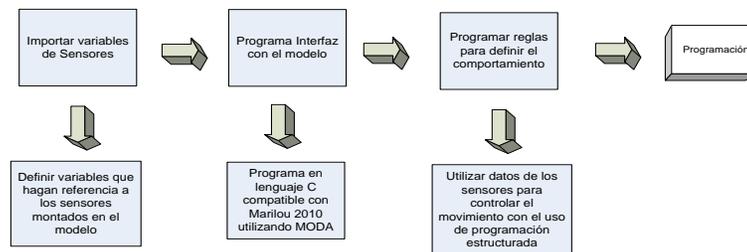


Figura 3. Proceso de Programación.

Una de las aportaciones del proyecto consistió en el empleo de técnicas, hasta ahora solo utilizadas en gestión de proyectos de software, a un proyecto de robótica, que además emplea también procesos de simulación. Con estas técnicas fue posible

formular y evaluar cinco prototipos distintos de robots móviles; las características en común de estos prototipos fue la inclusión de sensores ultrasónicos y laser, motores independientes en cada rueda y un modelado siempre emparentado con modelos físicos reales. Las características cambiantes de estos prototipos fueron sus configuraciones móviles, entre las cuales se probaron, principalmente las de tipo diferencial, con una rueda loca atrás o una combinación de dos diferenciales, una al frente y otra atrás; otras características que se modificaron frecuentemente fueron las relativas al software, aplicado a los algoritmos de navegación autónoma y mapeo implementados en la lógica del robot móvil.

Actualmente se han realizado otros proyectos con el software Marilou Robotics 2010 [13], en donde la simulación y la robótica van de la mano para evaluar y experimentar nuevas técnicas de navegación, modelos de robots y escenarios, en los que se pretende probar posteriormente, un sistema o conjunto de sistemas, como el trabajo presentado por Marcela Urdaniz [16], donde se describe una simulación realizada con el mismo software aquí descrito. Sin embargo poco se menciona acerca de la gestión del proyecto, la obtención de requerimientos, la forma en cómo se validaron los resultados y el control de versiones que se utilizó durante el proyecto.

En los últimos años ha crecido el interés por los robots móviles con capacidad de identificar su entorno (localización), y con capacidad de reaccionar respecto al mismo, sin la necesidad de intervención humana. El proceso de localización puede ser dividido de forma general en seguimiento de la posición, localización basada en mapas y SLAM (Simultaneous Localization And Mapping) [17]. Cada uno de estos niveles constituye un amplio campo de investigación que se ha extendido en función a la manera como la comunidad científica ha enfocado el problema. Usualmente los robots diseñados para operar en entornos interiores cuentan con sensores, que les permiten determinar la posición inicial para que el robot pueda estimar su posición a través del tiempo. Sin embargo, la estimación del proceso de localización se ve afectado por errores y factores del entorno que producen que el error de la posición estimada y la posición real del robot crezca con el paso del tiempo [18].

3 Método.

La navegación del robot móvil empleó el algoritmo fastSLAM que trabaja mediante filtros de partículas, para la estimación se implementó un sistema híbrido donde únicamente se utilizó la posición anterior $St-1$ y la actual St para determinar la posible posición del robot $St+1$ almacenando el mapa durante todo el periodo de la exploración. Este modelo permite que el procesamiento de los datos sea más rápido y se pueda generar un mapa en tiempo real (online), una de las modificaciones que se realizaron dentro del algoritmo fastSLAM es la inclusión de la variable Rt que aporta una nueva solución al problema de la asociación de datos mediante las etiquetas RFID usadas como referencias (Figura 4). La característica principal de los sistemas de navegación se basa en la dependencia de los datos que los sensores generan. Por tanto, si los sensores emiten lecturas erróneas, la posición estimada tendrá un alto grado de error. Debido a este problema, se propuso incluir un dispositivo auxiliar (RFID) para aumentar la fiabilidad de los datos del entorno.

El algoritmo es el encargado de estimar la posición actual del robot en un entorno semi-dinámico y utiliza las etiquetas RFID como información adicional para reconocer objetos dentro del mapa mediante un identificador único (ID).

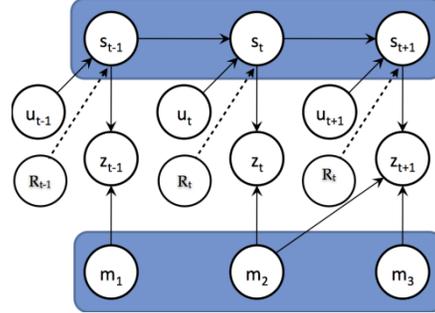


Figura 4. Red Bayesiana Dinámica SLAM-R.

En la Figura 4 se muestra el modelo de red bayesiano dinámico utilizado para representar las conexiones que existen entre el mapa del entorno (M), los datos del sensor láser (Z), el uso de las etiquetas RFID como referencia (R), la posición del robot (S) y la odometría (U) todos ellos dispuestos en un determinado tiempo (t). Para realizar la estimación de la posición del robot es necesario implementar la etapa de la predicción (Ec 1):

$$S_t^{[m]} \sim p(S_t | S_{t-1}^{[m]}, U_t, R_t) \quad \text{Ec 1}$$

Y posteriormente la de actualización de pesos y re muestreo (Ec 2),

$$W_t^{[m]} = \frac{p(S_t^{[m]} | Z^t, U^t, R^t)}{p(S_t^{[m]} | Z^{t-1}, U^{t-1}, R^{t-1})} \quad \text{Ec 2}$$

Además de emplear el uso del RFID para la identificación de los objetos del entorno, se mejoró el algoritmo de navegación, la identificación de los objetos del entorno se realiza mediante el ID, que contienen cada una de las etiquetas, registrando cada una de ellas en una base de datos, donde se registran las características físicas del objeto, para posteriormente incluirlo de forma virtual dentro del mapa generado y obtener una mejor representación gráfica del entorno real (Figura 5).

Para realizar las pruebas del funcionamiento del algoritmo propuesto, se programó en C++, por ser uno de los lenguajes que cuenta con soporte para implementar librerías SDL (Simple Directmedia Layer), que permiten la creación de gráficos en 2D y la generación del mapa. Como se mencionó anteriormente, para la simulación se utilizó el software Marilou, que permite modelar el funcionamiento de robots, dispositivos electrónicos y entornos para crear un ambiente similar al real, en el cual se diseñó el robot móvil basado en las características físicas del robot Pioneer 3AT y de los sensores que éste incluye adicionando el telemetro láser Hokuyo URG04LX. Por otro lado, también se modelaron diversos entornos estáticos con y sin objetos para analizar los posibles cambios en el sistema de navegación.

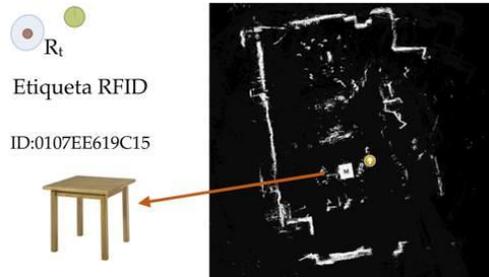


Figura 5. Reconocimiento de obstáculos mediante RFID.

Para la conexión de ambos sistemas (Algoritmo + Simulador) fue necesario utilizar un conjunto de librerías MODA (Marilou Open Devices Access), suministradas por el software Marilou, que sirven para utilizar métodos y tipos de datos que manipulan el funcionamiento de los dispositivos electrónicos simulados. Gracias a este conjunto de librerías, se logró la comunicación de ambos sistemas para transmitir la información del telémetro láser hacia el algoritmo de navegación. Uno de los problemas fue incluir el dispositivo RFID, debido a que el simulador Marilou no cuenta con dicho dispositivo, por tal motivo, se utilizó un puerto virtual para conectar el dispositivo RFID físico mediante el puerto serial de la computadora (Figura 6).

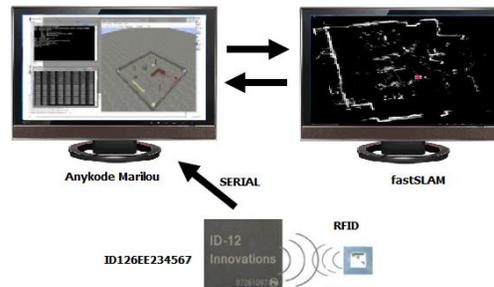


Figura 6. Sistema de interconexión entre el algoritmo de navegación y el simulador Marilou.

4 Pruebas y resultados.

El proceso de experimentación del proyecto consistió en el diseño de escenarios, las restricciones impuestas a éstos y el montaje del modelo del robot móvil dentro del escenario.

El diseño del escenario se realizó con el mismo software con el que se modela el robot Marilou Robotics2010 [13], la ventaja de este software fue la adaptación a las condiciones físicas como gravedad, masa de los objetos inmersos en el escenario, fuerzas que pueden inferir sobre el modelo, etc. Todas estas condiciones pueden ser definidas y cambiadas para observar el comportamiento del modelo dentro del escenario y corregir anomalías del sistema de navegación, que reside en la

programación del robot. Este proceso se detalla en la Figura 7.

La evaluación de los resultados obtenidos se logró mediante la generación de un mapa virtual con diversos objetos y se comparó el mapa virtual generado por el algoritmo de navegación, determinando la similitud entre ambos.

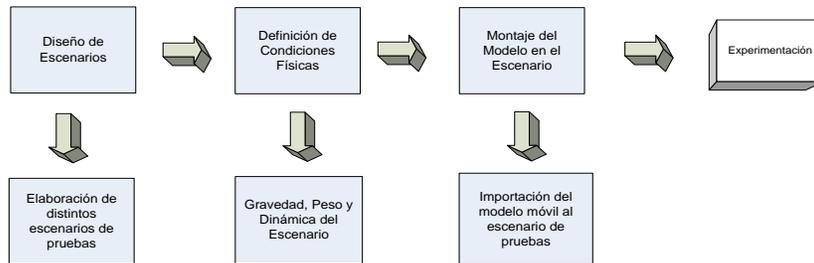


Figura 7. Proceso de Experimentación.

Se realizaron pruebas con 3 entornos diferentes dando como resultado una similitud de 95%, comparada con el mapa virtual, quedando únicamente incompletos los objetos no etiquetados mediante RFID. Así mismo para probar el algoritmo de SLAM, se utilizaron archivos repositorios (DataSet) internacionales, como dataset Málaga 2006 [19] o como los del laboratorio de robótica de mines paristech [20].

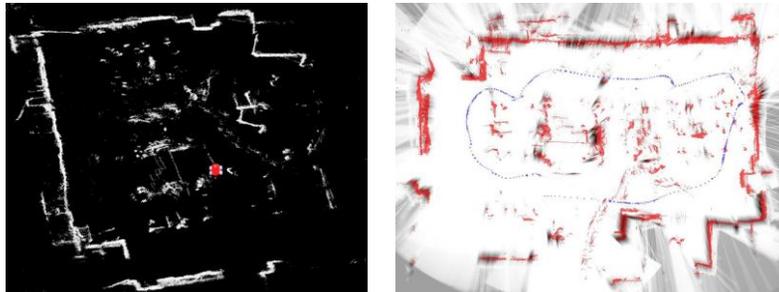
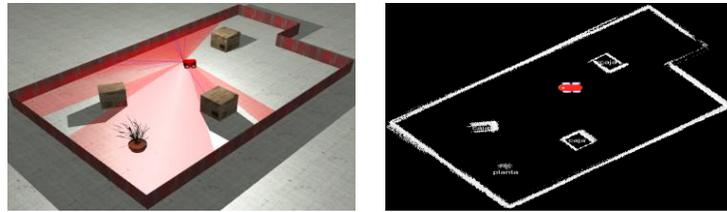


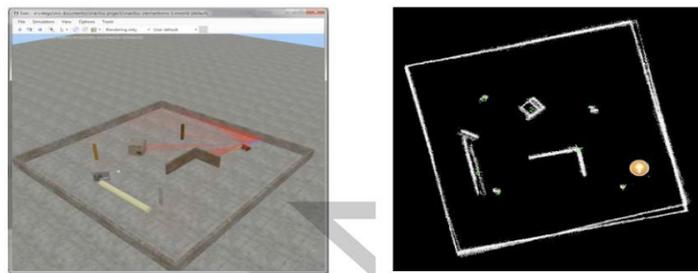
Figura 8. Resultado del algoritmo SLAM sin utilizar el RFID sin objetos en el entorno, utilizando el repositorio del laboratorio de robótica de mines paristech.

Una de las limitaciones del algoritmo propuesto fue la dependencia de las etiquetas RFID que fueron colocadas en el entorno para la identificación de objetos, sin embargo, esto no implicó que el algoritmo no pudiera continuar con la navegación ya que, en caso de que un objeto no se encuentre etiquetado, éste será representado como un obstáculo, sin las características propias del mismo, que le indicará al robot que no puede desplazarse por dicho lugar. La Figura 9 se muestra un entorno virtual sin objetos, con la finalidad de comprobar la independencia del algoritmo SLAM con el sistema de RFID.

Posteriormente se diseñó un entorno con diversos obstáculos para verificar si el algoritmo era capaz de interpretar las etiquetas y reconocer dichos objetos durante la navegación. En la Figura 10 se muestran los resultados de la prueba.



(a) Entorno virtual marilou (b) Mapa generado por el algoritmo SLAM-R
Figura 9. Resultados del algoritmo propuesto SLAM-R sin objetos en el entorno.



(a) Entorno virtual marilou (b) Mapa generado por el algoritmo SLAM-R
Figura 10. Resultados del algoritmo propuesto SLAM-R con objetos en el entorno y etiquetas RFID.

5 Conclusiones.

De acuerdo a los resultados obtenidos en las pruebas simuladas, se concluyó que las etiquetas RFID en un entorno interior son de gran utilidad para la identificación de objetos y para disminuir el error en la estimación de la posición del robot dentro del mapa, sin embargo, se recomienda etiquetar únicamente los objetos de interés necesarios para realizar determinadas tareas (ubicación de objetos, transporte de materiales, reconocimiento de áreas, etc.), ya que etiquetar todos los objetos del entorno, puede convertirse en proceso complicado y tedioso. Por otro lado, es importante mencionar que el trabajo realizado se probó únicamente en un modelo virtual (simulación), sin embargo, uno de los objetivos es comprobar que los resultados obtenidos corresponden al comportamiento de un modelo físico.

La utilización del modelo en espiral como respuesta a la necesidad de formalizar la gestión del proyecto, tuvo como principales objetivos, mejorar el análisis de la información, apoyar a la toma de decisiones, la re-utilización del código de programación, la evaluación y validación de cada uno de los prototipos; así como la reformulación de nuevos requerimientos que se adaptaron mejor a cumplir el propósito principal del proyecto; el cual estuvo enfocado al estudio e implementación de un algoritmo de navegación autónoma para un robot móvil, así como un algoritmo de mapeo y su representación gráfica en tiempo real de la simulación; con todas las aplicaciones que el empleo de estas técnicas pudieron aportar para proyectos futuros en forma de componente o como base de investigación para desarrollar otros proyec-

tos de robótica con ésta misma gestión, enfocados al aprovechamiento de la simulación como una técnica precedente a la manufactura eficiente de robots móviles y el ahorro de costos

Un objetivo más de la investigación como trabajo futuro es considerar la navegación del robot tanto en entornos interiores como exteriores.

Agradecimientos

Agradecemos el apoyo a la DGEST por aprobar el proyecto “Prototipo de un sistema virtual de un par de mini-actuadores de músculos artificiales neumáticos en antagonismo basados en el modelo de McKibben”, con clave TOL-DCIET-2010-102, de la convocatoria “Registro de proyectos 2010”. De igual manera agradecemos el apoyo a la DGEST por aprobar el proyecto “La Ingeniería de Sistemas como modelo integral para el desarrollo de aplicaciones adaptativas”, con clave 4389.11-P, que ha servido como base para continuar en investigaciones relacionadas a este proyecto.

Referencias

1. O. Wijk and H. I. Christensen. Localization and navigation of a mobile robot using natural point landmark extracted from sonar data. In *Robotics and Autonomous Systems*, 1(31), pp. 31-42, 2000.
2. J.J. Leonard and H.F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(4), 1991.
3. S. Thrun. A probabilistic online mapping algorithm for teams of mobile robots. In *Int. Journal of Robotics Research*, 20(5), pp. 335-363, 2001.
4. R. Triebel and W. Burgard. Improving simultaneous mapping and localization. In *Proc. Of the National Conference on Artificial Intelligence (AAAI)*, 2005.
5. J. Little, S. Se, and D.G. Lowe. Global localization using distinctive visual features. In *IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 2002.
6. A. Gil, O. Reinoso, W. Burgard, C. Stachniss, and O. Martínez Mozos. Improving data association in rao-blackwellized visual SLAM. In *IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 2006.
7. J. A., “Sistema de navegación y modelado del entorno para un robot móvil”, noviembre 2006, UPM Autonomous Systems Laboratory.
8. Ian Sommerville, *Ingeniería del Software, “Ingeniería de Requerimientos”*, (Pearson, 7ª Ed., México), pags. 105-150 (2005).
9. AnyKode, “Marilou, simulate your robotics applications.” <http://www.anykode.com/index.php>, 2012.
10. Hokuyo, “Hokuyo urg-04lx datasheet.” http://www.hokuyo-aut.jp/02sensor/07scanner/download/data/URG-04LX_spec.pdf, Febrero 2012.
11. Sparkfun, “Rfid innovations id-12 datasheet.” <http://www.sparkfun.com/datasheets/Sensors/ID-12-Datasheet.pdf>, Febrero 2012.
12. Robert Shannon; Johannes J. D. *Systems simulation: the art and science*, (*IEEE Transactions on Systems, Man and Cybernetics*. Vol. 6(10), Inglaterra)pp. 723-724(1999).
13. Marilou Robotics Studio, anyKode, Francia2008-2010, www.anykode.com.
14. Adept, “Pioneer 3-at datasheet.” <http://www.mobilerobots.com/Libraries/Downloads/Pioneer3AT-P3AT-RevA.sflb.ashx>, 2012 Febrero.
15. Tortoise Hg, Mercurial, <http://tortoisehg.bitbucket.org>.

16. Marcela V. Urdaniz, María A. Ortiz, et al., Diseño y Simulación de una Órtesis Robótica para Rehabilitación de la Marcha, (Universidad Nacional de entre Ríos, Revista Iberoamericana de Ingeniería Mecánica, Argentina)(2007).
17. D. H. and B. T., "Simultaneous localization and mapping: Part 1," IEEE Robotics and Automation, 2006.
18. B. J. and F. L., "Correction of systematic odometry errors in mobile robots," IEEE/RSJ International Conference, 1995.
19. Dataset: Málaga dataset 2006, http://www.mrpt.org/robotic_datasets
20. Dataset: ParisTech, <http://cogrob.ensta-paristech.fr/loopclosure.html>

Desacoplamiento dinámico y control de robot RRR de arquitectura abierta

Daniel De la Cruz-Muciño¹, Miguel Gabriel Villarreal-Cervantes¹, Jesús Said Pantoja-García², Nancy Donaji Hidalgo-Vázquez¹

¹CIDETEC-IPN, D.F., México

danielclcm@msn.com.mx, mvillarrealc@ipn.mx, donaji@mmm.com

²ESIME-ZAC-IPN, D.F., México

salosono23@hotmail.com

Paper received on 20/07/12, Accepted on 13/08/12.

Resumen. En este trabajo se desarrolla el modelo dinámico desacoplado de un robot tipo RRR de tres grados de libertad (x, y, z) con base en la formulación de Euler-Lagrange. Se valida experimentalmente el modelo obtenido sobre una plataforma de arquitectura abierta, presentándose una comparación empírica de los resultados en simulación y experimental.

Palabras Clave: Arquitectura abierta, robot RRR, dinámica, desacoplamiento dinámico.

1 Introducción.

Desde hace muchas décadas atrás, la robótica ha sido importante en el entorno industrial. Sin embargo, actualmente la robótica de servicios [1] ha tenido gran impacto en la vida cotidiana debido al decremento del costo de los componentes que conforman al robot.

Algunos aspectos importantes que se consideran en los sistemas robóticos es la "inteligencia" y la reconfigurabilidad. Se entiende como "inteligencia" a la capacidad que tiene el robot de aprender por sí misma por medio de sensores y procesamiento de la información. Por otra parte, se refiere a reconfigurabilidad a la capacidad de definir funciones propuestas por el usuario tanto en el hardware así como en los programas. Es así, que la mayoría de los sistemas robóticos en el mercado tienen arquitectura cerrada de control en donde los detalles técnicos del sistema están ocultos y generalmente utilizan un protocolo propio de comunicación, por lo que cambiar la estrategia de control no es una tarea fácil. En [2] se propone un sistema de arquitectura abierta para una máquina de control numérico. El resultado obtenido muestra que es fácil implementar estrategias de control y características del monitoreo con el sistema propuesto.

El estudio de la robótica [3] es una parte indispensable para el entendimiento de su cinemática, dinámica, planeación de trayectorias y control. Sin embargo, su estu-

dio formal requiere de sistemas robóticos de arquitectura abierta para llevar a la práctica la teoría. Un sistema de arquitectura abierta se define como aquel sistema en donde su planeación de trayectoria y controlador es flexible y reconfigurable haciendo énfasis en su facilidad para su programación utilizando algún lenguaje de programación.

Por otra parte en [4] se analiza y se desarrolla la dinámica completa del robot SCARA, considerando la dinámica de los actuadores. Se muestran únicamente resultados en simulación y en animación por lo que el trabajo experimental queda como trabajo futuro. Además se enfatiza el hecho de realizar primeramente simulaciones con el modelo del robot antes de llegar a la etapa experimental para evitar algún daño al sistema por una mala sintonización del controlador.

En [5] se desarrolla la dinámica de un manipulador móvil. Se analizan los efectos de las fuerzas que interactúan entre los dos subsistemas (manipulador y robot móvil). Los resultados en simulación muestran que la dinámica del manipulador se afecta por las fuerzas y/o pares ejercidos en el robot móvil pero no sucede lo mismo en caso contrario, es decir, la dinámica del robot móvil no se afecta por los pares y/o fuerzas del manipulador.

Por tal motivo en este trabajo se analiza y desarrolla la dinámica de un robot RRR en forma desacoplada, es decir, como si fueran un robot tipo RR y uno R. Se implementa una arquitectura de control abierta para su fácil implementación experimental, dando como resultado que el modelo desacoplado obtenido en forma analítica es válida en forma experimental.

El artículo se organiza de la siguiente manera: En la sección 2 se desarrolla la cinemática y dinámica del robot. La estrategia de control a utilizar se establece en la Sección 3. La discusión de resultados se muestran en la sección 4 y finalmente en la sección 5 se presentan las conclusiones.

2 Cinemática y Dinámica.

Considere el robot mostrado en la figura 1 donde $O_i \forall i = 0, \dots, 4$ es el origen del sistema de coordenadas del sistema inercial del i -ésimo eslabón y del efector final, respectivamente. La longitud, longitud del centro de masa, la masa y el tensor de inercia expresada en el centro de masa del i -ésimo eslabón se expresan como $l_i \in R, l_{c_i} \in R, m_i \in R, {}^c I_i \in R^{3 \times 3}$, respectivamente y $h \in R$ es la distancia de la base al primer eslabón.

Definiendo las coordenadas y velocidades generalizadas como $q = [q_1 \ q_2 \ q_3]^T$ y $\dot{q} = [\dot{q}_1 \ \dot{q}_2 \ \dot{q}_3]^T$ se puede obtener la matriz de transformación [6] (1) que describe la cinemática directa del robot, donde $A \in R$ es el desplazamiento lineal por vuelta del husillo del actuador final sobre el eje Z.

$${}_{0_0}T_{O_4} = \begin{bmatrix} \cos(q_1 + q_2) & -\sin(q_1 + q_2) & 0 & l_2 \cos(q_1 + q_2) + l_1 \cos q_1 \\ \sin(q_1 + q_2) & \cos(q_1 + q_2) & 0 & l_2 \sin(q_1 + q_2) + l_1 \sin q_1 \\ 0 & 0 & 1 & h + \frac{1}{2\pi} A q_3 - l_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

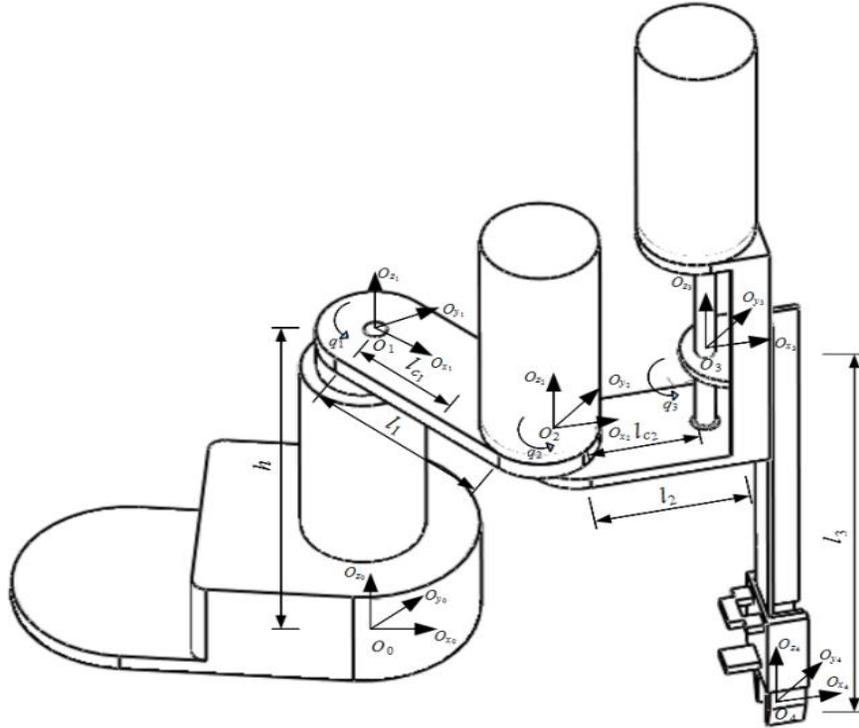


Figura 1. Diagrama esquemático del robot RRR.

Debido a que la inercia del eslabón 3 es pequeña comparada con las inercias de los otros dos eslabones y con el propósito de simplificar el modelo dinámico del robot RRR, se propone desacoplar su dinámica en dos subsistemas, por lo que se establece que el robot RRR consiste de dos robots, uno de tipo RR y otro de tipo R.

Para obtener el modelo dinámico del robot RR por Lagrange se establece la velocidad lineal del centro de masa del eslabón 1 y 2 con respecto al sistema inercial ${}^{0_0}V_{c_i} \forall i = 1, 2$, dado por (2) y (3).

$${}^{0_0}V_{c_1} = [-l_{c_1} \sin(q_1) \dot{q}_1 \quad l_{c_1} \cos(q_1) \dot{q}_1 \quad 0]^T \quad (2)$$

$${}^{0_0}V_{c_2} = \begin{bmatrix} -l_1 \dot{q}_1 \sin(q_1) - l_{c_2} (\dot{q}_1 + \dot{q}_2) \sin(q_1 + q_2) \\ l_1 \dot{q}_1 \cos(q_1) + l_{c_2} (\dot{q}_1 + \dot{q}_2) \cos(q_1 + q_2) \\ 0 \end{bmatrix} \quad (3)$$

La energía cinética K_i del robot RR está dada por (4), donde $\omega_1 = [0 \ 0 \ \dot{q}_1]^T$ y $\omega_2 = [0 \ 0 \ \dot{q}_1 + \dot{q}_2]^T$ es la velocidad angular del i -ésimo eslabón escrita en su mismo sistema de coordenadas.

$$K_1 = \sum_{i=1}^2 \left(\frac{1}{2} m_i {}^{0_0}V_{c_i}^2 + \frac{1}{2} w_i^T c_i I_i w_i \right) \quad (4)$$

Es importante mencionar que en el robot RR la energía potencial es cero debido a que el centro de masa de los eslabones 1 y 2 medido en el eje O_{z_0} permanece constante. Entonces el Lagrangiano \mathcal{L}_1 es igual a la energía cinética. Por lo que las ecuaciones de movimiento de Lagrange están dadas por (5) y el modelo dinámico del robot se expresa en (6), donde $\bar{\tau} = [\tau_1 \ \tau_2]^T$ es el vector de pares y fuerzas de entrada,

$\bar{M} \in R^{2 \times 2}$ es la matriz de inercia, $\bar{C}\dot{\bar{q}} \in R^2$ es el vector de fuerzas centrífugas y de Coriolis, $\bar{q} = [q_1 \ q_2]^T$, $\dot{\bar{q}} = [\dot{q}_1 \ \dot{q}_2]^T$, $\ddot{\bar{q}} = [\ddot{q}_1 \ \ddot{q}_2]^T$ son los vectores de posición, velocidad y aceleración de los eslabones 1 y 2, respectivamente,

$\bar{\tau}_f = [k_{c_1} \text{sgn}(\dot{q}_1) \ k_{c_2} \text{sgn}(\dot{q}_2)]^T$ es el vector de fuerza de fricción de Coulomb y $k_{c_i} \forall i = 1, 2$ es la constante de fricción de Coulomb.

$$\bar{\tau}_i = \frac{d}{dt} \left(\frac{d\mathcal{L}}{dq_i} \right) - \frac{d\mathcal{L}}{dq_i} \forall i = 1, 2 \quad (5)$$

$$\bar{\tau}_i = \bar{M}(\bar{q})\ddot{\bar{q}} + \bar{C}(\bar{q}, \dot{\bar{q}})\dot{\bar{q}} + \bar{\tau}_f \quad (6)$$

Donde:

$$\bar{M}_{11} = m_2 l_1^2 + 2 m_2 \cos(q_2) l_1 l_{c_2} + m_1 l_{c_2}^2 + I_{zz_1} + I_{zz_2}$$

$$\bar{M}_{12} = \bar{M}_{21} = m_2 l_{c_2}^2 + l_1 m_2 \cos(q_2) l_{c_2} + I_{zz_2}$$

$$\bar{M}_{22} = m_2 l_{c_2}^2 + I_{zz_2}$$

$$\bar{C}_{11} = -2\dot{q}_2 l_1 l_{c_2} m_2 \sin(q_2)$$

$$\bar{C}_{12} = -\dot{q}_2 l_1 l_{c_2} m_2 \sin(q_2)$$

$$\bar{C}_{21} = \dot{q}_1 l_1 l_{c_2} m_2 \sin(q_2)$$

$$\bar{C}_{22} = 0$$

En el modelo dinámico del robot R, la energía cinética sólo incluye términos de la derivada momento angular, dando como resultado (7), donde $J \in R$ es la suma del momento de inercia del husillo y del rotor del motor del efector final.

$$K_2 = \frac{1}{2} J \dot{q}_3^2 \quad (7)$$

Definiendo el Lagrangiano como $\mathcal{L}_1 = K_2$, el modelo dinámico del robot R se expresa en (8), donde $\bar{M} = J$, $\hat{\tau}_f = k_{v_3} \dot{q}_3$ es el par de fricción viscosa y $k_{v_3} \in R$ es el coeficiente de fricción viscosa.

$$\tau_3 = M \ddot{q}_3 + \tau_f \in R \quad (8)$$

Finalmente, el modelo completo desacoplado se observa en (9), donde $\tau = [\bar{\tau} \ \tau_3]^T \in R^3$.

$$\tau = \begin{bmatrix} \bar{M}_{11} & \bar{M}_{12} & 0 \\ \bar{M}_{21} & \bar{M}_{22} & 0 \\ 0 & 0 & \bar{M} \end{bmatrix} \ddot{\bar{q}} + \begin{bmatrix} \bar{C}_{11} & \bar{C}_{12} & 0 \\ \bar{C}_{21} & \bar{C}_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{\bar{q}} + \begin{bmatrix} \bar{\tau}_f \\ \hat{\tau}_f \end{bmatrix} \quad (9)$$

3 Sistema De Control.

Se propone un controlador PID cuyo esquema utiliza el error de posición angular $e_i = q_{d_i} - q_{r_i}$ para calcular el par de los actuadores de acuerdo a (10).

$$u_i = k_{p_i} e_i + K_{d_i} \dot{e}_i + K_{i_i} \int_0^{t_f} e_i dt \quad \forall i = 1,2,3 \quad (10)$$

Donde q_{d_i} , \dot{q}_{d_i} y q_{r_i} , \dot{q}_{r_i} es la posición y velocidad angular deseada y medida del i -ésimo eslabón, respectivamente, $k_{p_i} \in R$; $k_{d_i} \in R$ y $k_{i_i} \in R$ es la ganancia proporcional, derivativa e integral y $\dot{e}_i = \dot{q}_{d_i} - \dot{q}_{r_i}$ es el error de velocidad angular.

4 Resultados.

El sistema experimental de arquitectura abierta consiste de un robot RRR el cual contiene tres servomotores Pittman, dos de ellos de la serie GM9234S033 y el otro de la serie GM9236S015, ambos con decodificador óptico de 500 pulsos por vuelta, una unidad de procesamiento central (CPU por sus siglas en inglés) Pentium IV a 3.2GHz con 4GB de RAM, de una tarjeta de adquisición de datos Sensoray 626 con interconexión de componentes periféricos (PCI por sus siglas en inglés) y de la etapa de potencia conformado por tres convertidores voltaje-corriente LMD18245, conectados al puerto de salidas analógicas de la Sensoray.

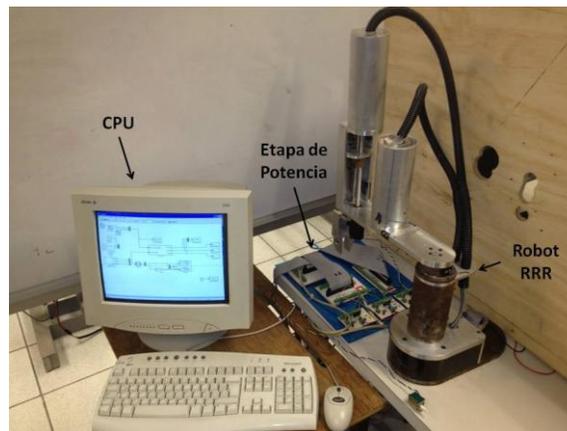


Figura 2. Sistema experimental de arquitectura abierta.

Para realizar las simulaciones numéricas se consideraron los parámetros del robot y del sistema de control mostrados en la Tabla 1. Cabe mencionar que los parámetros del robot se obtienen con la ayuda del programa de diseño paramétrico SolidWorks y el problema de control se aborda como un problema de seguimiento de trayectorias sin restricciones, donde las ganancias del controlador se obtienen de acuerdo al procedimiento presentado en [7]. En la Fig. 3 se muestra el diagrama esquemático del sistema experimental.

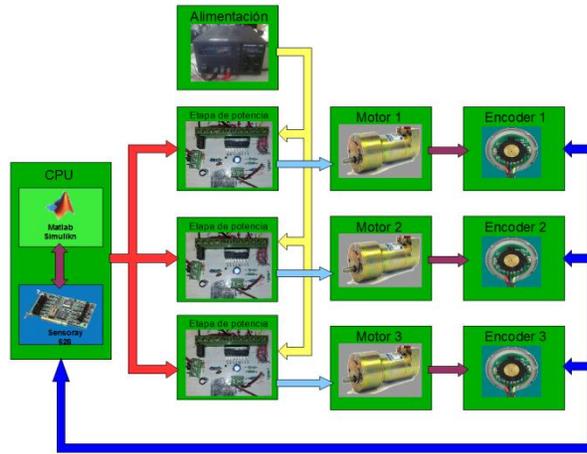


Figura 3. Diagrama esquemático del sistema experimental.

Se utilizó Matlab-Simulink para realizar la simulación numérica con el método de Runge-Kutta de 4to orden y un tiempo de integración de $\Delta t_s = 5ms$, un tiempo final $t_f = 120s$ y condición inicial $q(t_0 = 0) = [0 \ 0 \ 0]^T$. Las trayectorias deseadas a seguir por el robot RRR se muestran en (11). Para los resultados experimentales se propuso un tiempo de muestreo de $\Delta t_e = 5ms$ y se programó la ley de control y la trayectoria deseada en Matlab-Simulink.

$$\begin{aligned}
 q_{d1}(t) &= \frac{\pi}{2} \sin(0.2t) \\
 q_{d2}(t) &= -\frac{\pi}{2} \sin(0.5t) \\
 q_{d3}(t) &= 0.03 \sin\left(\frac{2\pi}{50}t\right)
 \end{aligned} \tag{11}$$

Tabla 1. Parámetros del robot RRR y ganancias del controlador PID.

Eslabón 1	$l_1 = 0.152 \text{ m}$	$l_{c1} = 0.114$	$m_1 = 1.472 \text{ kg}$	$I_{zz1} = 0.005 \text{ Kg m}^2$
	$k_{p1} = 2$	$k_{i1} = 0.1207$	$k_{d1} = 2$	$k_{c1} = 0.199 \text{ Nm}$
Eslabón 2	$l_2 = 0.1342 \text{ m}$	$l_{c2} = 0.105 \text{ m}$	$m_2 = 2.207 \text{ Kg}$	$I_{zz2} = 0.003 \text{ Kg m}^2$
	$k_{p2} = 2.8$	$k_{i2} = 0.4$	$k_{d2} = 2$	$I_{zz2} = 0.003 \text{ Kg m}^2$
Eslabón 3	$l_3 = 0.37841 \text{ m}$	$J = 0.00028 \text{ Kg m}^2$	$h = 0.2841 \text{ m}$	$A = 0.002 \text{ m}$
	$k_{p3} = 0.003$	$k_{i3} = 0.0002$	$k_{d3} = 0$	$k_{v3} = 0.0000555 \text{ Nms}$

En las figuras 4, 5 y 6 se muestran los resultados en simulación y en forma experimental del seguimiento de la trayectoria deseada con el controlador propuesto. Se observa que el resultado experimental con el resultado numérico presenta un comportamiento similar, con lo cual podemos validar el modelo dinámico obtenido. Sin embargo, la discrepancia entre los resultados de simulación y experimental se debe principalmente a que no se incluye en el modelo la dinámica asociada con el juego de engranes, además el modelo de fricción propuesto es ideal.

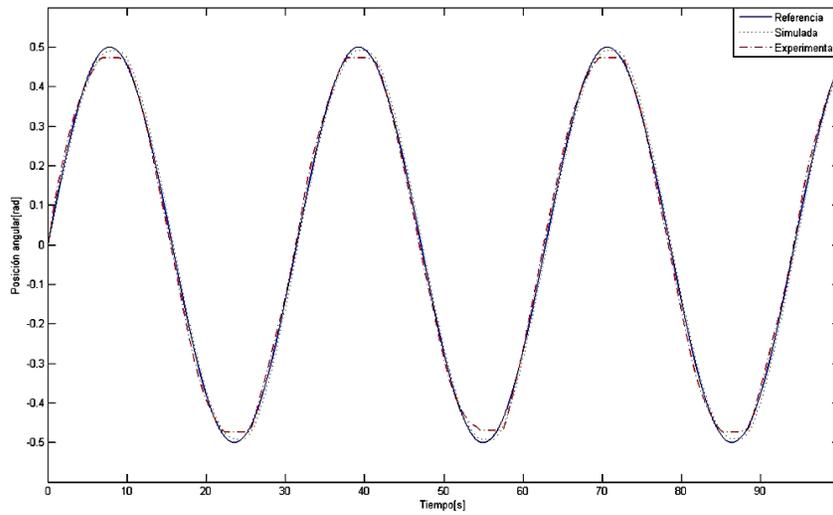


Figura 4. Resultados en simulación y experimental del seguimiento de la trayectoria para q_1 .

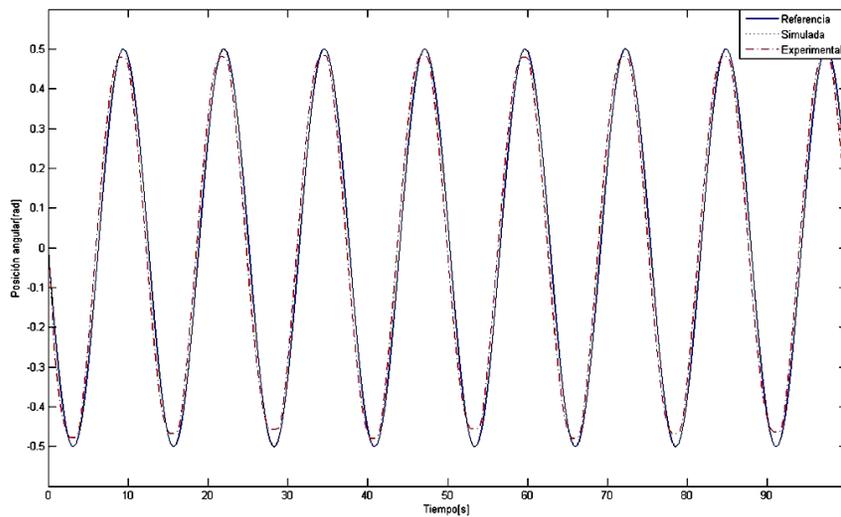


Figura 5. Resultados en simulación y experimental del seguimiento de la trayectoria para q_2 .

5 Conclusiones.

En este artículo se presenta el modelo desacoplado del robot RRR y se muestran los resultados experimentales en una plataforma de arquitectura de control abierta.

Comparando los resultados en simulación con los resultados experimentales se observa que es factible realizar el desacoplamiento dinámico del robot RRR en dos robots, uno de tipo RR y otro de tipo R.

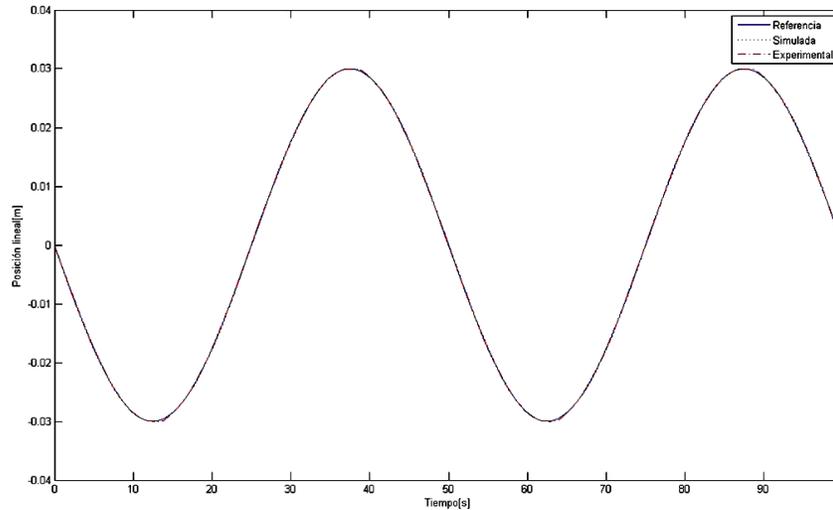


Figura 6. Resultados en simulación y experimental del seguimiento de la trayectoria para q_3 .

Se mostró el hecho de que es muy útil realizar primeramente simulaciones del sistema a controlar antes de llevar a cabo los resultados experimentales. Sin embargo es muy importante que se obtenga e identifique correctamente los parámetros del modelo del sistema para que los resultados en simulación correspondan a los resultados experimentales.

Algunos beneficios de realizar simulaciones es su fácil reconfiguración, se realiza en forma rápida, y es más conveniente, seguro y barato que estar probando con el sistema experimental.

La arquitectura de control abierta hace posible que la implementación experimental sea más rápida y eficiente.

Agradecimientos

Los autores agradecen el apoyo otorgado por la Secretaría de Investigación y Posgrado del Instituto Politécnico Nacional bajo el proyecto SIP 20120663 y al CONACYT bajo el proyecto 084060.

Se agradece el apoyo del Dr. Gabriel Sepúlveda Cervantes y del CNAD.

Referencias

1. Takashi Kabe, Kohichi Tanaka, Hiroyasu Ikeda, Noboru Sugimoto, Consideration on safety for emerging technology – Case studies of seven. *Safety Science*. 2010, Vol. 48, pp. 296-301.
2. Luis Morales-Velazquez, Rene de Jesus Romero-Troncoso, Roque Alfredo Osornio-Rios, Gilberto Herrera-Ruiz, Eduardo Cabal-Yepez, Open-architecture system based on a reconfigurable hardware–software multi-agent platform for CNC machines. *Journal of Systems Architecture*. 2010, Vol. 56, pp. 407-418.

3. Mark W. Spong, Seth Hutchinson, M. Vidyasagar, Robot Modeling and Control. Wiley. 2005.
4. Mahdi Salman Alshamasin, Florin Ionescu, Riad Taha Al-Kasasbeh, Kinematic Modeling and Simulation of a SCARA Robot by Using Solid Dynamics and Verification by MATLAB/Simulink. European Journal of Scientific Research. 2009, Vol.37 No.3 (2009), pp.388-405
5. Miguel G. Villarreal-Cervantes, Carlos A. Cruz-Villar, S. A. Saldívar Bárcenas, L. E. García Sosa, Dynamic coupling between a parallel manipulator and an omnidirectional mobile platform. XV Latinamerican Control Conference, Lima, Perú, Octubre 2012.
6. John J. Craig, Introduction to robotics: mechanics and control, Prentice Hall. 2004.
7. Miguel G. Villarreal-Cervantes, Jesús Said Pantoja García, Análisis comparativo entre un control heurístico y un PID para un sistema mecatrónico, 3th International Supercomputing Conference. Guanajuato, México, Marzo 2012.

Aplicación de Técnicas Inteligentes de Control en un Motor de Corriente Directa Incluyendo Modulación por Ancho de Pulsos

R. Morales-Caporal, J. I. Paredes-Tabales, O. Sandre-Hernández,
E. Bonilla-Huerta y J. F. Ramírez-Cruz

Instituto Tecnológico de Apizaco
División de Estudios de Posgrado e Investigación
Av. Instituto tecnológico s/n, Apizaco, Tlax. C.P. 90300
rmcaporal@ieee.org. Tel: (01241) 41-7-20-10 ext. 138
Paper received on 14/06/12, Accepted on 03/09/12.

Resumen. En este trabajo, se presenta la aplicación de dos técnicas de inteligencia artificial (Lógica Difusa y Redes Neuronales Artificiales) para controlar la velocidad y la posición del eje de un motor de Corriente Directa (CD). Primero se muestra el diseño de un controlador clásico (Proporcional-Integral-Derivativo) con el que se realizan comparaciones, después se explica el diseño de los controles basados en inteligencia artificial. Se incluye una técnica de modulación de ancho de pulsos (basada en banda de histéresis) a la salida de los controladores, la cual es necesaria durante la implementación, ya que el motor debe ser accionado con dispositivos de electrónica de potencia. Las técnicas de control propuestas, el accionamiento electrónico y el motor se simulan digitalmente con la herramienta computacional Matlab/Simulink[®]. Los resultados simulados muestran que con las técnicas inteligentes de control se logra un mejor desempeño de la máquina en estado transitorio y estacionario, además se determina el controlador más factible para una posterior implementación digital.

Palabras Clave: Control de velocidad, control de posición, lógica difusa, modulación por ancho de pulsos, motor de CD, redes neuronales artificiales.

1 Introducción.

Los motores de CD son altamente usados en aplicaciones industriales, tales como en máquinas herramienta, empacadoras, enrolladoras, manipuladores robóticos, etc., donde se requiere un control de la velocidad, posición y par del eje del motor con un alto desempeño [1]. Tradicionalmente, estas variables de la máquina se han regulado con controladores clásicos (Proporcional-Integral [PI] o Proporcional-Integral-Derivativo [PID]), gracias a su bajo costo y su simplicidad.

© J. C. Hernández-Hernández, J. F. Ramírez Cruz,
A. Cortés Fernández, J. H. Sossa Azuela. (Eds.).
Advances in Intelligent and Information Technologies.
Research in Computing Science 60, 2012, pp. 121 -130.



Sin embargo, para sintonizar correctamente estos controladores es necesario obtener un modelo matemático y parámetros precisos de la máquina, los cuales se obtienen comúnmente “off-line”. Por lo tanto, cuando los parámetros de la máquina varían debido a su operación, estos controles pierden su sintonización y disminuye gradualmente su ajuste, dando como resultado un error incremental en las variables controladas [2]. Para corregir esto, se han propuesto diversas técnicas, entre las que destacan: control adaptivo [3], control repetitivo [4], sliding mode control [5], [6], control difuso [7]-[9] y control neuronal [10]-[12]. Sin embargo, puesto que estas últimas técnicas modernas de control se han implementado en máquinas diferentes, no es posible encontrar en la literatura una comparación real del desempeño que ofrece cada uno de los controladores antes mencionados en la misma máquina, sobre todo para el caso del control de la posición. Por eso, en este trabajo se analizan tres técnicas diferentes de control en la misma máquina, para evaluar su desempeño, además se incluye una técnica de modulación del ancho de los pulsos, ya que esta última parte representa una no linealidad en el sistema real.

2 Control Clásico de Velocidad y Posición del Motor de CD.

2.1 Modelo del Motor.

El modelo matemático de la máquina de CD de imanes permanentes bajo estudio puede escribirse como sigue [2]:

$$\frac{di_a}{dt} = \frac{V_a}{L_a} - \frac{R_a \cdot i_a}{L_a} - \frac{K_b}{L_a} \omega_m; \quad K_b = K \cdot \phi \quad (1)$$

$$\frac{d\omega_m}{dt} = \frac{K_b \cdot i_a}{J} - \frac{\tau_L}{J} = \frac{\tau_e}{J} - \frac{\tau_L}{J}; \quad \frac{d\gamma}{dt} = \omega_m \quad (2)$$

Donde K representa una constante proporcional que depende de la construcción física de la máquina, R_a , L_a e i_a , representan la resistencia, la inductancia, y la corriente de armadura respectivamente; K_b se define como la constantes del motor; τ_e y τ_L , representan el par electromagnético y el par de carga respectivamente; V_a , es el voltaje de armadura, J es la inercia del motor; ω_m y γ son la velocidad y posición angular del eje del rotor respectivamente.

2.2 Control Clásico.

En este trabajo se usa primero un controlador clásico PID para controlar la velocidad y la posición del eje de la máquina, ya que muestra mejor desempeño que un controlador PI. La ecuación matemática del controlador PID está dada por [13]:

$$u(t) = K_p e(t) + K_i \int e(t) dt + K_d \frac{de(t)}{dt} \quad (3)$$

Donde $e(t)$ es el error entre el valor de referencia y la salida del sistema, $u(t)$ es la salida del controlador, K_p , K_i , K_d , son la ganancia proporcional, la ganancia integral y la ganancia derivativa respectivamente. En este caso, para determinar los valores de las ganancias del controlador se utilizaron las técnicas de sintonización de Ziegler-Nichols [13]. La Fig. 1a, muestra el diagrama de bloques del control clásico de velocidad; en esta figura se observa que la entrada del controlador es el error de la velocidad y su salida es la corriente de referencia i_a^* , la cual, a su vez se compara contra la corriente medida para generar la modulación del ancho de pulso por medio de un regulador de histéresis. El diagrama de bloques del controlador de la posición del eje del motor se muestra en la Fig. 1b, este tiene como entrada el error de la posición y como salida también la corriente de referencia i_a^* . La cual genera los pulsos para accionar la etapa de potencia, como en el caso anterior.

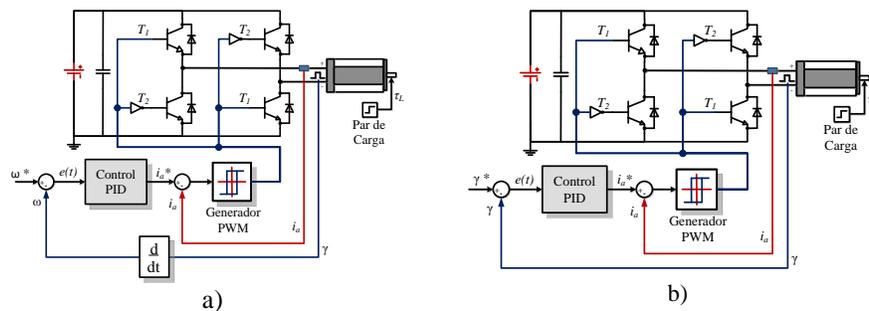


Figura 1. Diagrama de bloques del control clásico. a) De velocidad. b) De posición.

3 Control Inteligente de Velocidad y Posición del Motor de CD.

El Control Inteligente (CI) es una disciplina donde los métodos de control desarrollados tienden a emular características importantes de la inteligencia humana, entre las que se encuentran: adaptación y aprendizaje, respuesta óptima bajo altas incertidumbres, y hacer frente a grande cantidad de datos. El CI ha sido aplicado en áreas como robótica y automatización, comunicaciones, manufactura, control de tráfico, etc. En el control de máquinas eléctricas, la aplicación del CI, particularmente la lógica difusa y las RNA, ha demostrado ser una herramienta eficaz, sobre todo cuando hay variación de parámetros, variaciones en el par de carga o modos anormales de operación [14].

3.1 Control Difuso de Velocidad del Motor de CD.

El controlador difuso de velocidad del motor de CD contiene dos antecedentes y una consecuencia, el error de velocidad $e(t)$, la derivada del error $de(t)/dt$ y la corriente de armadura i_a respectivamente. El diagrama de bloques del controlador difuso de velocidad propuesto para el motor de CD se muestra en la Fig. 2a.

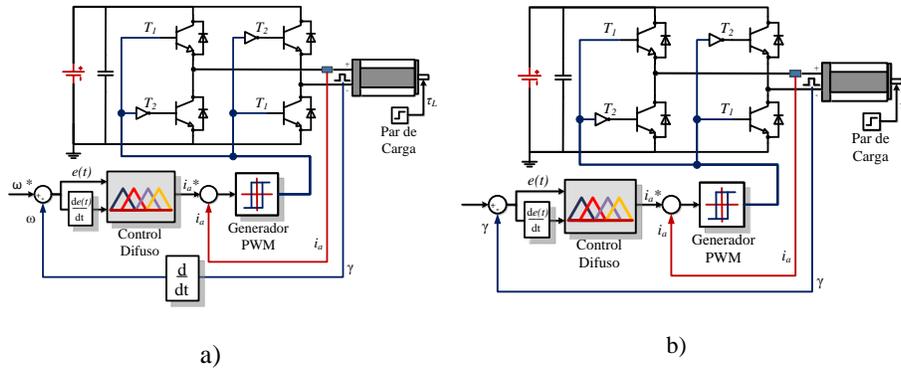


Figura 2. Diagrama de bloques del control difuso. a) De velocidad. b) De posición.

La estrategia de fuzzificación convierte las variables de entrada y salida en los siguientes conjuntos difusos: NG (Negativo Grande), NP (Negativo Pequeño), CE (Cero), PP (Positivo Pequeño), PG (Positivo Grande). Se han seleccionado funciones de membresía triangulares para cada uno de los cinco conjuntos difusos de cada variable [8]. En la Fig. 3 se observa los parámetros de las funciones de membresía de cada una de las variables. La base de reglas es determinada para cada uno de los conjuntos difusos de cada variable, y después de una optimización [2], se obtuvieron un total de 25 reglas difusas como se muestra en la Tabla 1.

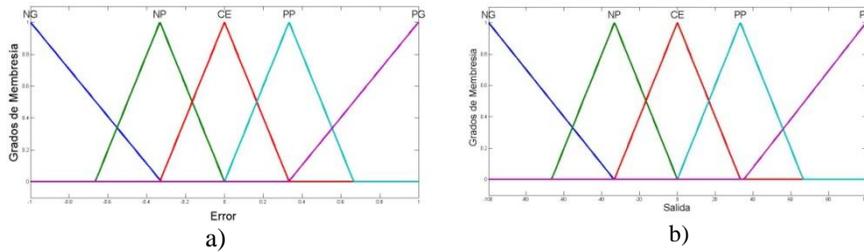


Figura 3. Funciones de membresía del control difuso de velocidad. a) Error y derivada del error. b) Salida control difuso (corriente de referencia i_a).

Tabla 1. Base de reglas del control de velocidad.

$e_{\omega}(k)$ \ $\Delta e_{\omega}(k)$	NG	NP	Z	PP	PG
NG	NG	NG	NG	NP	NG
NP	NG	NG	NP	Z	PP
Z	NG	NP	Z	PP	PG
PP	NP	Z	PP	PG	PG
PG	PG	PP	PG	PG	PG

3.2. Control Difuso de Posición del Motor de CD.

El controlador difuso de posición del motor de CD contiene dos antecedentes y una consecuencia, el error de velocidad $e(t)$, la derivada del error $de(t)/dt$ y la corriente de armadura i_a respectivamente. El diagrama de bloques del controlador difuso de posición del motor de CD se muestra en la Fig. 2b.

La estrategia de fuzzificación convierte las variables de entrada y salida en los siguientes conjuntos difusos: NG (Negativo Grande), NP (Negativo Pequeño), CE (Cero), PP (Positivo Pequeño), PG (Positivo Grande). También, se han seleccionado funciones de membresía triangulares para cada uno de los cinco conjuntos difusos de cada variable. En la Fig. 4 se observa los parámetros de las funciones de membresía de cada una de las variables. La base de reglas utilizadas se muestra en la Tabla 1.

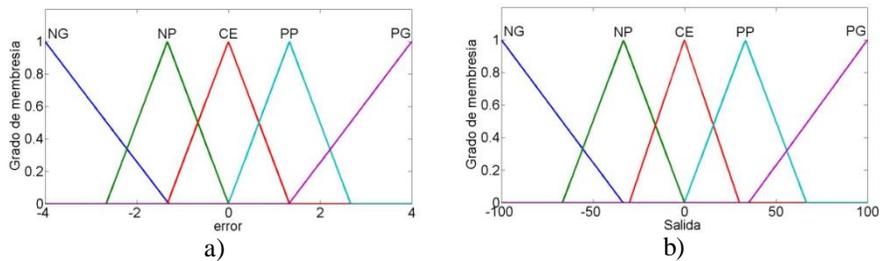


Figura 4. Funciones de membresía del control difuso de posición. a) Error y derivada del error. b) Salida control difuso (corriente de referencia i_a).

3.3 Control Neuronal de Velocidad del Motor de CD.

Las RNA tienen aplicaciones en la teoría de control, gracias a su capacidad de aprendizaje y adaptabilidad, de modo que es posible, entrenar una RNA para que emule el comportamiento de un controlador [10]. Últimamente, se han propuesto diversas arquitecturas de RNA para muchas aplicaciones de control industrial, siendo el control de máquinas eléctricas uno de los variados campos de aplicación.

Para controlar la velocidad del motor de CD se propone un neurocontrolador directo de tres capas, el diagrama de bloques del neurocontrolador propuesto se muestra en la Fig. 5a. Los datos de entrenamiento han sido tomados del controlador difuso, con los datos obtenidos se entrena la red por medio del algoritmo de retro propagación, las especificaciones del neurocontrolador se muestran en la tabla 2.

Tabla 2. Especificaciones del Neurocontrolador.

Especificaciones del Neurocontrolador	
Numero de Neuronas capa de Entrada	2
Numero de Neuronas capa de Salida	1
Numero de Neuronas capa Oculta	20
Ejemplos de Entrenamiento	5000
Numero de Épocas	1000

3.4 Control Neuronal de Posición del Motor de CD.

El neurocontrolador de posición tiene la misma arquitectura y entrenamiento que el neurocontrolador de velocidad (mostrados en la Tabla 2), sin embargo los datos de entrenamiento han sido tomados del controlador difuso de posición. El diagrama de bloques del neurocontrolador de posición se muestra en la Fig. 6b. Una vez entrenado el control basado en RNA, se sustituye en lugar del controlador difuso.

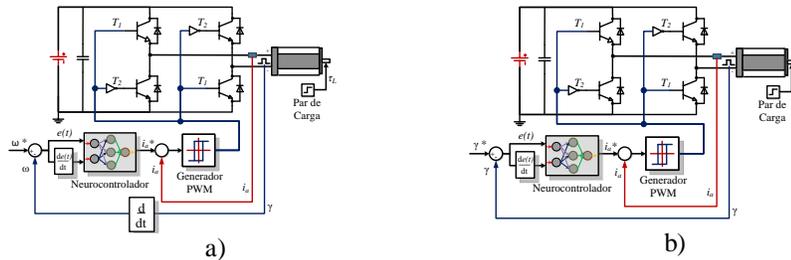


Figura 5. Diagrama de bloques del neurocontrolador. a) De velocidad. b) De posición.

4 Resultados.

Los parámetros del motor de CD utilizado en este trabajo se muestran en la tabla 3 y algunos programas desarrollados se muestran en el apéndice.

Tabla 3. Parámetros del motor de CD.

R_a	0.6Ω	J	1.2 Kg m^2
L_a	12 mH	K_b	1.8 Vs/rad
V_a	240V	T_L	29 Nm

4.1 Resultados de las Pruebas de Velocidad.

Para realizar las pruebas, se ajusta la velocidad de referencia a 60 rad/s. En $t=2.5$ s, se aplica un par de carga de 8Nm. La Fig. 6 muestra los resultados de la velocidad del motor de CD con los diferentes esquemas de control. La Fig. 6a muestra la señal de velocidad con el controlador PID, se observa que la velocidad de referencia es alcanzada rápidamente, y al aplicar el par de carga no se observa perturbación alguna. Las Figs. 6b y 6c muestran los resultados de la velocidad con el controlador difuso y neuronal respectivamente, se observa que la velocidad de referencia es alcanzada casi tan rápido como con el controlador PID, y también sin perturbarse al aplicar el par de carga. Las simulaciones demuestran que el control de velocidad, en este caso, no es un problema complejo, ya que tanto el control clásico, como el control inteligente, alcanzan la velocidad de referencia rápidamente, y aún cuando se aplica el par de carga, la velocidad del motor se mantiene en su referencia.

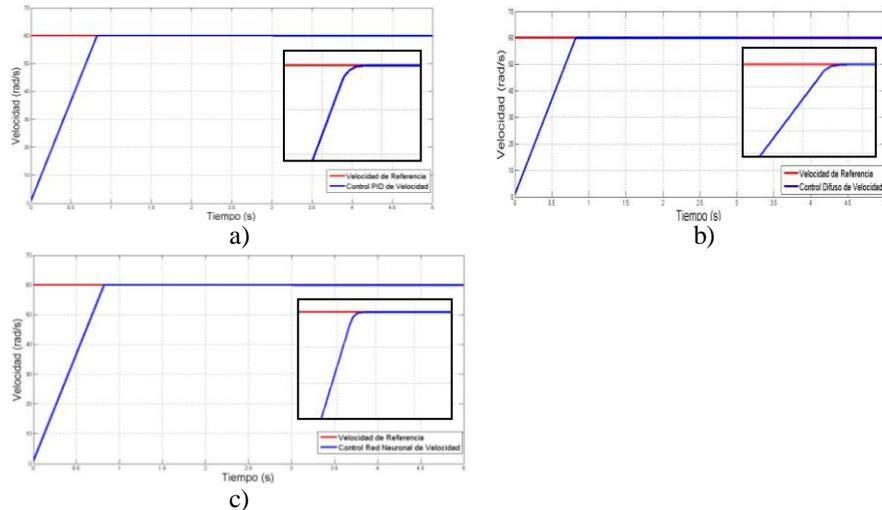


Figura 6. Simulación de la γ del motor de CD con control a) Clásico, b) Difuso, c) Neuronal.

4.2. Resultados de las Pruebas de Posición.

Para estas pruebas, se ajusta la posición de referencia a 2 rad. En $t=4$ s, se aplica un par de carga de 8 Nm. La Fig. 7 muestra los resultados de la posición del motor de CD bajo los diferentes esquemas de control propuesto. La Fig. 7a muestra la respuesta de la señal de posición del eje del motor con el control PID, se observa un sobretiro al momento de arranque, así como una caída en la posición de la máquina al aplicar el par de carga. La Fig. 7b y 7c muestran los resultados obtenidos del control de la posición del eje del motor de CD con control difuso y neuronal respectivamente. El controlador difuso muestra un mejor desempeño, puesto que no se observa un sobretiro que sobrepase la referencia, al momento de aplicar el par de carga la posición del eje de la máquina se mantiene en su referencia sin mostrar perturbaciones. Por último debido a que controlador neuronal fue entrenado tomando la salida del controlador difuso, la respuesta de la posición de la máquina es similar a la respuesta obtenida al aplicar el controlador difuso.

5 Conclusiones.

De acuerdo a los resultados de la simulación se puede concluir lo siguiente:

- El control de velocidad del motor de CD no representa un problema complejo, ya que el control clásico (PID), como el control inteligente (Difuso y RNA), muestran un desempeño satisfactorio manteniendo la velocidad del motor en la referencia aun durante un transitorio de carga.

- El control clásico de posición no presenta un buen desempeño, si considera un método complejo para obtener las ganancias del controlador PID, quizá el rendimiento del sistema mejore.
- El control difuso de posición es con el cual se obtiene un mejor desempeño al controlar la posición del motor, aunque presenta un pequeño sobrepaso al alcanzar la posición de referencia, no muestra una caída de posición durante la aplicación del par de carga.
- El control neuronal presente un buen desempeño tanto para la velocidad como para la posición, sin embargo, cuando se aplica el par de carga existe una pequeña diferencia respecto a la referencia, además, se necesita generar un conjunto de datos de entrenamiento para poder realizarlo, lo cual lo hace menos atractivo para su implementación.

Con las conclusiones obtenidas, es posible decir, que el mejor controlador para el motor de CD es el controlador difuso, lo cual lo hace el mejor candidato para su implementación en un sistema digital.

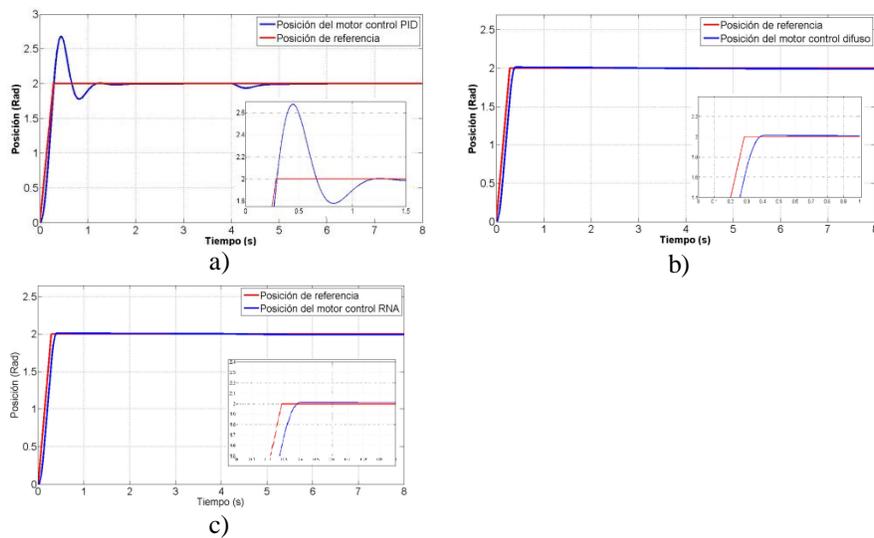
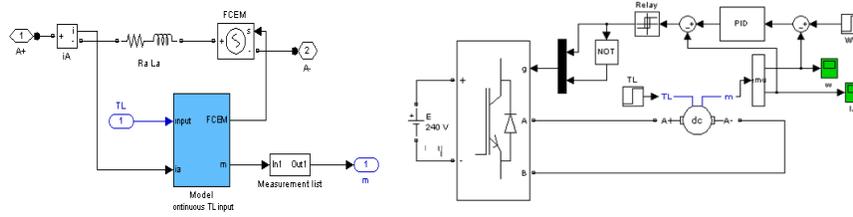


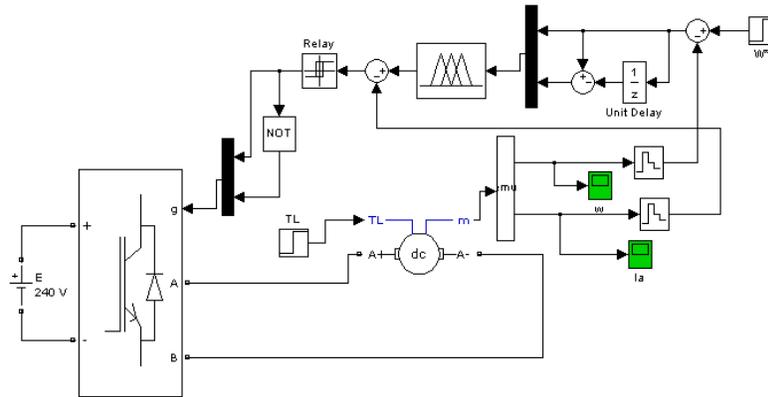
Figura 7. Simulación de la γ del motor de CD con control a) Clásico, b) Difuso, c) Neuronal.

Apéndice.

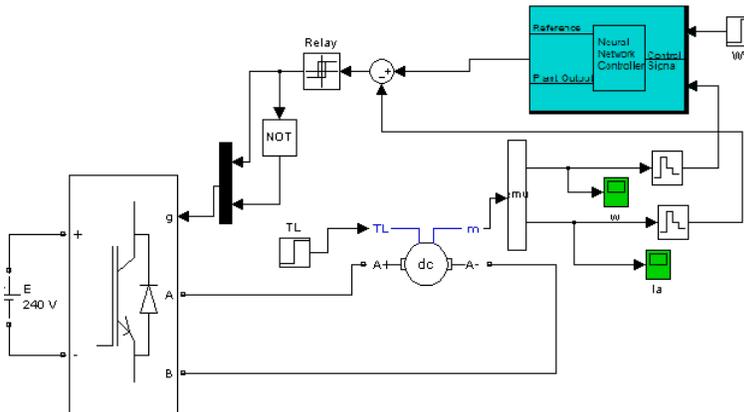
Programas en Matlab/Simulink



Motor de CD (MCD) de imanes permanentes MCD con control PID de velocidad permanentes



MCD de imanes permanentes con control difuso de velocidad, incluyendo modulación por ancho de pulsos.



MCD de imanes permanentes con control neuronal de velocidad, incluyendo modulación por ancho de pulsos.

Referencias.

1. R. Morales-Caporal, E. Bonilla-Huerta, and M. A. Morales-Caporal "DSP-Based Digital Torque/Motion Control of DC Motors for Direct-Drive Industrial Robotics Applications", In. Proc. IEEE Electronics, Robotics and Automotive Mechanics Conference (CERMA 2010). vol. 1. pp. 613-618. Cuernavaca, México, 2010.
2. I. Hernández, "Desarrollo de Diferentes Técnicas de Control Inteligente del Movimiento/Par de un Motor de Corriente Directa para Aplicaciones de Robótica Industrial,". Tesis de Maestría en S. C. Ins. Tec. de Apizaco, Tlaxcala, México 2011.
3. D. R. Tutakne, H. M. Suryawanshi, and S. G. Tarnekar, "Adaptive Pulse Synchronizing Control for High-Power-Factor Operation of Variable Speed DC-Drive" IEEE Transactions on Power Electronics, vol. 22, no. 6, pp. 2499-2510. Nov. 2007.
4. W. Wu, "DC Motor Drive Speed Regulator: Using a Repetitive Control Application in a Flatbed Scanner" IEEE Industry Applications Magazine, vol. 18, no. 2 pp. 38-46. Mar/Apr. 2012.
5. B. M. Patre, V. M. Panchade and R. M. Nagarale. "Sliding Mode Control of DC Drives, Sliding Mode Control" Prof. Andrzej Bartoszewicz (Ed.), ISBN: 978-953-307-162-6, InTech, Available from: <http://www.intechopen.com/books/sliding-mode-control/sliding-mode-control-of-dc-drives>. 2011.
6. C. E. Castaneda, A. G. Loukianov, E. N. Sanchez, and B. Castillo-Toedo, "Discrete Time Neural Sliding-Mode Block Control for a DC Motor with Controlled Flux" IEEE Transactions on Industrial Electronics, vol. 59, no. 2, pp. 11924-1207. Feb. 2012.
7. F. Cupertino, A. Lattanzi, and L. Salvatore, "A New Fuzzy Logic-Based Controller Design Method for DC and AC Impressed-Voltage Drives", IEEE Transactions on Power Electronics, vol. 15, no. 6, pp. 974-982. Nov. 2000.
8. R. Morales-Caporal, I. Hernández-Davila, E. Bonilla-Huerta and J. F. Ramírez-Cruz, "Control PI-Fuzzy de la Posición del Rotor de un Motor de Corriente Directa de Capacidad Industrial". Research on Computing Science, IPN, vol. 54. pp. 3-11. October 2010.
9. R. Malhotra and T. Kaur, "DC Motor Control Using Fuzzy Logic Controller", International Journal of Advanced Engineering Sciences and Technologies, vol. 8, n° 2, pp. 291-296, 2011. Available from: <http://www.ijaest.iserp.org/archieves/14-Jul-1-15-11/Vol-No.8-Issue-No.2/27.IJAEST-Vol-No-8-Issue-No-2-DC-MOTOR-CONTROL-USING-FUZZY-LOGIC-CONTROLLER-291-296.pdf>.
10. S. Weerasooriya, and M. A. El-Sharkawi, "Identification and Control of a DC Motor Using Back-Propagation Neural Networks," IEEE Transactions on Energy Conversion, vol. 6, no. 4, pp. 663-669. Dec. 1991.
11. G. M. Rao, and B. V. S. Ram, "A Neural Network Based Speed Control for DC Motor", International Journal of Recent Trends Engineering, vol. 2, no. 6, pp. 121-124. Nov. 2009.
12. R. Morales-Caporal, I. Hernández-Davila, E. Bonilla-Huerta y J. F. Ramírez-Cruz, "Control de Velocidad y Posición de un Motor de Corriente Directa Usando Lógica Difusa y Redes Neuronales Artificiales". 9o Congreso Intenacional Sobre Innovación y Desarrollo Tecnológico (CIINDET 2011). En CD. Paper. 636. 2011.
13. B. C. Kuo, *Sistemas de Control Automatico*, Prentice Hall, 1996.
14. Y. Kang, Y. Chen y M. Chu, "The Artificial Neural Network Applied to Servo Control Systems.", *New Developments in Robotics, Automation and Control*, InTech, 2008

Detección y Cuantificación de Hemorragias Intracerebrales en Tomografía Axial Computarizada

Enrique Irwin Illán García¹, María Guadalupe Medina Barrera²,
Federico Ramírez Cruz³ y José Juan Hernández Mora⁴

Instituto Tecnológico de Apizaco, División de Estudios de Posgrado e Investigación

^{1,2,3,4}Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México

enrique.illan@gmail.com, posgrado@itapizaco.edu.mx,

framirez@itapizaco.edu.mx, jjhmora@itapizaco.edu.mx

Paper received on 24/07/12, Accepted on 27/08/12.

Resumen. En este artículo se presenta un sistema de detección y cuantificación automático de hemorragias intracerebrales en estudios de tomografía axial computarizada. El objetivo es facilitar al especialista la identificación y medición de la región hemorrágica de forma rápida y precisa. Para validar la detección y cuantificación automática se comparan los resultados calculados automáticamente contra la segmentación y cuantificación manual realizada por el especialista. Los resultados demuestran que el método propuesto puede ser utilizado como una herramienta de apoyo para el diagnóstico de hemorragias intracerebrales en estudios de tomografía axial computarizada.

Palabras Clave: Tomografía Axial Computarizada, Hemorragia Intracerebral, MLP, Cuantificación, Detección, Sistema de Visión Artificial.

1 Introducción

La Hemorragia Intracerebral (HIC) es una colección de sangre dentro del parénquima cerebral producida por una rotura vascular espontánea, no traumática. La Tomografía Axial Computarizada (TAC) es la técnica más utilizada para el diagnóstico de una HIC.

Las HIC son urgencias médicas que deben ser atendidas dentro de las primeras 3 horas del inicio de los síntomas. El reconocimiento y cuantificación de una HIC es de vital importancia ya que determina el tratamiento que debe proporcionarse al paciente. Sin embargo, cuando la Hemorragia presenta un aspecto borroso con un contorno irregular o el especialista no tiene suficiente experiencia se puede estimar erróneamente las dimensiones de la región hemorrágica. Las consecuencias pueden ir hasta determinar la intervención quirúrgica innecesaria poniendo en riesgo la vida del paciente. Para mejorar la certeza en el diagnóstico clínico se propone el desarrollo de un Sistema de Visión Artificial (SVA) para la detección y cuantificación automática de la HIC en estudios de TAC.

2 Trabajos relacionados

La detección y cuantificación de HIC en estudios de TAC ha sido una de las áreas de investigación más estudiadas dentro del diagnóstico clínico de enfermedades cerebro-vasculares. Desde la década de los 90's se han propuesto sistemas de cómputo para la segmentación y clasificación de las regiones.

Lee [4] propone un análisis entre los algoritmos de clustering K-means y FCM para la detección de potenciales regiones hemorrágicas. En ambos casos se realiza una clasificación de las potenciales regiones hemorrágicas tomando en cuenta características geométricas así como estadísticos de primer orden a partir del histograma de la imagen.

Kabara [3] propone el uso del análisis de texturas mediante el uso de la matriz de co-ocurrencia de Haralick. En la etapa de clasificación se utiliza el algoritmo C4.5. Los resultados muestran que las características de textura son capaces de formar dos grupos: tejido normal y regiones hemorrágicas.

Gong [1] propone el uso de un algoritmo de umbralización automática para la detección de potenciales regiones hemorrágicas así como de siete características geométricas para la clasificación en tres clases: tejido normal, calcificación y región hemorrágica.

Li [5] proponen un enfoque en donde se hace uso del algoritmo FCM para la segmentación de la masa intracerebral. Para segmentar las potenciales regiones hemorrágicas se hace uso del algoritmo de Máxima Entropía. Como método de clasificación se utiliza el algoritmo C4.5. Los resultados mencionan una tasa de detección de la región hemorrágica del 100% pero la salida muestra que aún se tienen artefactos en la imagen que no forma parte de la región hemorrágica. Este trabajo muestra que con el uso de técnicas basadas en umbralización automática se pueden detectar potenciales regiones hemorrágicas.

Liao [6] utiliza un método de segmentación basado en un análisis multi-resolución. Posteriormente realiza el cálculo de un modelo geométrico elíptico alrededor de la región afectada para obtener la longitud del eje mayor, el eje menor y la excentricidad de la elipse que se utilizan como características de las potenciales regiones hemorrágicas. Adicionalmente, se realiza un análisis comparativo entre los resultados de la segmentación manual y la segmentación automática.

Romana y Korrapati [9] describen una segmentación mediante umbralización automática en donde los parámetros son calculados de acuerdo a los niveles de gris de la masa Intracerebral. Posteriormente se utiliza un conjunto de características similares a las propuestas por Gong [1]. Para clasificar el tipo de región hemorrágica se hace uso de la Red Neuronal MLP.

En la mayoría de trabajos propuestos se utilizaron técnicas de segmentación y clasificación de las potenciales regiones hemorrágicas, sin embargo, desde el punto de vista clínico el problema no consiste en detectar la ubicación de la región hemorrágica sino en cuantificar el tamaño de la misma. A continuación se muestra una metodología tanto para la detección como para la cuantificación que ayude al especialista a mejorar el diagnóstico clínico de HIC.

3 Metodología propuesta

La metodología propuesta consiste en el desarrollo de un Sistema de Visión Artificial el cual mediante técnicas de Procesamiento Digital de Imágenes y una Red Neuronal Artificial pueda detectar y cuantificar automáticamente la HIC. En la Fig. 1 se muestran las diferentes etapas que lleva a cabo el SVA tanto para la identificación de la región hemorrágica, como para la cuantificación de la misma.



Figura 1. Esquema general del Sistema de Visión Artificial.

3.1 Selección del corte axial

El primer paso consiste en revisar cada uno de los cortes axiales del estudio TAC y seleccionar el corte en donde se visualice la HIC con mayores dimensiones. A partir del corte axial seleccionado se realizan las operaciones de segmentación de la masa Intracerebral, detección de potenciales regiones hemorrágicas, clasificación y cuantificación de la región hemorrágica. En la Fig. 2 se muestra la selección del corte axial más significativo de un caso de estudio.

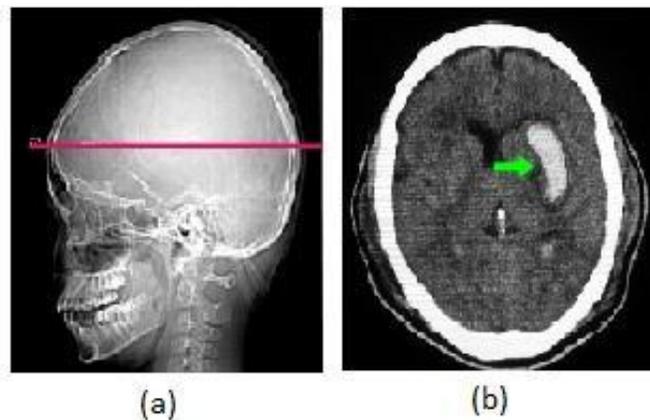


Figura 2. Corte axial con mayor tamaño. En la parte (a) se muestra la altura del corte axial y en la parte (b) la región hemorrágica.

3.2 Segmentación de la masa Intracerebral

En esta etapa se realiza la segmentación de la masa Intracerebral mediante un algoritmo denominado como máxima circularidad (ver Fig. 3). Esta técnica utiliza operaciones de umbralización con α -cortes que permitan encontrar la región de interés con mayor circularidad. En la Fig. 4 se muestra el cálculo de la circularidad para 3 α -cortes.

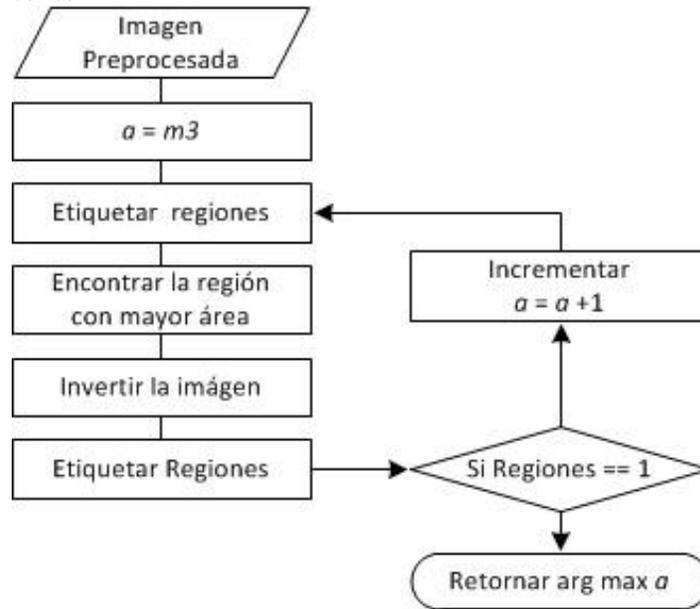


Figura 3. Algoritmo de máxima circularidad.

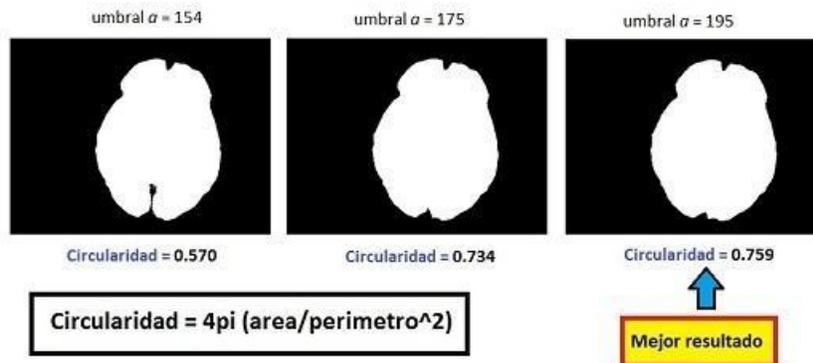


Figura 4. Cálculo de la circularidad con diferentes α -cortes

3.3 Detección de potenciales regiones hemorrágicas

La etapa de segmentación tiene por objetivo separar la masa Intracerebral de las potenciales regiones hemorrágicas. El método consiste en dividir el histograma de la imagen en tres clases mediante el algoritmo de Otsu Multiumbral [6]. Para el caso de umbralización multinivel se debe considerar que existen al menos $M-1$ posibles umbrales, $\{t^1, t^2, \dots, t^{M-1}\}$, los cuales pueden dividir la imagen en M clases: C^1 para $[0, \dots, t^1]$, C^2 para $[t^1+1, \dots, t^2]$, ..., C^i para $[t^{i-1}+1, \dots, t^i]$, ..., y C^M para $[t^{M-1}+1, \dots, L-1]$, en donde los valores óptimos para los umbrales $\{t^1, t^2, \dots, t^{M-1}\}$ se encuentran maximizando σ_B^2 de la siguiente forma:

$$\{t_1^*, t_2^*, \dots, t_{M-1}^*\} = \arg \max \{\sigma_B^2(t_1, t_2, \dots, t_{M-1})\},$$

$$0 \leq t_1 \leq \dots \leq t_{M-1} \leq L-1$$

En donde:

$$\sigma_B^2 = \sum_{k=1}^M \omega_k (\mu_k - \mu_T)^2$$

Definiendo a ω^k y μ^k como:

$$\omega_k = \sum_{i \in C_k} P_i \quad \mu_k = \sum_{i \in C_k} \frac{i \cdot P_i}{\omega_k}$$

En la Fig. 5 se puede observar la salida del algoritmo de segmentación Otsu multiumbral a partir de la masa Intracerebral. Aunque el caso de estudio presenta una región hemorrágica con un área pequeña y un aspecto borroso, la salida del algoritmo de segmentación encuentra de forma satisfactoria la región hemorrágica.

Adicionalmente se puede observar una serie de regiones que se consideran como falsos positivos los cuales se eliminan en la fase de clasificación.

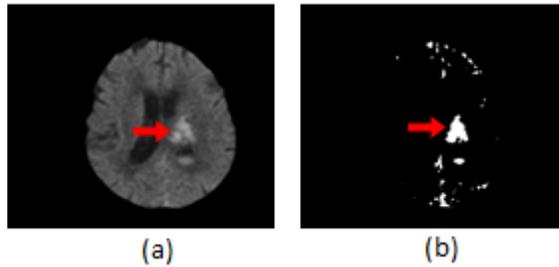


Figura 5. Segmentación mediante Otsu multiumbral. En la parte (a) se visualiza la masa Intracerebral y en la parte (b) el resultado de la segmentación automática.

3.4 Extracción de características

Para realizar la identificación de las regiones hemorrágicas, se hace uso de las características visuales como el tamaño, forma y la textura de las potenciales regiones hemorrágicas.

Características geométricas. Gong [1] propuso el uso de siete medidas geométricas como características principales para realizar una clasificación de las potenciales regiones hemorrágicas. Las características que se utilizaron fueron el área, la longitud del eje mayor, la longitud del eje menor, excentricidad, solidez, extensión y adyacencia al cráneo.

Estadísticos de primer orden. Uno de los métodos más sencillos para describir una textura es el uso de momentos estadísticos del histograma de la región que se esté analizando. Por cada una de las potenciales regiones hemorrágicas se extraen un total de 7 parámetros estadísticos: media, varianza, suavidad, asimetría, curtosis, energía y entropía.

Estadísticos de segundo orden. Para realizar el análisis de las texturas de cada una de las potenciales regiones hemorrágicas se hace uso de la matriz coocurrencia GLCM, a partir de la cual se obtienen los siguientes 7 parámetros estadísticos: máxima probabilidad, energía, entropía, varianza, correlación, homogeneidad y contraste.

3.5 Clasificación mediante Red Neuronal Artificial

En la etapa de clasificación se utiliza una Red Neuronal Artificial MLP para la clasificación de potenciales regiones hemorrágicas en dos clases: HIC y tejido normal. En la Fig. 6 se muestra la arquitectura de la RNA la cual consiste en 21 neuronas en la capa de entrada, 40 neuronas en la capa oculta y 1 neurona en la capa de salida. La metodología para el entrenamiento y prueba de la red MLP así como la elección de los parámetros de configuración se tomaron a partir del trabajo realizado por Illán y Medina [2].

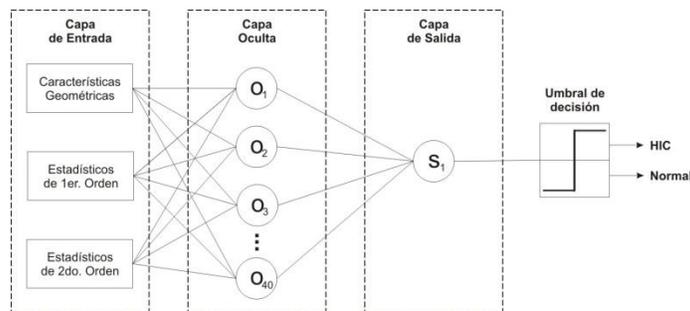


Figura 6. Arquitectura del modelo de clasificación MLP.

El número total de muestras fue de 1028 que corresponden con el número de regiones binarias que se encontraron en la etapa de detección de poten-

ciales regiones hemorrágicas. La distribución del conjunto de datos se muestra con más detalle en la Tabla 1.

Tabla 1. Distribución del conjunto de datos

Conjunto	Clase	Muestras
Entrenamiento	Hemorragia	128
	Normal	34
	Total	570
Validación	Hemorragia	64
	Normal	17
	Total	285

3.6 Cuantificación de la región hemorrágica

Normalmente las HIC tienden a formar coágulos esféricos cuyo volumen puede ser calculado mediante una técnica conocida como ABC2 [8]. Para realizarla cuantificación de las dimensiones de la región hemorrágica se calcula un modelo geométrico elíptico que circunscribe a la región hemorrágica. Posteriormente se calcula el volumen de la región hemorrágica mediante la siguiente expresión:

$$volumen = \frac{A \cdot B \cdot C}{2}$$

En donde A es el segmento más largo que se puede trazar en la región hemorrágica, B es el segmento perpendicular al segmento A , ver Fig. 7, y C es el número aproximado de cortes axiales con grosor aproximado de 8mm.

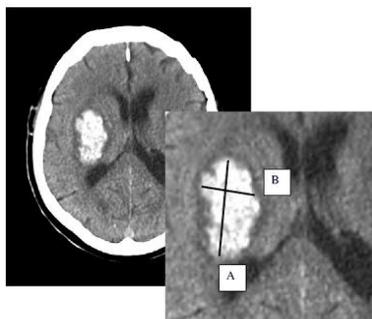


Figura 7. Trazo del segmento A y B para el cálculo volumétrico de la HIC.

4 Resultados experimentales

El conjunto de datos experimental está conformado por dos grupos: moderado y complejo. Estos grupos definen la complejidad del diagnóstico clínico desde el punto de vista médico. Cada grupo está compuesto por 10 casos de estudio los cuales fueron seleccionados por el especialista. En la Fig. 8 se muestra la comparación entre la segmentación y cuantificación manual realizada por el especialista y la segmentación y cuantificación automática realizada por el Sistema de Visión Artificial.

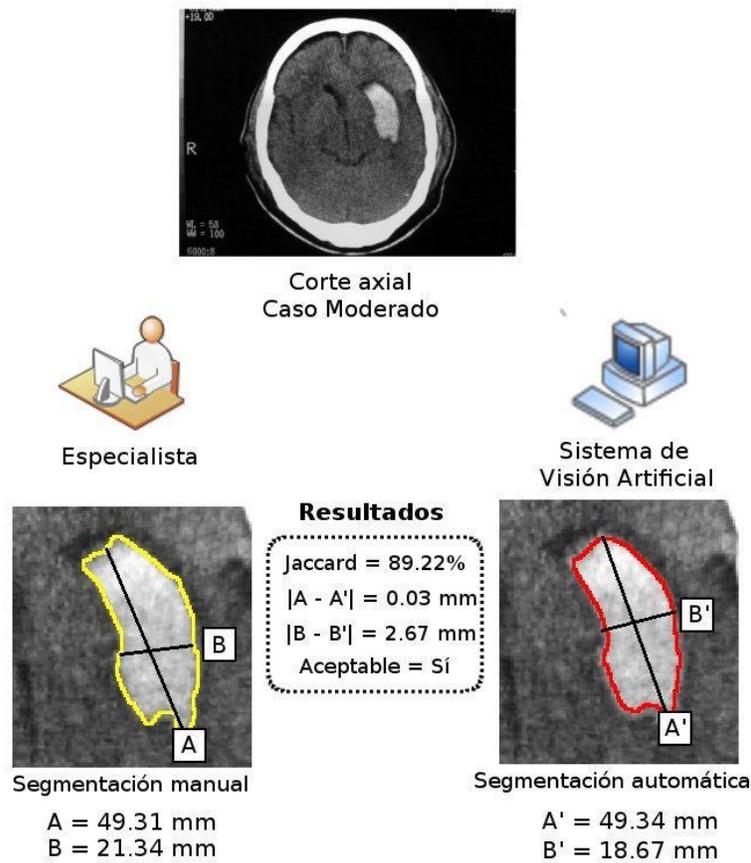


Figura 8. Comparación de la segmentación y cuantificación de HIC

Los casos de diagnóstico moderado incluyen los estudios tomográficos en donde se visualiza una HIC de forma clara y un contorno bien definido que se distingue con facilidad del resto de las estructuras intracerebrales. En la Tabla 2 se muestran los resultados de la segmentación y cuantificación de HIC en los casos de diagnóstico moderado.

Tabla 2. Resultados de la segmentación y cuantificación de HIC.

Casos de diagnóstico moderado.

Caso de Estudio	Especialista		SVA		Diferencia		Jaccard	Aceptable
	A	B	A'	B'	A-A'	B-B'		
1	49.31	21.34	49.34	18.67	0.03	2.67	89.22%	Sí
2	26.16	15.65	34.90	25.50	1.55	2.50	80.85%	Sí
3	65.87	40.13	62.50	36.90	3.37	3.23	90.25%	Sí
4	56.87	26.50	54.10	22.15	2.77	4.35	88.75%	Sí
5	39.82	19.20	36.61	17.20	3.21	2.00	89.40%	Sí
6	75.30	43.75	70.56	39.80	4.74	3.95	83.63%	Sí
7	47.65	22.38	45.83	20.64	1.82	1.74	80.79%	Sí
8	49.87	21.59	46.37	19.43	3.50	1.94	84.98%	Sí
9	53.89	26.84	48.71	23.90	5.18	2.94	88.65%	Sí
10	62.30	29.47	57.43	26.75	4.87	2.71	89.91%	Sí

Los casos de diagnóstico complejo incluyen los estudios tomográficos en donde se visualiza una HIC de forma borrosa con un contorno irregular que dificulta la identificación de la región hemorrágica. En la Tabla 3 se muestran los resultados de la segmentación y cuantificación de HIC en los casos de diagnóstico complejo.

Tabla 3. Resultados de la segmentación y cuantificación de HIC.

Casos de diagnóstico complejo.

Caso de Estudio	Especialista		SVA		Diferencia		Jaccard	Aceptable
	A	B	A'	B'	A-A'	B-B'		
1	76.8	42.72	90.93	28.17	14.13	14.55	65.11%	No
2	15.67	7.50	27.12	14.30	11.45	6.80	60.30%	No
3	22.54	10.32	35.10	16.89	12.56	6.57	70.45%	No
4	18.90	8.50	25.20	14.45	6.30	5.95	58.90%	Sí
5	24.90	13.70	26.74	11.89	1.84	1.81	63.46%	Sí
6	16.72	7.50	27.25	14.87	10.53	7.37	67.89%	No
7	17.60	9.82	26.62	15.63	9.02	5.81	61.50%	Sí
8	18.83	9.40	29.56	13.72	10.73	4.32	71.13%	No
9	22.52	13.20	35.25	18.74	12.73	5.54	69.82%	No
10	19.50	9.45	31.20	13.28	11.70	3.83	57.90%	No

5 Conclusiones y trabajo futuro

La presente investigación propone una metodología para la identificación y cuantificación de HIC en estudios de TAC. En la metodología se propuso el algoritmo de Máxima Circularidad para segmentar la masa intracerebral, así como el uso del algoritmo de Otsu Multiumbral para la detección potenciales regiones hemorrágicas. Adicionalmente se propone que el análisis de resultados debe llevarse a cabo mediante la comparación de la segmentación manual y automática, así como de la diferencia de la longitud de los segmentos AB realizada por el especialista y los segmentos A'B' realizada por el SVA.

Los resultados muestran que el método propuesto es capaz de identificar y cuantificar de forma adecuada la región hemorrágica en los casos de diagnóstico

moderado. Sin embargo, existen algunos casos de diagnóstico complejo en los cuales por la irregularidad del contorno de la región hemorrágica la segmentación de la región hemorrágica no se asemeja a la segmentación manual realizada por el especialista. A pesar de este tipo de dificultades, los casos de diagnóstico complejo aparecen de forma menos frecuente en la práctica clínica por lo que el método se considera adecuado desde el punto de vista del especialista.

En trabajos futuros se realizará un estudio de los niveles de la densidad de los tejidos y estructuras intracerebrales que permitan determinar con mayor precisión la forma y ubicación de la región hemorrágica en los diagnósticos de casos complejo.

Referencias

1. Gong T.: Classification of CT Brain Images of Head Trauma. Proceedings of the 2nd IAPR international conference on Pattern recognition in bioinformatics, pp. 401-408, 2007
2. Illán G. E., Medina G. B.: Clasificación de Hemorragias Intracerebrales en Tomografía Axial Computarizada mediante Redes Neuronales Artificiales. XXXIV Congreso Nacional de Ingeniería Biomédica, pp 33-36, 2011
3. Kabara S. A.: CT Image Texture analysis of Intracerebral Hemorrhage. Proceedings of the 2003 Finnish Signal Processing Symposium, pp. 190-194, 2003
4. Lee T. H.: A Segmentation of CT Brain Images Using Unsupervised Clusterings. The Visualization Society of Japan Journal of Visualization, pp. 131-138, 2009.
5. Li Y., Hu Q.: A hybrid approach to detection of brain hemorrhage candidates from clinical head CT scans. Sixth International Conference on Fuzzy Systems and Knowledge Discovery, pp. 361-365, 2009.
6. Liao C. C.: A Knowledge Discovery Approach to Diagnosing Intracranial Hematomas on Brain CT: Recognition, Measurement and Classification. Proceedings of the 1st international conference on Medical biometrics, pp. 73-82, 2009.
7. Liao P.S.: A Fast Algorithm for Multilevel Thresholding. Journal of Information Science and Engineering, pp. 713-721, 2001.
8. Lozano J.C.: Aspecto clinicos y epidemiologicos del trauma craneo encefalico. Sistema Nacional de Vigilancia Epidemiológica, pp. 1-4, 2006.
9. Ramana K.V., Korrapati R.: Neural Network Based Classification and Diagnosis of Brain Hemorrhages. International Journal of Artificial Intelligence and Expert Systems (IJAE), pp.7-25, 2010.

Identificación de displasias en análisis de muestras de Papanicolaou usando ACO.

Maricela Gress-Roldán¹, Edmundo Bonilla-Huerta², Roberto Morales-Caporal³ y José Crispín Hernández-Hernández⁴.

^{1,2,3,4} Laboratorio de Investigación en Tecnologías Inteligentes,
Instituto Tecnológico de Apizaco,
Apizaco, Tlaxcala, México
{maricela,edbonn,moralescaporal}@itapizaco.edu.mx, josechh@yahoo.com
Paper received on 30/07/12, Accepted on 16/08/12.

Resumen. El Papanicolaou es una prueba muy sencilla, la desventaja es que esta sujeta a errores a la hora de la interpretación de las muestras durante su observación en el microscopio, debido a esta problemática se han desarrollado herramientas de cómputo, donde la etapa de segmentación llega a ser una de las etapas más importantes y difíciles durante el proceso de identificación de las imágenes. En este artículo se presenta una metodología que se apoya en el Procesamiento Digital de Imágenes (PDI) y el algoritmo de Optimización por Colonia de Hormigas (ACO por sus siglas en inglés) para identificar y extraer las características más importantes de las células examinadas con la finalidad de obtener una tasa de reconocimiento satisfactoria.

Palabras Clave: Procesamiento Digital de Imágenes, Citología, ACO, reconocimiento automático.

1 Introducción.

El cáncer cérvicouterino (CaCu) afecta el sistema reproductivo de la mujer. La causa predominante de éste se debe a ciertos derivados del virus del papiloma humano (VPH), una de varias enfermedades venéreas. Por lo general, el sistema inmunológico del organismo evita que el virus afecte de forma grave al sistema reproductor, sin embargo en algunos casos éste logra sobrevivir el suficiente tiempo para afectar las células localizadas en la superficie del cérvix, transformándolas en cancerígenas. Cerca del 50% de los casos de cáncer cérvico se da en mujeres cuya edad va de los 35 a los 55 años [9].

El CaCu es el resultado de un proceso dinámico que se inicia probablemente a partir de una infección viral, aunada a factores de riesgo, pasando por una etapa pre invasora. Este proceso es asintomático y de evolución muy lenta, lo que permite diagnosticar y tratar oportunamente este padecimiento. El CaCu tiene una evolución bien identificada, con la presencia de lesiones precursoras denominadas displasias, las cuales se clasifican en leves, moderadas y severas, según la gravedad de las alteraciones.

2 Estado del arte.

El uso de herramientas computacionales en la medicina es un tema de suma importancia hoy en día, los análisis de muestras de Papanicolaou han sido un problema que se ha tratado de atacar en muchos trabajos utilizando diferentes técnicas de visión artificial, por ejemplo en 2009, Gallegos [3] buscó clasificar el tipo de células y obtuvo como resultado el 76.0% utilizando el método Naive Bayes, y del 83.2% de precisión con la red Perceptrón Multicapa. A principios de 2011, Pérez [4] trató de resolver el problema de segmentación de imágenes de células con Sobel. Los resultados fueron aceptables ya que de las 60 imágenes de prueba, fueron segmentadas correctamente 54, que es una tasa de segmentación del 90 %.

En el artículo presentado por F. Gallegos-Funes [5] en el año de 2009, se usó una red neuronal con la función de base radial tipo M para el proceso de reconocimiento automático de muestras de Papanicolaou. Las características extraídas fueron área, perímetro, circularidad, relación núcleo/citoplasma y finalmente los valores máximos y mínimos de brillo en el núcleo. Los resultados muestran que la red obtuvo 20% de error.

Durante el presente año se han desarrollado proyectos donde podemos encontrar el de I. Muhimmah [6] quien ha desarrollado una metodología basada en Clustering fuzzy C-Means. Los resultados de clasificación arrojaron los resultados en la tasa de sensibilidad de 90.86% y una tasa de especificación de 81.62% en comparación con el experto. En el artículo donde B. Mahanta [7] presentó una metodología capaz de extraer las características que le interesa a los patólogos (tamaño del núcleo, forma del núcleo y distribución del núcleo celular). A partir de las imágenes binarizadas se logró obtener cada una de las características buscadas.

El método de Ant Colony Optimization para segmentación de imágenes ha sido probado por algunos autores quienes han demostrado que este algoritmo resulta ser una buena alternativa para una de las tareas más difíciles en las etapas de la visión artificial.

Tian [8] en el año 2008 planteó una solución a base de segmentación de imágenes por medio de Optimización por Colonia de Hormigas (ACO). Y en 2010, Bateria [2] propuso extraer los bordes de imágenes utilizando ACO (Ant Colony Optimization), ya que la segmentación basada en ACO, donde afirmaron que este algoritmo tiene el potencial de superar las limitaciones de los métodos convencionales. Los resultados se que se obtuvieron fueron satisfactorios, y el tiempo de respuesta fue de 3 segundos.

3 Metodología.

La metodología propuesta se muestra en el diagrama de la figura 1, posteriormente se describe cada una de las etapas que la componen.

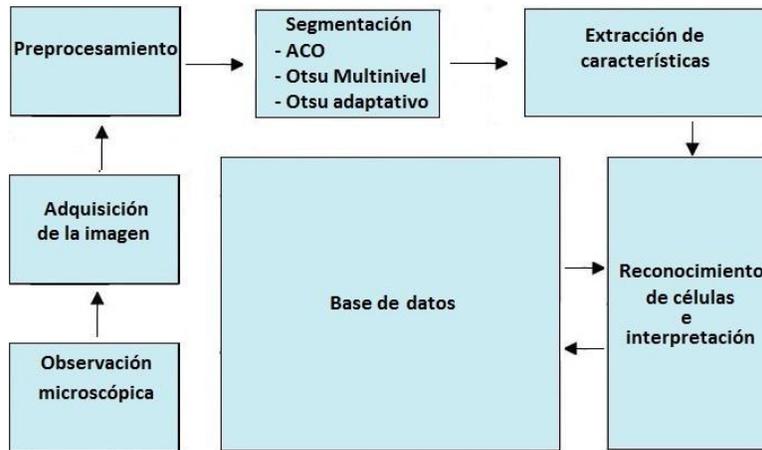


Figura 1. Metodología propuesta.

3.1 Adquisición de la imagen.

Para realizar el análisis es necesario digitalizar las muestras, las imágenes fueron obtenidas de un banco de imágenes proporcionadas por el Instituto Mexicano del Seguro Social (IMSS) del laboratorio de citología del Hospital General de zona No.1, Tlaxcala, las imágenes son de formato JPG, tomadas con zoom óptico que varía de 40x a 100x con dimensión de 2592 X 1944 píxeles.

3.2 Preprocesamiento.

En primer lugar se busca eliminar la sobre-iluminación que se encuentra en el área de interés haciendo una resta de imágenes (original menos imagen del escenario sin célula), posteriormente la imagen se convierte a formato de color Lab donde se elige el canal de color L que muestra las imágenes de manera más clara y definida, y se realiza una mejora de la imagen usando un filtro de punto medio [11].

3.3 Segmentación.

El objetivo de la segmentación es obtener una partición de la imagen en regiones coherentes como paso previo al análisis de su contenido. En este artículo se propone el uso del algoritmo ACO que consiste en las siguientes etapas (ver figura 2).

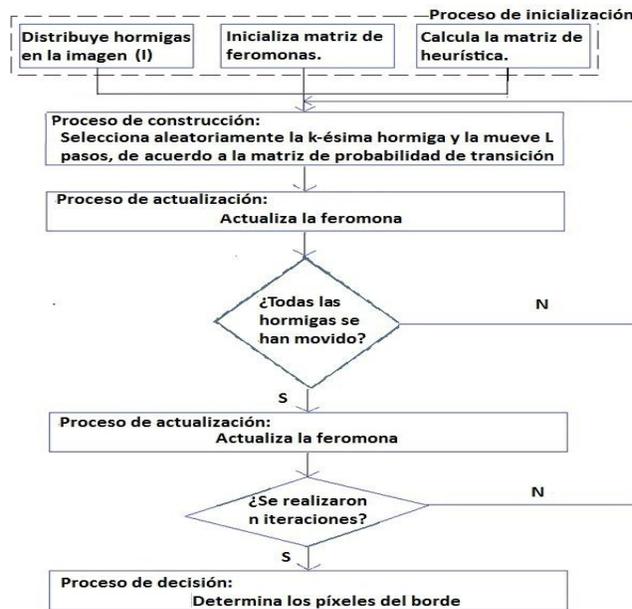


Figura 2. Diagrama descriptivo de ACO.

Las fórmulas y valores asignados para cada una de las etapas se describen a continuación:

***Proceso de inicialización:** Esta etapa se compone de las siguientes sub etapas:

* *Distribuye las hormigas en la imagen:* La distribución de las hormigas se realiza de manera aleatoria, sobre toda la imagen. El número de hormigas que se distribuyen en la imagen se calcula de la siguiente manera:

$$K = \lfloor \sqrt{M1 \cdot M2} \rfloor \quad (1)$$

Donde:

-M1=Largo de la imagen.

-M2=Ancho de la imagen.

**Inicializa matriz de feromonas:* Se inicializa la matriz de las feromonas, con un valor fijado por el usuario, el valor asignado en este trabajo a esta constante se muestra en la tabla 1.

*Calcula la matriz heurística:

$$n_{i,j} = \frac{1}{Z} V_c(I_{i,j}) \quad (2)$$

Donde:

- $Z = \sum_{i=1:M1} * \sum_{i=1:M2}$.
- $I_{i,j}$ = valor de intensidad del pixel en su posición (i,j).
- $V_c(I_{i,j})$ = función de un grupo local de píxeles (vecinos), su valor depende de las variaciones de los valores de intensidad de la imagen.

*Proceso de construcción: En este proceso una hormiga es aleatoriamente seleccionada de k hormigas, esta hormiga será consecutivamente movida sobre la imagen por L pasos o movimientos. Estas hormigas se mueven del nodo (l,m) a su nodo vecino (i,j) de acuerdo a la probabilidad de transición dada por:

$$P_{(l,m),(i,j)}^n = \frac{(T_{i,j}^{(n-1)})^\alpha (\eta_{i,j})^\beta}{\sum_{(i,j) \in \Omega(l,m)} (T_{i,j}^{(n-1)})^\alpha (\eta_{i,j})^\beta} \quad (3)$$

Donde:

- $T_{i,j}^{(n-1)}$ = Valor de la feromona en el nodo (i,j).
- $\Omega(l,m)$ = Es la vecindad de nodos del nodo (l,m).
- $\eta_{i,j}$ = Información heurística del nodo (i,j)
- α Y β = La influencia de la matriz de feromonas y la matriz heurística.

*Proceso de actualización: En esta etapa se lleva a cabo dos operaciones para actualizar la matriz de feromonas. La primera actualización se realiza después del movimiento de cada hormiga dentro de cada paso de construcción. Cada componente de la matriz de feromonas es actualizado de acuerdo a:

$$T_{i,j}^{(n-1)} = \begin{cases} (1 - \rho) \cdot T_{i,j}^{(n-1)} + \rho \cdot \Delta_{i,j}^{(k)} & \text{Si (i,j) es visitado por la k-} \\ T_{i,j}^{(n-1)} & \text{ésima hormiga actual.} \\ & \text{En otro caso} \end{cases} \quad (4)$$

Donde:

- ρ = tasa de evaporación de la feromona.
- $\Delta_{i,j}^{(k)} = \eta_{i,j}$ es determinada por la matriz de heurística.

La segunda actualización es llevada a cabo después del movimiento de todas las hormigas dentro de cada paso de construcción de acuerdo a:

$$T^{(n)} = (1 - \varphi) \cdot T^{(n-1)} + \varphi \cdot T^{(0)} \quad (5)$$

Donde:

- φ es el coeficiente de decadencia de la feromona.

***Proceso de decisión:** En este paso, se realiza un proceso de decisión binario en cada píxel para determinar si éste es un borde o no, por la aplicación de un umbral (T) en la matriz final de feromonas. El umbral (T) es determinado por el cálculo de la media de los valores del píxel de la imagen, el cálculo se realiza a través del histograma de la imagen.

$$E_{i,j} = \begin{cases} 1, & \text{Si } T_{i,j}^{(N)} \geq T^l \\ 0 & \text{En otro caso} \end{cases} \quad (6)$$

Una vez encontrada la región del citoplasma, se calculan sus coordenadas dentro de la imagen, esto con la finalidad de realizar posteriormente la detección de color en el área de interés. Los parámetros utilizados en ACO para segmentar las imágenes fueron elegidos de acuerdo a los resultados obtenidos en pruebas previamente realizadas, y son los siguientes (ver tabla 1):

Tabla 1. Parámetros utilizados en ACO

Parámetro	Cantidad
Número de hormigas	141
Valor de inicial de la feromona	0.0000001
Factor de peso de la información de la feromona	1
Factor de peso de la información heurística.	1
Conectividad de la vecindad	8
Tasa de evaporación de la feromona	0.1
Decadencia del coeficiente de feromonas	0.05
Número de movimientos de las hormigas en cada paso de construcción	60

3.4 Extracción de características.

En esta etapa el objetivo es extraer las características más importantes del núcleo y citoplasma (área, perímetro, circularidad, relación núcleo/citoplasma [10] y color)

Por ello en la metodología propuesta para determinar el color de la célula, se extraen los valores del tono en la imagen original. Posteriormente se calcula la longitud de onda dominante comparando el valor del píxel con el valor asignado a cada uno de los colores que forman parte del espectro de color visible en el que se descompone la luz blanca.

3.5 Reconocimiento de células e interpretación.

Se utiliza una red neuronal de retropropagación, en la cual, la capa de entrada será del tamaño de número de características obtenidas, mientras que la capa de salida consta de 6 neuronas que corresponden al número de clases a las que puede pertenecer la célula, mientras que para la capa intermedia se realizaron pruebas con diferentes número de neuronas, considerando la que obtiene mejores resultados.

4 Resultados.

Imagen original. Las imágenes fueron catalogadas previamente por un citotecnólogo de la siguiente manera: cáncer in situ, carcinoma invasor, displasias graves, displasias moderadas, displasias leves. En la figura 3 se muestran dos imágenes de ejemplo (denominadas muestra 1 y 2 respectivamente) donde se observa la sobreiluminación en el área de interés (célula).

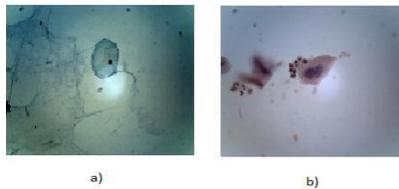


Figura 3. Imágenes observadas en el microscopio.

Preprocesamiento. Los resultados de la etapa del preprocesamiento se muestran en la figura 4, donde los incisos *a* y *e* muestran las imágenes convertidas a formato Lab, incisos *b* y *f* extracción del canal L, incisos *c* y *g* extracción del canal a y finalmente en los incisos *d* y *h* la extracción del canal b de la muestra 1 y muestra 2 respectivamente).

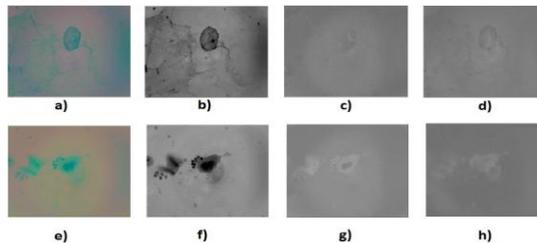


Figura 4. Conversión a formato Lab y extracción de cada uno de sus canales.

Segmentación. Los resultados de la segmentación usando ACO se muestran en la figura 5, donde el inciso *a* corresponde a la muestra 1 y el inciso *b* corresponde a la muestra 2.

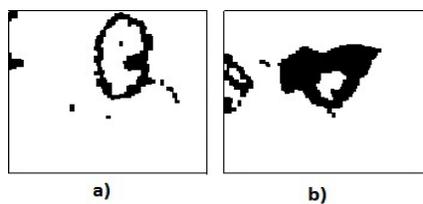


Figura 5. Imágenes segmentadas con ACO.

Extracción de características. En la tabla 2 se muestran las características obtenidas y los resultados de la identificación de color en la muestra 1 se muestran en la figura 6.

Tabla 2. Parámetros calculados en la muestra 1.

Características	Citoplasma	Núcleo
Área	2004	41
Perímetro	218	34
Circularidad	0.530	0.446
Relación núcleo/citoplasma	0.006	0.006

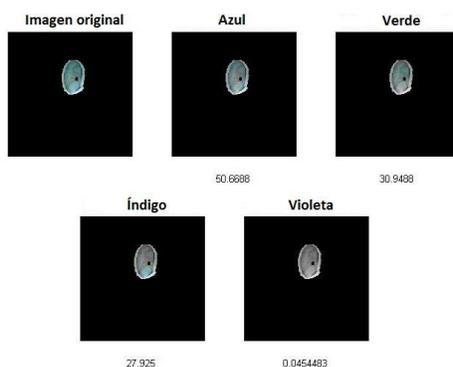


Figura 6. Colores identificados en muestra 1.

También se muestran los resultados de la extracción de características de la muestra 2 que se observan en la tabla 3 y en la figura 7 se muestran los resultados de la identificación de color.

Tabla 3. Parámetros calculados en la muestra

Características	Citoplasma	Núcleo
Área	1751	446
Perímetro	292	106
Circularidad	0.258	0.499
Relación núcleo/citoplasma	0.276	0.276

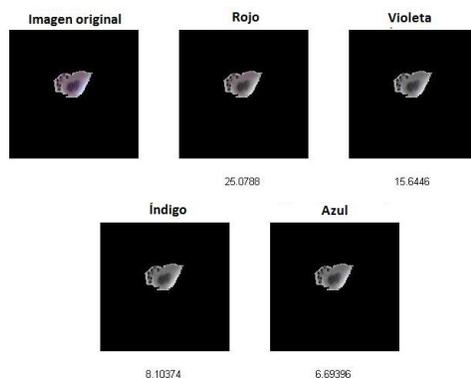


Figura 7. Colores identificados en muestra 2.

Reconocimiento de células e interpretación. Los resultados obtenidos en esta etapa se muestran en la siguiente matriz de confusión (ver tabla 4) donde se puede observar el desempeño de la red al clasificar las imágenes.

Tabla 4. Matriz de confusión del desempeño de la red neuronal de retropropagación.

Muestras	Número de muestras	Cáncer in situ	Carcinoma invasor	Displasias graves	Displasias leves	Displasias moderadas	Otras	%
Cáncer in situ	10	8	1	-	-	-	1	80%
Carcinoma Invasor	5	1	3	-	1	-	-	60%
Displasias Graves	5	1	-	3	-	-	-	80%
Displasias Leves	10	1	-	-	8	-	1	80%
Displasias Moderadas	10	1	-	-	-	9	-	90%
Otras	20	-	1	2	1	1	15	75%

5 Conclusiones y trabajos futuros.

Se logró extraer de manera exitosa cada una de las características más relevantes para realizar la clasificación de nuevas imágenes donde se utilizó la red backpropagation obteniendo una tasa de clasificación del 77.5%, por ello se desea obtener más muestras y probar la efectividad de otros clasificadores tales como SVM y otras redes neuronales hasta obtener niveles mucho más elevados de clasificación para poder garantizar la efectividad y confiabilidad de los resultados.

Referencias

1. Olvera, A.: Citodiagnóstico cervicovaginal. Trillas. 2009.
2. Instituto Mexicano del Seguro Social, cáncer cérvicouterino, <http://www.imss.gob.mx/salud/Cancer/tipos/Pages/cancercervico.aspx>, octubre 2009
3. E. Gallegos: Aplicación de la teoría Wavelets en el análisis de muestras citológicas de Papanicolaou, In Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México. 2009.
4. Z. Perez: "Segmentación de células traslapadas en imágenes microscópicas de cérvix con PDI", In Instituto Tecnológico de Apizaco, División de Estudio de Posgrado e Investigación, Maestría en Sistemas Computacionales. Apizaco, Tlaxcala, México. 2011.
5. F. Gallegos, M. Gómez, J. López y R. Cruz: "Rank M-Type Radial Basis Function (RMRBF) Neural Network for Pap Smear Microscopic Image Classification", In Apeiron, Vol. 16, No. 4, Octubre 2009.
6. I. Muhimmah, R. Kurniawan y Indrayanti: "Automatic Epithelial Cells Detection of Pap smears images using Fuzzy C-Means Clustering", In 2012 4th International Conference on Bioinformatics and Biomedical Technology IPCBEE vol. 29 IACSIT Press, Singapore, 2012.
7. L. Mahanta, D. Nath y C. Nath: "Cervix Cancer Diagnosis from Pap Smear Images Using Structure Based Segmentation and Shape Analysis", In Journal of Emerging Trends in Computing and Information Sciences, ISSN 2079-8407, 2012.
8. J. Tian, W. Yu y S. Xie: "An Ant Colony Optimization Algorithm For Image Edge Detection". In IEEE Explore, 2008
9. A. Bateria y C. Oppus: "Ant Colony Optimization for Image Edge Detection" In International Journal of Circuits, Systems and Signal processing, pp 25-33. 2010.
10. C. Lacruz y J. Fariña: "Citología ginecológica de Papanicolaou a Bethesda", In Editorial Complutense, ISBN 84-7491-717-4.
11. B. Shinde y A. Dani: "Noise Detection and Removal Filtering Techniques in Medical Images", In International Journal of Engineering Techniques in Medical Images, ISSN:2248-9622. Vol. 2, Issue 4, pp. 311-316. July-August 2012.

Segmentación en imágenes microscópicas

José A. Márquez¹, Beatriz A. Sabino¹, Jesús M. Campos¹ y Manuel I. Martín²

¹Universidad de la Cañada, Teotitlán de Flores Magón, Oaxaca, México
{albertomarquez, beatriz_sabino, campos}@unca.edu.mx

²Benemérita Universidad Autónoma de Puebla-FCC, Puebla, México
mmartin@cs.buap.mx

Paper received on 11/07/12, Accepted on 05/09/12.

Resumen. En este documento se presenta un método de segmentación usando la librería OpenCV para imágenes microscópicas de diferentes células, el proceso se realiza mediante un conjunto de funciones llamado batería. Se observó que la batería propuesta proporciona un mejor resultado en objetos con relleno no uniforme, como es el caso de las muestras que se analizaron, mismas que son de importancia en el área de la Microbiología, Biotecnología, entre otras. Este trabajo es parte de un proyecto de investigación que se realiza en la Universidad de la Cañada, el cual pretende realizar un sistema que contabilice la levadura *Saccharomyces cerevisiae*, a través de una imagen.

Palabras Clave: Procesamiento de Imágenes, Segmentación, Microbiología, *Saccharomyces cerevisiae*.

1 Introducción

El manejo de las imágenes digitales se ha convertido en estos años en un tema de interés en las diferentes áreas de las ciencias naturales, ciencias médicas, en las aplicaciones tecnológicas, entre otras. En la actualidad es posible explotar plataformas de bajo costo y obtener resultados de gran calidad y con ello, crear aplicaciones de gran utilidad, versátiles y flexibles, así como aplicaciones de software de propósito específico para atender las necesidades del usuario y/o especialistas. Algunas áreas donde el procesamiento digital de imágenes (PDI) ayuda a analizar, deducir y tomar decisiones, están las siguientes: Medicina, Biología, Fisiología, Biometría, Astronomía, Ciencias Ambientales, Robótica, Metalurgia, Física, Electrónica y Microbiología.

Dentro de la Microbiología es común hablar de conteo celular, esto se lleva a cabo por diferentes métodos, uno de ellos es utilizando una cámara de Newbauer, éste se relaciona con la exactitud y la velocidad, ya que depende de la experiencia de la persona que examina la muestra debido a que se efectúa de forma manual, resul-

tando largo y cansado (visualmente hablando) al efectuar múltiples conteos, cabe señalar que esto se realiza sin ningún tipo de software que pueda auxiliar al técnico, ya que es el que realiza el seguimiento del crecimiento celular. Además, se consideran otros parámetros, en el caso de la fermentación se toma en cuenta la viabilidad celular y la contaminación bacteriana [1-3].

El análisis automático o semiautomático de imágenes plantea varios retos técnicos respecto al procesamiento requerido debido a factores como:

1. Iluminación no homogénea.
2. Oclusión de objetos.
3. Variación en la forma, el tamaño y la orientación de los objetos.
4. Variación de intensidades.

Debido a estas características, los sistemas de análisis de imágenes de microscopía generalmente plantean enfoques híbridos en el área de segmentación de imágenes y Reconocimiento de Patrones (RP) para resolver el problema particular que enfrentan. La mayoría de los casos el esquema general de trabajo para analizar una imagen es el siguiente:

- Realizar un preprocesamiento a la imagen para disminuir el ruido y suavizar los bordes de los objetos buscados.
- Realizar una clasificación de la imagen, detectando diferencias de manera aproximada los distintos objetos de la imagen.
- Identificar bordes o contornos de los objetos de interés.
- Segmentar los objetos usando la información obtenida.

Debido a la particularidad de cada problema, los métodos usados en cada fase difieren considerablemente, pasando por aplicaciones de teoría de conjuntos, filtros de imágenes métodos de segmentación, métodos de aprendizaje computacional, entre otras.

Ahora bien, existe software que facilita en gran manera el aprendizaje e implementación de diferentes técnicas de visión por computadora, tal es el caso de OpenCV (*Open Computer Vision Library*) que es una librería cuyas funciones están principalmente dirigidas a las técnicas del procesamiento de imágenes, además de ser código abierto, dicha librería proporciona funciones de alto nivel para el procesamiento de imágenes como: operaciones básicas de apertura, creación, acceso a píxeles, filtrado, reducción mediante muestreo, cambio de espacio de color, entre otras [4 y 5].

Para este trabajo se utilizó la librería antes mencionada, las pruebas incluyen imágenes microscópicas de: 2 cepas de levadura (*Saccharomyces cerevisiae*), la primera fue aislada en la Región de la Cañada Oaxaqueña con aplicación biotecnológica para la producción de etanol (UCT1), y la segunda es una levadura comercial usada para la panificación (LC2). Ambas cepa se llevaron a cabo en un cultivo por lote en matraces Erlenmeyer de 500 mL con 300 mL de medio sintético (Glucosa 70g/L, KH_2PO_4 5g/L, $(\text{NH}_4)_2\text{SO}_4$ 2g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0.4g/L y extracto de levadura 1 g/L) a 30°C y 200 rpm, la muestra fue obtenida a las 10 horas donde la levadura alcanza su máxima velocidad de crecimiento; otra de las imágenes se obtuvo de una muestra de pulque de la Región de la Cañada Oaxaqueña, el cual se observó la presencia de levaduras, bacterias (bacilos y cocos) y filamentos de hongos (UCP3). Esto fue investigada con la finalidad de poder utilizar cualquier muestra que contenga

levadura y poder reconocer a *Saccharomyces spp.* Además se incluyó una muestra de sangre enfocando a los glóbulos rojos (MS4).

En este trabajo se presenta un método de segmentación en imágenes microscópicas usando la librería de OpenCV, proporcionando una serie de filtros que sirvieron para detectar los objetos de estudio.

2 Segmentación

La segmentación de una imagen, es un proceso por el cual una imagen digital se subdivide en partes, zonas u objetos disyuntos que la constituyen; es decir, se agrupan los píxeles en función del nivel de intensidad luminosa, el nivel al que se lleva a cabo la subdivisión depende del problema a resolver. Muchas de las veces es necesario separar en una imagen el fondo del objeto para obtener información; por ejemplo, en el análisis de la imagen de un documento se requiere extraer caracteres y líneas del fondo del papel, a este proceso se le llama binarización, puesto que clasifica cada píxel de la imagen en dos clases [6].

La librería OpenCV contiene una diversidad de funciones que permiten el PDI, para este trabajo, se utilizaron un conjunto de filtros que permiten segmentar la imagen separando los objetos de interés, en este caso las células de las diversas muestras.

2.1 Filtros

La aplicación de filtros o funciones permite obtener la información solamente de los objetos de estudio contenidos en la imagen, mismo que facilita la segmentación [7], la serie de filtros conforma una batería final con la cual se trabaja y compara los resultados obtenidos.

El filtrado es una técnica para modificar o mejorar una imagen, éste puede resaltar o atenuar algunas características. A continuación se describen algunos filtros para la segmentación de microorganismos en una muestra de pulque de la Región de la Cañada Oaxaqueña (UCP3), ya que la librería de OpenCV ofrece diferentes funciones, sólo se mostrarán las que proporcionaron mejores resultados para la segmentación.

Umbralización. Es el más sencillo de los filtros implementados, es conocido como umbralización fija (*threshold*) [8]. En su forma más simple, la imagen binaria resultante $B(i, j)$ se define a partir de la versión en escala de grises de la imagen digital original $I(i, j)$, en función de un valor U que corresponde al umbral de separación seleccionado.

$$B(i, j) = \begin{cases} 0, & \text{si } I(i, j) \geq U \\ 1, & \text{si } I(i, j) \leq U \end{cases} \quad (1)$$

En dicha fórmula el 0 indica el color blanco y 1 el color negro, la función de umbralización de OpenCV agrega un segundo valor de umbral, de manera que se consideran blancos aquellos píxeles cuyo nivel de gris encuentre dentro del rango

definido por los dos valores del umbral U_1 , y U_2 . En la Fig. 1 se presenta la muestra de pulque de la Región de la Cañada.

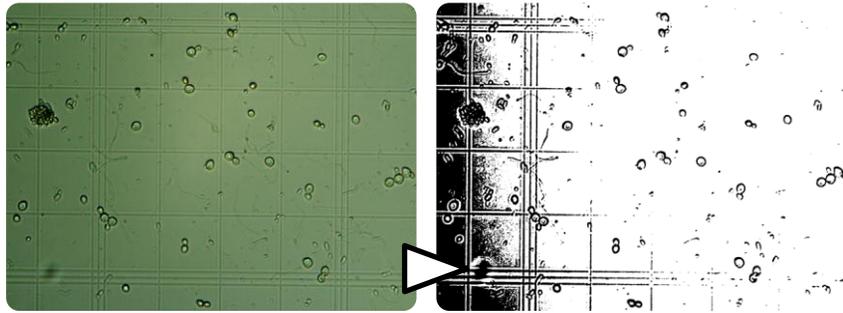


Figura 1. Muestra original (imagen izquierda); aplicación de la Función de la Umbralización (imagen derecha).

Existen diferentes funciones de umbralización que la librería de OpenCV proporciona, ver Tabla 1.

Tabla 1. Presentación de los tipos de Umbralización.

Imagen Original	Tipos
	Binario
	Binario invertido
	Truncar
	Umbral a Cero
	Umbral a Cero invertido

Ésta función se usa mucho en los diferentes ámbitos de segmentación, realizar dicha operación resulta útil, en tanto que la mayoría de los problemas buscan distinguir dos objetos diferentes en la imagen: el objeto de interés y el fondo [9]. El valor

del umbral debe ser escogido de acuerdo a un criterio particular, no sólo a las variaciones en la tinción o en la iluminación de la muestra.

Erosión. Provoca en las imágenes el afinado de las regiones de píxeles negros. Esto es de utilidad para adelgazar los trazos de un dibujo. Dicha función toma como parámetro la imagen a erosionar I y un elemento estructural B , que determina la forma de un entorno de píxeles, del cual el mínimo es elegido y colocado en el lugar del píxel procesado:

$$C(x, y) = \min\{I(x + s, y + t) | (s, t) \in B\} \quad (2)$$

En caso de no especificarse el elemento estructural (como es en éste), la función utiliza una matriz por defecto de 3×3 . El efecto de erosión se puede aplicar cierta cantidad de veces, que se especifica como parámetro de la función (ver Fig. 2).

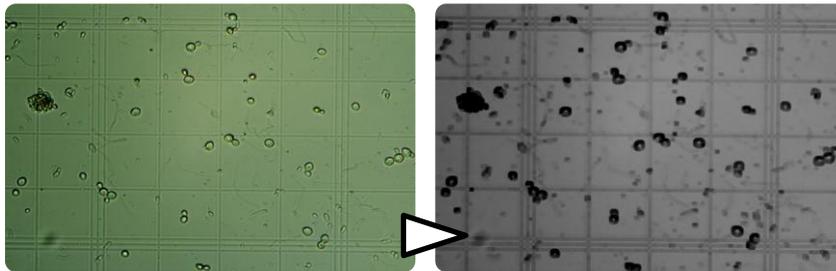


Figura 2. Muestra original (imagen izquierda); aplicación de la Función Erosión (imagen derecha).

Si bien, las herramientas que provee OpenCV son varias, se debe tener en cuenta los puntos antes mencionados para realizar la segmentación de microorganismos, por ello se llevaron a cabo una serie de pruebas con los diferentes filtros que proporciona la librería para obtener así un solo conjunto de filtros.

2.2 Batería o conjunto de filtros

Consiste de un conjunto de filtros que permite obtener la información de interés de una imagen, en este caso la batería que se presenta realiza la segmentación de diferentes imágenes microscopias, generando como resultado final una imagen binaria (blanco y negro).

Para este trabajo, se propusieron una serie de baterías, dentro de éstas se seleccionó una, que proporcionó los mejores resultados, ver Fig. 3. Cabe mencionar que el orden de los filtros da un resultado distinto.

Para conseguir este conjunto final se utilizaron varias funciones que proporciona la librería de OpenCV, la batería final con la que se trabajó cubrió las necesidades requeridas para la segmentación sin utilizar un número excesivo de funciones.

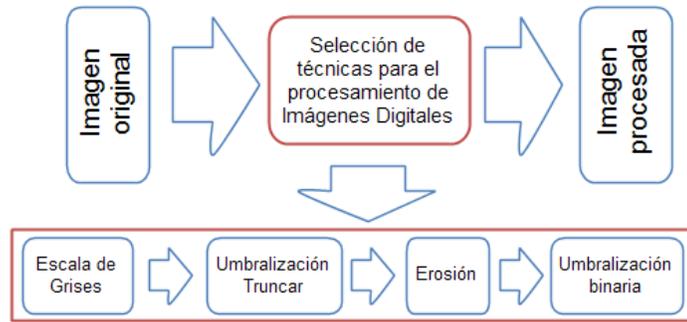


Figura 3. Selección de las funciones proporcionadas por OpenCV para el procesamiento de imágenes microscópicas.

En la Tabla 2 se muestra la aplicación de la batería a las muestra UCP3.

Tabla 2. Generación de imágenes a partir de la aplicación de las funciones, mejorando así la búsqueda de los objetos.

Imagen original	Sección de la muestra original	Descripción
		Muestra del pulque elaborada en la Región de la Cañada (UCP3)
		Imagen en escala de grises.
		Imagen con Umbralización Truncar
		Aplicación de Erosión



Cabe mencionar que las otras baterías se descartaron porque en las imágenes resultantes se obtuvo ruido y pérdida de los objetos de estudio. El resultado final que se obtiene, en este caso la imagen binaria, es necesario para un posterior uso.

3. Resultados

En esta sección se presenta la aplicación de la batería elegida con las diferentes imágenes microscópicas, cabe resaltar que los umbrales de todas las funciones para cada tipo de imagen son distintos, ver Tabla 3, debido a la variación de iluminación en la imagen, tamaño y forma de los objetos.

Tabla 3. Aplicación del conjunto de filtros a los diferentes.

Nom- bre	Imagen Original	Batería 1	Batería 2
(A) UCT1			
(B) LC2			
(C) UCP3			
(D) MS4			

La aplicación de los filtros en un orden diferente proporciona una imagen distinta, es por ello que se establece la serie de filtros agrupados en una batería. Además, la Batería 1 fue propuesta anteriormente, y a comparación de la Batería 2 que usa las funciones de OpenCV.

Los resultados obtenidos al aplicar la batería 1 en la muestra A (UCT1) reduce considerablemente el ruido en comparación de la batería 2 propuesta en [7], esto se percató en la esquina inferior izquierda de la imagen; con respecto a la muestra B (LC2) se obtuvo mejoras, entre ellas están la eliminación de ruido presente en la imagen (causado por la iluminación no uniforme del fondo de la misma) y los objetos de estudio con forma definida.

En la muestra C (UCP3) se observó diversos microorganismos de diferentes formas, por lo que al aplicar dicha batería se eliminaron los microorganismos bacterianos (cocos y bacilos), así como también los filamentos de los hongos, quedando de esta manera la información de interés (levadura).

Finalmente, en la muestra D (MS4), se percató que no es tan nítida (refiriéndose a los objetos de estudio), por lo cual al aplicar la batería 1 se pierde información, aunque con la batería 2 se obtuvieron mejores resultados.

Es importante señalar con respecto a las pruebas realizadas, aplicando la batería 1 se examinó en las muestras A, B, y C, que el objeto de estudio (levadura) está presente, sin embargo, no es el mismo resultado en la muestra D; por otro lado la batería 2 realizó la segmentación de manera correcta en las muestras A y D, excepto para la B y C.

Como se observó, los microorganismos presentan diferentes formas, relleno y vecindades [10]. Por ello es interesante e importante tratar con el PDI para la segmentación de muestras de microorganismos. Algunos de éstos están bien definidos y esto facilita la segmentación [10 - 14], en el caso de las imágenes propuestas en este trabajo no es de la misma manera, por esta razón, se hace una propuesta de ésta batería.

4 Conclusiones

El uso de las herramientas de librería de OpenCV facilita la implementación de código, mediante la cual se desarrolló la batería propuesta, además de haber efectuado diversas pruebas para obtener la batería final.

Es importante señalar, con respecto a las pruebas planteadas, la batería 1 realiza la segmentación de diferentes tipos de imágenes microscópicas y de esta forma se evita definir un nuevo conjunto para cada tipo de muestra. La Batería 2 solo lleva a cabo la correcta segmentación de la primera muestra y con las siguientes no se obtienen un buen resultado, esto podría atribuirse a la no uniformidad de iluminación.

En trabajos posteriores si se requiere contabilizar las bacterias, es necesario modificar esta batería por una que permita la segmentación de todos los microorganismos presentes en la imagen.

Dicho trabajo proporciona una utilidad en la Informática y en el área de Microbiología, como una herramienta de aplicación para el conteo celular. Debido que en la actualidad este conteo se realiza de forma manual además se pretende que la

herramienta realice la distinción de células vivas y muertas (Viabilidad celular) ya que es un parámetro muy importante dentro de la fermentación.

Además, uno de los propósitos de este trabajo es utilizar las herramientas proporcionadas con software libre, para facilitar el PDI, sin embargo, existen otras que no están implementadas y resulta necesario crear nuevas herramientas que faciliten el PDI.

A futuro se plantea que el sistema se automatice, es por ello que se realizan pruebas con diferentes imágenes para integrarlo.

Referencias

1. Chang, I.S., B. H. Kim, y P. K. Shin (1996). Use of sulfite and hydrogen peroxide to control bacterial contamination in ethanol fermentation. *Appl. Environ. Microbiol.* 63, (1), 1-6.
2. Campos J, Establecimiento de un proceso de producción de etanol a partir de jugo de caña y miel intermedia B con *Saccharomyces cerevisiae* ITV-01. Tesis de Maestría, UNIDA-ITV, Mexico (2008).
3. Strehaiano, U. (1984). Phénomènes d'inhibition et fermentation alcoolique, These Dr. És Sciences. I.N.P Toulouse, Francia.
4. J.R. Parker, *Algorithms for Image Processing and Computer Vision*, Second Edition, Wiley Publishing Inc. 2011, ISBN: 978-0-470-64385-3.
5. Open Source Software. <http://opencvlibrary.sourceforge.net/>.
6. Oswaldo Rojas Camacho. 2000, "Segmentación de Imágenes usando Algoritmos Difusos", *Revista ECM* 6(1):51-63.
7. Beatriz A. Sabino; José A. Márquez; Jesús M. Campos. 2010. Segmentación de Células de la Levadura *Saccharomyces cerevisiae*. *Temas de Ciencia y Tecnología.* 15(45):3-8.
8. Nicolás Sierra, Enrico Pizzorno, Juan Daniel Ferré. Desarrollo de una aplicación para trabajar con la técnica de "Stop Motion". Proyecto de Grado. Montevideo, Uruguay, 2004.
9. Javier Rojas. "Segmentación y reconocimiento de patrones en imágenes de histología", (2006).
10. Victor González, Enrique Alegre, Patricia Morala-Argüello, Sir Alexei Suárez Castañón. "Segmentación de cabezas de espermatozoides de Verraco mediante combinación de umbralización y transformada watershed".
11. Virginia L. Ballarin, Gustavo J. Meshino, Guillermo N. Abras y Lucia I. Passoni. "Segmentación de imágenes cerebrales de Resonancia Magnética basada en Redes Neuronales de Regresión Generalizada.
12. Humberto Bustince, Gerardo Artoln, Miguel Pagola, Edurne Barrenechea, Hamid Tizhoosh. "Sistema neuro-difuso intervalo-valorado aplicado a la segmentación de imágenes de ultrasonido". 2008.
13. Moreira Quiroz, José, Valencia Delgado, Vladimir, Chavez Burbano, Patricia "Implementación de un algoritmo para la detección y conteo de células en imágenes microscópicas".
14. Mariana Andrés Tribuj. Segmentación de imágenes texturizadas. Tesis de Licenciatura en Ciencias de la Computación, Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Computación, Buenos Aires.

Contratos traslúcidos: Implementación de Diseño por Contrato en el Desarrollo Orientado a Aspectos

Guadalupe Isaura Trujillo-Tzanahua y Ulises Juárez-Martínez

División de Estudios de Posgrado e Investigación,
Instituto Tecnológico de Orizaba
Orizaba, Ver., México
gittz87@hotmail.com, ujuarez@ito-depi.edu.mx
Paper received on 24/07/12, Accepted on 10/09/12.

Resumen. Un contrato traslúcido es el resultado de combinar 1) la Programación Orientada a Aspectos (POA) la cual ofrece ventajas como: separación de asuntos, facilidad para razonar conceptos, reutilización, mejor mantenimiento del software, entre otras; 2) implementar el Diseño por Contrato (DbC) de forma dinámica, es decir que conforme al evento que identifique una regla pueda aplicar o no. En este artículo se presentan los elementos propuestos en el Enfoque de Temas y los resultados obtenidos en el desarrollo de software para una aplicación de Domótica combinando la POA y los contratos traslúcidos en Ptolemy.

Palabras Clave: Contratos Traslúcidos, Diseño por Contrato, POA, Ptolemy.

1 Introducción

El objetivo de la Ingeniería de Software es mejorar la calidad del software mediante la reducción de costos de producción, facilitar su mantenimiento y evolución. Para lograr este objetivo han surgido diversos paradigmas de programación que buscan reducir la complejidad del software y promueven la reutilización y desarrollo de software fiable. Dentro de los paradigmas de programación se encuentra la Programación Orientada a Objetos (POO) la cuál busca el desarrollo de software de calidad a través de la abstracción, encapsulación, polimorfismo y descomposición, no obstante presenta limitaciones que dificultan el mantenimiento y evolución del software. La descomposición aplicada en la POO consiste en dividir el sistema en varios subsistemas que corresponden a diferentes partes del dominio, sin embargo no se adapta a algunos requisitos no funcionales tales como: manejo de excepciones, seguridad, sincronización, registro (*logging*), persistencia, entre otros. Con el advenimiento de la Programación Orientada a Aspectos (POA) que cada vez es más utilizada en áreas como Aviación, Domótica, Redes de telecomunicaciones, por

mencionar algunas, debido a que permite una adecuada modularización de las aplicaciones encapsulando los asuntos de corte, es decir asuntos que interfieren con la funcionalidad de otros. Tras observar que el Diseño por Contratos (DbC) es una técnica muy conocida pero poco utilizada en el desarrollo de software, se infiere la utilidad de combinar estos conceptos (DbC y POA) para aprovechar al máximo las ventajas que cada uno ofrece. La combinación del DbC y el enfoque orientado a aspectos da origen a un mecanismo más robusto denominado contrato traslúcido cuya implementación es provista por Ptolemy, un lenguaje orientado a aspectos con soporte completo para la programación basada en eventos. Sin embargo, no se reporta una guía para el desarrollo de software orientado a aspectos utilizando contratos traslúcidos.

La contribución principal de este trabajo consiste en especificar mediante un caso de estudio aplicado a la Domótica, cómo construir correctamente software en Ptolemy aplicando este concepto desde las etapas de análisis y diseño del desarrollo de software orientado a aspectos (DSOA) utilizando el Enfoque de Temas.

Este artículo está organizado de la siguiente forma: en la sección 2 se describe el lenguaje de programación orientado a aspectos llamado Ptolemy. La sección 3 se refiere a la técnica del Diseño por Contratos. La sección 4 describe a los contratos traslúcidos. La sección 5 presenta el Enfoque de Temas. La sección 6 muestra los resultados obtenidos mediante un caso estudio. La sección 7 presenta los trabajos relacionados. Finalmente, en la sección 8 se presentan las conclusiones y el trabajo a futuro.

2 Ptolemy

Ptolemy es un lenguaje de programación cuyo objetivo es mejorar la capacidad de un ingeniero para separar los asuntos conceptuales. Asimismo, añade una nueva noción de contratos flexibles denominados contratos traslúcidos [1].

Cuando se implementan algunos requisitos utilizando las técnicas orientadas a objetos por ejemplo: manejo de excepciones, sincronización, compartir recursos, *logging*, por mencionar algunos, el código de esos requisitos se dispersa en todo el software y se confunde con el código de otros requisitos. Estos requisitos se denominan **asuntos de corte** (*crosscutting concerns*). Cuando los desarrolladores atienden las solicitudes de mantenimiento a este tipo de requisitos es necesario que examinen un gran número de módulos del sistema para identificar los cambios que se requieran.

El objetivo de Ptolemy es establecer mecanismos que permitan a los desarrolladores de software implementar los asuntos de corte en módulos separados llamados *handlers*, los cuales mejoran la separación de asuntos [2].

Ptolemy extiende a Java con nuevos mecanismos para la declaración de eventos y avisos a estos eventos. Asimismo permite a los desarrolladores escribir contratos traslúcidos como parte de la definición del tipo de evento.

Ptolemy se distingue de lenguajes OO como Java debido que ofrece el soporte para el manejo de eventos y de los Lenguajes Orientados a Aspectos como AspectJ y CaesarJ porque solventa dificultades identificadas tales como:

1. **Cuantificación (quantification):** referenciar los lugares en el código donde un evento es disparado utilizando el identificador del evento. Esto se hace sin tener que nombrar a las clases que disparan el evento.
2. **Inconsciencia (obliviousness):** liberar al programador de las especificaciones puntuales donde una sentencia o conjunto de sentencias del programa tendrá efecto.
3. **Fragilidad en los cortes:** dependencia de los cortes respecto a cómo se estructura el código sobre el que actuarán los aspectos.
4. **Sombras en los puntos de unión (Join point shadows):** instrucción en el bytecode donde los avisos se pueden aplicar.
5. **Control de avisos:** los avisos deben ser comprendidos con el fin de razonar el flujo de un programa y cómo un aviso puede interferir con la ejecución de otro aviso [3].
6. **Razonamiento modular:** la formalización del Diseño por Contrato en la POA dificulta el razonamiento de los efectos de un aspecto sobre otros [4], es decir se debe considerar:
 - Todos los posibles puntos de unión en el código base.
 - Todos los aspectos aplicables en cada punto de unión.
 - Por cada aspecto, el control de flujo de los aspectos deben ser comprendidos, es decir el desarrollador es capaz de conocer todo lo que ocurre en un módulo funcionalmente independiente (alta cohesión y bajo acoplamiento).

3 Diseño por Contrato (DbC)

El Diseño por Contrato (DbC) es una técnica diseñada por Bertrand Meyer y característica central del lenguaje de programación Eiffel, que él desarrolló [5]. El DbC no es específico para Eiffel, puede ser utilizado con cualquier lenguaje de programación para desarrollar software correcto, es decir que se comporte de acuerdo a su especificación.

Un contrato es una especificación que establece las condiciones de uso e implementación (obligaciones y beneficios) que clientes y proveedores de un componente deben cumplir. Consecuentemente, un contrato protege al cliente por especificar lo que necesita y por el otro al proveedor por especificar los servicios que debe implementar.

Una regla importante para el DbC es el principio de no cláusulas ocultas [5], es decir, la precondition es el único requisito que un cliente debe cumplir para obtener un servicio, por lo tanto no hay sorpresas.

Una aseveración es una condición que debe cumplirse. Su incumplimiento invalida totalmente el software. En el DbC se utilizan tres tipos de aseveraciones:

1. **Precondición:** expresa las restricciones bajo las cuales una rutina funcionará correctamente.
2. **Postcondición:** describe el efecto de la rutina.
3. **Invariante de clase:** estado que deben cumplir los objetos de una clase, son restricciones que deben satisfacerse por cada objeto tras la ejecución de los métodos y constructores.

4 Contratos traslúcidos

Un contrato traslúcido para un tipo de evento es un algoritmo abstracto que describe el comportamiento de los aspectos que se aplican a una interface orientada a aspectos. El algoritmo es abstracto porque suprime muchos detalles de la implementación. Esto permite a la especificación decidir qué detalles ocultar y cuáles revelar. Un fragmento de código satisface la especificación de un algoritmo abstracto si el código refina la especificación, pero usa de forma limitada el refinamiento que requiere de una estructura similar, lo cual permite el control de efectos [2]. Un contrato traslúcido se puede expresar como:

DbC + POA= Contratos traslúcidos

En [6] se propone a C# para la aplicación de contratos traslúcidos debido que son independientes de su contexto original (Ptolemy) y son aplicables a otras interfaces orientadas a aspectos como Crosscut Programming Interfaces (XPIs), Open Modules (OMs) y Aspect-Aware Interfaces (AAIs).

Los compromisos que deben cumplir los contratos traslúcidos son [7]:

1. Los contratos traslúcidos deben revelar todas las llamadas a **invoke**.
2. Todos los *handlers* deben refinar el contrato de un evento.

4.1 Definición de un contrato traslúcido

La sintaxis para la definición de contratos traslúcidos se muestra a continuación:

```
01 <Evento>:= [Modificador *] [Tipo] event [Identificador] {[Contrato]}
02 <Contrato>:= requires <Expresión> assumes <Sentencia> ensures <Expresión>
```

En la figura 1 se muestra la implementación del contrato traslúcido para el caso de estudio de la aplicación de Domótica que se encarga de monitorizar cuando algún elemento de la casa (climatización, iluminación o alarma) es encendido o apagado.

```
01 public void event EventoCasa {
02     DomoticaElement eventoCasa;
03     //Contrato traslúcido para este tipo de evento
04     requires eventoCasa != null //precondición
05     assumes { //invariante de clase
06         next.invoke();
07         establishes next.eventoCasa() == old(next.eventoCasa());
08     }
09     ensures eventoCasa != null //postcondición
10 }
```

Figura 1. Implementación de un contrato traslúcido para la aplicación Domótica.

5. Enfoque de Temas

El Enfoque de Temas permite identificar y modelar aspectos independientemente del lenguaje orientado a aspectos a utilizar. Clarke y Baniassad establecen que el enfoque de temas se compone de dos partes [8]:

1. **Theme/Doc (Análisis):** conjunto de heurísticas para el análisis de los requisitos de software.
2. **Theme/UML (Diseño):** permite al desarrollador modelar en UML temas (características y aspectos) de un sistema.

Un **tema** [9] es un elemento de diseño, es decir una colección de estructuras y comportamientos que representan una característica. Varios temas pueden integrarse o combinarse para formar un sistema. El modelo de temas considera dos tipos de temas [10]:

1. **Temas Base:** los cuales comparten algunas estructuras y comportamiento con otros temas base.
2. **Temas de Corte:** aspectos, es decir temas que poseen comportamiento que corta transversalmente la funcionalidad de los temas base.

5.1 Theme/Doc

El proceso de análisis Theme/Doc tiene dos actividades principales:

1. Identificar los temas principales en un sistema
2. Determinar si las responsabilidades de cierto tema deben ser modelados como un aspecto.

Una vista es un diagrama que engloba requisitos. Theme/Doc se compone de 3 principales vistas [8]:

1. **Vistas de relaciones:** muestran las relaciones entre los temas.
2. **Vistas de corte:** los requisitos comunes entre temas describen el comportamiento que corta transversalmente a otros temas (corte). Es decir los requisitos se encuentran dispersos en varios temas (asuntos de corte) del software. Sin embargo, si estos temas de corte se identifican correctamente pero no se manejan adecuadamente se presentan las dificultades anteriormente citadas (sección 2). Para solventar estos problemas desde las fases de análisis y diseño, se consideró utilizar los contratos traslúcidos, por lo cual es necesario definir extensiones al metamodelo de UML. En la sección 6.2 se presentan las extensiones al metamodelo UML para la aplicación de contratos traslúcidos.
3. **Vista individual:** muestra los requisitos de un solo tema (sin cortes).

5.2 Theme/UML

Pasos del proceso Theme/UML:

1. Diseño de temas.
2. Especificar las relaciones entre los temas.
3. Composición de temas para la verificación de propósitos.

6 Resultados experimentales

En la figura 2 se presenta el diagrama de clases del caso de estudio de un sistema de software con aplicación Domótica, diseñado para demostrar la aplicación y utilidad de los contratos traslúcidos.

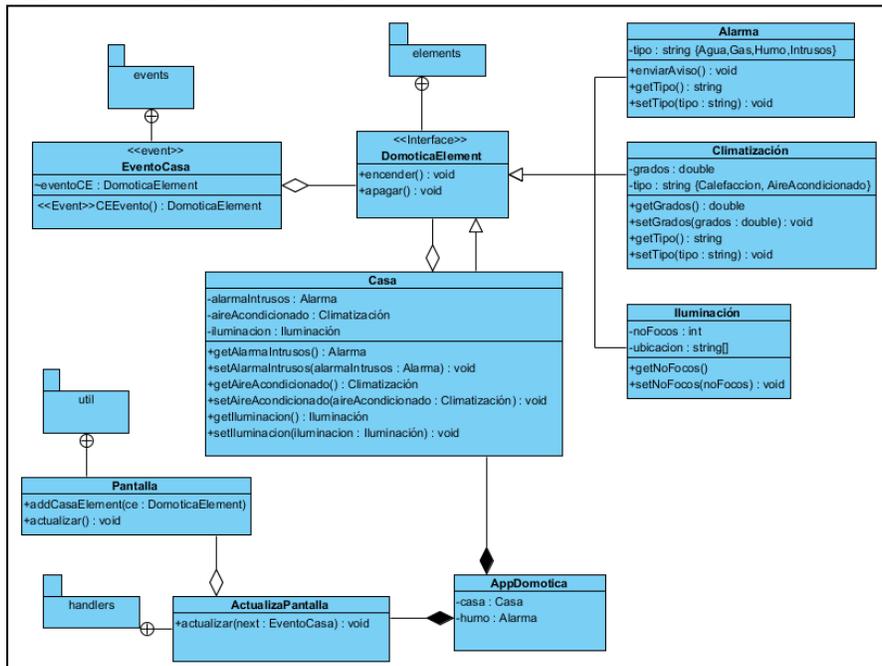


Figura 2. Diagrama de clases para el caso de estudio.

6.1 Requisitos de la aplicación de Domótica

A continuación se listan los requisitos que se implementan en la aplicación Domótica:

- R1. El sistema **registra** en pantalla cuando algún elemento de la casa (climatización, iluminación o alarma) es encendido o apagado.
- R2. El usuario **enciende** el elemento Climatización, seleccionar el tipo deseado (Calefacción o Aire Acondicionado).
- R3. El usuario **enciende** el elemento Iluminación: debe indicar el número de focos y seleccionar la ubicación.
- R4. El usuario **enciende** el elemento Alarma: es necesario indicar el tipo de alarma (agua, gas, humo o intrusos).

- R5. El usuario **enciende** el elemento Casa que integra todos los elementos deseados a activar al salir de la casa.
- R6. El usuario **apaga** el elemento Climatización, elige el tipo a desactivar.
- R7. El usuario **apaga** el elemento Iluminación: debe indicar el número de focos y seleccionar la ubicación a desactivar.
- R8. El usuario **apaga** el elemento Alarma: debe indicar el tipo de alarma a desactivar.
- R9. El usuario **apaga** el elemento Casa para desactivar todos los elementos de la configuración.

En la figura 3 se presenta el diagrama de vistas de acción para la aplicación Domótica. Una vista de acción consiste en 2 elementos: acciones (rombos) y requisitos (rectángulo redondeado). Para desarrollar las diversas vistas, se identifican en el requisito la acción clave y se conecta la acción con el requisito.

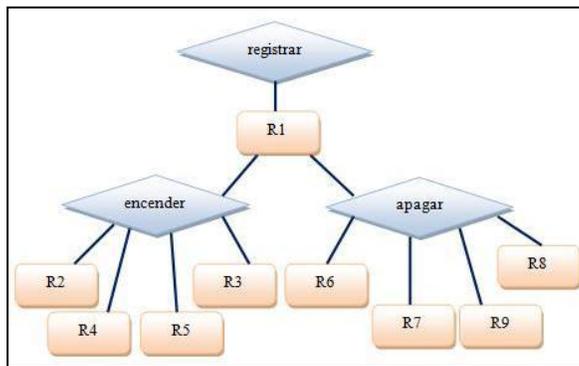


Figura 3. Diagrama de vistas de acción.

En la figura 4 se presenta el diagrama de vistas de corte para la aplicación Domótica. Se observa que entre R1, R5 y R9 existen requisitos comunes (encender y apagar), por lo que **casa** es un tema de corte, es decir el requisito se encuentra disperso en el software y se confunde (interfiere) con el código de otros requisitos.

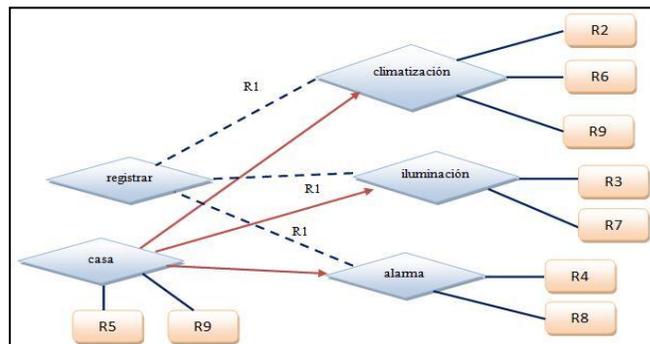


Figura 4. Diagrama de vistas de corte.

6.2 Elementos propuestos en el enfoque de temas para la aplicación de contratos traslúcidos

1. Elegir el tema dominante, es decir aquel tema base que posea una relación del tipo “**has-a**” (agregación o composición), en el caso de estudio el tema **registrar** el cual representará la definición del contrato traslúcido para el evento denominado **EventoCasa**.
2. Al identificar los temas base y de corte: casa, climatización, iluminación y alarma, se propone extender al metamodelo UML con el estereotipo aviso (<<announcement>>) para indicar que en cada clase es necesario refinar el contrato traslúcido y mantener el control de efectos.
3. Se recomienda añadir el estereotipo <<handler>> para indicar la implementación del asunto de corte. En el caso de estudio, la clase **ActualizaPantalla** representa el *handler*, el cuál se encarga de registrar el orden en que recibe los avisos.

6.3 AspectJ vs Ptolemy

En la tabla 1 se presenta la implementación de aspectos de la aplicación Domótica en los lenguajes AspectJ y Ptolemy.

Tabla 1. Cuantificación AspectJ vs Ptolemy

Cuantificación en AspectJ	Cuantificación en Ptolemy
pointcut encendido(): execution(* domotica*.encender()); (1)	when EventoCasa do update;
pointcut apagado(): execution(* domotica*.encender()); (2)	

Como se observa la identificación de puntos de unión en AspectJ se consigue mediante patrones de firmas (1 y 2) y en Ptolemy se realiza mediante la llamada al evento a identificar con las palabras reservadas **when** y **do**.

Al proceso de conocer el número de eventos a identificar se le conoce como **cuantificación**. Sin embargo AspectJ presenta el problema de la **fragilidad de cortes** debido a que cuantifica mediante expresiones (pseudo) regulares, ya que si se modifica un aspecto o el sistema cambia, el corte no garantiza que cumpla la regla original para identificar los eventos de interés y algunos pueden ser excluidos.

Lo anterior muestra que Ptolemy permite cuantificar de manera inconsciente, al no depender del patrón de firmas, garantizando que el sistema reaccione a los eventos registrados.

7 Trabajos relacionados

En [10] se analiza la complejidad de los temas relacionados con el DbC y la POA. La naturaleza dinámica de los aspectos junto con la inconsciencia limita a las

metodologías existentes del DbC para hacer frente a los programas orientados a aspectos. Un mecanismo del DbC para la POA tiene que abordar la ejecución de contratos (sobre clases y aspectos) así como las consecuencias que deben ser validados entre los contratos sobre aspectos y los contratos en clases. Todas estas ideas se integran en una herramienta llamada CONA, basada en el DbC con aspectos, la cual valida contratos y aspectos en tiempo de ejecución. CONA es una extensión del lenguaje Java y AspectJ. Sin embargo, en este trabajo se reconoce que es posible mejorar el razonamiento modular sobre los aspectos.

En [11] se describe un enfoque orientado a aspectos llamado Theme/UML que extiende a UML para ayudar a los diseñadores en la modularización de asuntos incluyendo cortes. Un tema es cualquier característica, asunto o requisito que tiene que cumplir el sistema. Las extensiones de Theme/UML a UML se centran alrededor de cómo los temas de corte se relacionan entre sí y cómo se definen las capacidades de composición. Para proporcionar esta capacidad Theme/UML define un nuevo tipo de relación llamada relación de composición que permite al diseñador identificar aquellas partes de los temas que se relacionan entre sí y por lo tanto deberían ser compuestas.

En [12] se aborda que Theme/Doc ofrece vistas de las especificaciones de los requisitos, cuyo propósito es exponer las relaciones entre comportamientos de los requisitos. Este trabajo presenta un caso de estudio que muestra que Theme/Doc es un método eficaz para identificar aspectos en las etapas iniciales del ciclo de vida del software (requisitos).

8 Conclusiones y trabajo a futuro

Para garantizar que el software a desarrollar sea fiable se requiere de mejores prácticas que gestionen de manera adecuada requisitos y asuntos. En la actualidad, existen metodologías como el Enfoque de Temas que apoyan la separación de asuntos (vistas de acción) e identificación de asuntos de corte (vistas de corte) para minimizar el acoplamiento entre ellos. Sin embargo es necesario conservar esa separación conceptual a lo largo del ciclo de vida del desarrollo de software orientado a aspectos, por lo cual es de gran importancia implementar correctamente los contratos traslúcidos desde las etapas de análisis y diseño.

La importancia de un contrato traslúcido radica en que combina las ventajas de la Programación Orientada a Aspectos y el Diseño por Contratos de forma dinámica porque el contrato es flexible y se aplica de acuerdo al evento identificado resolviendo mediante este mecanismo los problemas que presentan algunos Lenguajes Orientados a Aspectos como CaesarJ y AspectJ como la fragilidad de cortes, cuantificación, razonamiento modular y control de efectos.

Los resultados obtenidos permiten detectar las oportunidades que el enfoque de temas ofrece para contribuir a un adecuado tratamiento de requisitos. Como trabajo a futuro se considera establecer una metodología completa y robusta que integre los

conceptos de la POA y el DbC como una alternativa flexible para el desarrollo de software fiable.

Agradecimientos

Este trabajo cuenta con apoyo por parte del Consejo Nacional de Ciencia y Tecnología (CONACyT).

Referencias

1. H. Rajan, G. T. Leavens, R. Dyer, and M. Bagherzadeh, "Modularizing crosscutting concerns with ptolemy," in Proceedings of the tenth international conference on Aspect-oriented software development companion, ser. AOSD '11. New York, NY, USA: ACM, 2011.
2. I. S. U. of Science and Tecnology, "Ptolemy language," 2012, <http://ptolemy.cs.iastate.edu/docs/>.
4. M. Bagherzadeh, H. Rajan, G. T. Leavens, and S. Mooney, "Translucid contracts: expressive specification and modular verification for aspect-oriented interfaces," in Proceedings of the tenth international conference on Aspect-oriented software development, ser. AOSD '11. New York, NY, USA: ACM, 2011.
4. M. Bagherzadeh, H. Rajan, G. T. Leavens, and S. Mooney, "Translucid contracts for modular reasoning about aspect-oriented programs," in Proceedings of the ACM international conference companion on Object oriented programming systems languages and applications companion, ser. SPLASH '10. New York, NY, USA: ACM, 2010
5. B. Meyer, Object Oriented Software Construction. Prentice Hall, 2002.
6. M. Bagherzadeh, G. T. Leavens, and R. Dyer, "Applying translucid contracts for modular reasoning about aspect and object oriented events," in Proceedings of the 10th international workshop on Foundations of aspect-oriented languages, ser. FOAL '11. New York, NY, USA: ACM, 2011
7. M. Bagherzadeh, "Enabling expressive aspect oriented modular reasoning by translucid contracts", in Proceedings of the ACM international conference companion on Object oriented programming systems languages and applications companion, ser. SPLASH '10. New York, NY, USA: ACM, 2010
8. S. Clarke and E. Baniassad, Aspect Oriented Analysis and Design. Addison Wesley Professional, 2005.
9. E. Baniassad and S. Clarke, "Theme: An approach for aspect-oriented analysis and design," in Proceedings of the 26th International Conference on Software Engineering, ser. ICSE '04. IEEE Computer Society, 2004.
10. D. H. Lorenz and T. Skotiniotis, "Extending design by contract for aspect-oriented programming," 2005.
11. E. Baniassad and S. Clarke, "Finding aspects in requirements with theme/doc," 2004.
12. S. Clarke, "Aspect-oriented design with theme/uml," Junio 2004.

A Generic Model of Deliverable Artifacts through the Development of Graphical User Interfaces

Laura C. Rodríguez-Martínez¹, Ricardo Mendoza-González¹, Héctor Durán²
and Mario A. Rodríguez-Díaz¹

¹Instituto Tecnológico de Aguascalientes, Aguascalientes, Ags., México

²Universidad de Guadalajara, Zapopan, Jal., México

lrodriguez@mail.ita.mx

Paper received on 14/07/12, Accepted on 02/08/12.

Abstract. Every product requires a process model to guide its construction; to define such a model it is required the specification of its components (i.e. its phases, activities, deliverable artifacts, agents (people and tools), and roles). The construction of Graphical User Interfaces (GUI), as a product, is not an exception. In the Human-Computer Interaction (HCI) research area the development of GUIs is driven by a set of principles, guides and tools. However, the development process is presented in a general way (i.e. current approaches do not present a detailed development process), and in most of cases the development process is not explicitly presented. In recent proposals for constructing deliverable artifacts, it is suggested the use of tools like UML, sketching, prototyping among others. Nevertheless, such proposals do not specify the deliverable artifacts to be constructed neither its order of construction. This paper proposes a generalized model of deliverable artifacts of a GUI development Process (GDP). Our main aim is reaching an explicit and detailed definition of a GDP that indicates the deliverable products and their order of construction. The contributions of this work are: (1) the generalized model of deliverable artifacts (or products) of the GDP through the normalization of the deliverable products proposed in ten selected GDPs; (2) a general description of each deliverable artifact of such a generalized model. Finally, the validation of this proposal is based on the robustness of the revised theoretical base, the arguments for the generalization of the proposed model, and by comparing our approach with the International standard ISO 13407.

Key words: Human-Computer Interaction, Graphical User-Interface, GUI Development Process, GUI Design Process.

1 Introduction.

One of the main objectives to model the GUI development processes (GDPs) is to reach deliverable artifacts with quality in two senses. First, the quality of the development process for GUIs, and second, the quality of the proper interfaces developed with the specific GDP. On the other side, GUI design is considered a critical task for the successful development of software systems. Thus, we believe that

while the System Development Life Cycle (SDLC) literature scarcely addresses the task of GUI design, such type of studies are required and can enhance the required experience in modern highly interactive software systems to design GUIs. We consider that by including the “HCI experience” in the used SDLC, it can be reached a more “systematic application” of such an experience and it can be avoided the commonly intuitive GUI design realized by developers. Then, primarily it is required to know what is to be constructed and how, and as part of this knowledge it is necessary the generic and explicit definition of deliverable artifacts into the GDP to be modeled in this paper.

We review ten selected and relevant GUI design methodologies reported in the Software Engineering (SE) and Human-Computer Interaction (HCI) literature. Each one of these proposals, defines explicit or implicitly several artifacts as products of the process of design and implementation of GUIs in software systems. We do not intend to address the great open problem of HCI. However, we believe that an analysis of such a SwE and HCI literature from a methodological point of view, will contribute to advance our scientific knowledge on how to improve the usual SDLC on regarding the GUI design.

This paper continues as follow: in section II we explain the theoretical background about well-defined processes, and a previous work about normalized activities of GDP. Section III provides a cross-grained description of the deliverable artifacts of the ten selected GDPs (named from H1 to H10). Section IV describes the generalized model proposed. In section V, the paper concludes with recommendations for further work and limitations.

2 Theoretical background.

A well-defined process model defines its phases, activities, artifacts, roles, and tools, for building a software system [10]. Since 1970 several SDLCs have been proposed in the SwE area. According to [11] each new SDLC evolves from previous ones fostered by two main drivers: (i) the knowledge-gap driver and (ii) the technological-change driver. For instance, the emergence of the Unified Process (or its most known version Rational Unified Process) can be explained by the need of counting with an Object-oriented SDLC (e.g. the knowledge-gap driver) given that availability of Object-oriented technologies (e.g. the technological-change driver). A similar situation occurred for the SDLCs based on flow of data and functions (e.g. there were technologies based on functional paradigm). Following a similar logical argumentation, we claim that in the SwE area there is a gap referring to a plausible well-defined process to design and implements GUIs as part of a normal SDLC. Then, by capitilizing the experience of HCI on SDLC we claim that the study of GDPs is necessary and important too by the following consideration presented in [12]: *“We support the notion of that the Software System development cannot be separated of the development of its respective GUI. In 1987, Rosson identified a discussion point about the benefits to separate such task or keep it on the other systems’ functionalities. Its study revealed that there is not one unified and systematic GUI design decision concerning to its realization through a separate or included GUI design process from the system design process. In such a study on a sample of*

19 software developers, 8 designers recognized such separation as feasible and necessary, including the design of GUI in early development tasks, but the remainder 11 designers indicated that GUI must not be considered as a separate element of the system. Furthermore, they expressed that they were not able to imagine the difficulty of such a separation. Rosson concluded that if the system functionality is a product of an integrated design process of the system with its GUI, then this integration will provide help for supporting one user-centered design approach with the benefit of reaching a correct early user interaction in less number of design iterations. This notion of a non-separation of the GUI design task of the SLDC has been also supported by other studies. For instance, Harston & Hix reports that such a separation introduces a greater level of design difficulty rather than its integrated execution. Thus, since the GUIs expose functionality of software-systems and they are strongly linked to the total system functionality, both processes of design should not be separated. Hence, the concern of a GUI separated or included task in a SLDC is a relevant issue that deserves further study.”

We have revised ten selected GDPs, shown in Table 1, and the survey of GDPs presented in [12]. This survey presented a table with ordered and normalized activities for ten GDPs that are shown in Fig. 1. Such selected GDPs are named H1..H10.

Table 1. Ten GUI Development Processes (GDP) down study.

H1	“Designing for Usability: Key Principles and What Designers Think” Gould & Lewis, 1985 [4]
H2	“Designing for Designers: Analysis of Design Practice en the Real World” Rosson, Maass, & Kellogg, 1987 [13]
H3	“Human-Computer Interface Development: Concepts and Systems for its Management” Hartson & Hix, 1989 [6]
H4	“The Usability Engineering Life Cycle” Nielsen, 1992 [9]
H5	“Iterative Methodology and Designer Training in Human-Computer Interface Design” Bailey, 1993 [2]
H6	“Usability Engineering Turns 10” Butler, 1996 [3]
H7	“Design Methodology and Design Practice” Lowgren & Stolterman, 1999 [7]
H8	“Guía de actuación en el desarrollo de interfaces de usuario según la metodología centrada en el usuario INTEGRAM” Losada, López, & Martínez, 2004 [8]
H9	“Diseño de Sistemas Interactivos Centrados en el Usuario” Granollers, Lorés, & Cañas, 2005 [5]
H10	“A Survey on HCI in the Software Development Life Cycle: from Practitioner’s Perspective” Abd Majit, Md Noor, Wan Adnan & Mansor, 2009 [1]

We have revised this theoretical base as a means to identify the gaps on SwE related to a systematic and detailed process for developing GUIs.

3 Description of deliverable artifacts of the ten GDPs.

This section presents the deliverable artifacts for each of the GDPs assigned to the normalized activities presented in Figure 1. The artifacts are proposed in each of the GDPs in an explicit or implicit way. Such artifacts are contained in Table 2 to Table 11.

Each table shows in the first column, the activities proposed by each GDP (one table corresponds with one GDP); and the second column shows the deliverable arti-

facts. Each deliverable artifact proposed by the corresponding GDP is located on the row corresponding to the activity that produces it. For example, the deliverable artifact “Preliminary Specification of the User Interface” (located in the second row in Table 2 for GDP H1), is produced by the activity “Preliminary Specification of the User Interface”. It is important to note that different levels of detail are presented in Tables H1..H10, since the corresponding GDPs are defined at different levels of detail.

Identification of artifacts and relating them to particular activities of each GDP is not trivial since several of the artifacts are included in some proposals in an implicit way. Furthermore relating artifacts to activities is based on the description (implicit or explicit) of both, artifacts and activities in the GDP.

Macro-phase	Phase	H1 (1985) Gould	H2 (1987) Rosson	H3 (1989) Hartson	H4 (1992) Nielsen	H5 (1993) Bailey	H6 (1996) Butler	H7 (1999) Lowgren	H8 (2004) Losada	H9 (2005) Granollers	H10 (2009) Abd Majid
Definition	Requirements	Organize the Work.	Other Activities (i.e. marketing activities)	---	---	---	---	---	---	---	Project selection and planning
		Preliminary Specification of the User Interface.	---	---	---	---	---	---	---	---	Analysis of Needs
		Collect Critical Information About Users.	---	Requirements Specification	Pre-design	---	Analysis	Design Contextualization	Analysis: task analysis / functional analysis, requirements specification	Requirements Analysis	Analysis
		Develop Behavioral Goals.	---	Functional Analysis of Tasks	---	---	---	---	---	User Analysis and Task Analysis + Functional Analysis	---
Development	Design	---	Design	Formal Conceptual Design	---	Design	Design	Detailed Design	Design: conceptual design/formal design	Design	Design
	Construction	Iterative Development Phase	Implementation	Prototyping	Design	Prototyping	Construction	Implementation	Prototyping	Prototype	Implementation
Evolution	Operation		Evaluation	Evaluation	Postdesign	Evaluation	Evaluation	Evaluation	Evaluation	Evaluation	

Figure 1. Normalized activities of the ten GDPs.

Table 2. Artifacts deliver for the activities of the GDP H1 [4].

Activity	Deliverable Artifact
Organize the work	---
Preliminary Specification of the User Interface	Preliminary Specification of the User Interface
Collect Critical Information About User	Documentation of user characteristics (cognitive, conductual, anthropometric and actitudinal characteristics)
Develop Behavioral Goals	Behavioral objective definition: <ul style="list-style-type: none"> - Behavioral User description. - Task to be performed, and the circumstances in which they should be performed. - Definition of interesting metrics, such as learning time, errors, number of help requisitions. Definition of metrics criteria.
Iterative Development Phase	Presentation, review and verify. Prototyping Empirical testing.

Table 3. Artifacts deliver for the activities of the GDP H2 [13].

Activities	Deliverable Artifacts
Other activities (i.e. marketing activities)	---
Design	---
Implementation	---
Evaluation	Evaluation

Table 4. Artifacts deliver for the activities of the GDP H3 [6].

Activities	Deliverable Artifacts
Requirements specification	---
Functional analysis of tasks	Task analysis: Task list, Task hierarchy, Task description
Format conceptual design	<p>Structural description: Dialog objects (entradas, validaciones, mensajes)</p> <p>Interfaces representations:</p> <ul style="list-style-type: none"> - Details of the form - Content - Sequence (using interface representation schemes) <p>Dialog transactional model that includes constituent objects of UI (static or dynamic display objects, input objects and dialog computational objects).</p> <p>Presentation component: lexical and syntactic criteria that implies the relation between objects as transaction, interaction and actions.</p> <p>Input Dialog. Output Dialog.</p> <p>Dialog control component (dialog between human and computer)</p>
Prototyping	Application interface model.
Interface construction	Interface Construction
Evaluation	Evaluation

Table 5. Artifacts deliver for the activities of the GDP H4 [9].

Activities	Deliverable Artifacts
Pre-design	<p>Know user:</p> <ul style="list-style-type: none"> - individual user characteristics, - user' model of the task - análisis de funciones. <p>Competitive analysis.</p> <p>Definition of Usability Objectives.</p>
Design	<p>Developers' model of the task (including metaphors).</p> <p>Prototypes.</p> <p>Consistency that includes the definition of standards and the definition of the product identification.</p> <p>Definition of the guidelines of usability for heuristic evaluation:</p> <ul style="list-style-type: none"> - general guidelines - guidelines of specific category - guidelines of specific product
Post-design	Empirical testing (with prototype)

Table 6. Artifacts deliver for the activities of the GDP H5 [2].

Activities	Deliverable Artifacts
Design	---
Prototyping	Prototyping
Evaluation	Evaluation

Table 7. Artifacts deliver for the activities of the GDP H6 [3].

Activities	Deliverable Artifacts
Analysis	User's conceptual model of system functions: User's work Concepts, Data, Processes, Task Analysis, Hierarquical Decompostion, Contextual Analysis System functions and requirements of technology
Design	Computing specifications: <ol style="list-style-type: none"> 1. Functional descomposition (objects + operations + interfaces) : <ul style="list-style-type: none"> - Mapeo del "trabajo del mundo real" (real-world work) - Computational specifications assigned to objects: for each object is defined its layout, appareance, control and beha-vior. 2. Layout, appearance, control and behavior of the presentation.
Construction	Prototypes (since low fidelity even high fidelity): <ol style="list-style-type: none"> 1. User interface model 2. First prototypes with less interactivity. 3. Prototype evolvable using reusable components. Deliverable application
Evaluation	Evaluation: <ul style="list-style-type: none"> - Metrics of evaluation - Registry of objective data. - Registry of subjective data. - Assesing of suitability of design for supporting the planned task improvements, of to recommend how to improve design. - Diagnostics of problems.

Table 8. Artifacts deliver for the activities of the GDP H7 [7].

Activities	Artifacts
Design Contextualizacion	Function list, clasification of functions in principal, necessary, desirables, and innecessary)
Detailed Design	---
Implementation	---
Evaluation	---

Table 9. Artifacts deliver for the activities of the GDP H8 [8].

Activities	Artifacts
Analysis: task analysis / functional analysis, requirements specification	User Model of Task (HTA – Hierarquical Task Analysis)
Design: conceptual design / formal design	Dialog Model.
Prototyping	- Early prototype with presentation and navigation models which represented as primitive presentation. - Implementation of operative version.
Evaluation	Verification of design integrity, Logs, Problem identification.

Table 10. Artifacts deliver for the activities of the GDP H9 [5].

Activities	Deliverable Artifacts
Analysis of Needs	- Plan. - Functional goals and objectives.

Requirements Analysis	<ul style="list-style-type: none"> - Scenario analysis that represent task and procedures. - Platform Analysis. - Usability Objectives. - Interaction Style (boceto + storyboard). - Contextual Analysis.
User Analysis and Task Analysis + Functional Analysis	<ul style="list-style-type: none"> - Stakeholders analysis. - Task Analysis. - Hierarquical task analysis (HTA) - Task-Object-Rol-Agent-Event . - Functional Analysis.
Design	<ul style="list-style-type: none"> - Metaphors Analysis. - GUI Design (paper prototype + storyboard). - Navigation. - Information Architecture, that includes the layout and graphical design. - GUI design (digital mokap + storyboard navegacional).
Prototype	Software prototype with minimal functionality called horizontal prototype, and/or disposable prototypes with different technological alternatives.
Evaluation	Heuristic evaluation, trail observation, field observation, interviews, questionaries, test, focus group, thinking aloud.

Table 11. Artifacts deliver for the activities of the GDP H10 [1].

Activities	Deliverable Artifacts
Project selection and planning	Project Plan
Analysis	<ul style="list-style-type: none"> - User Requirements: Context Analysis, User Analysis, Task Analysis, Evaluation Metrics - Formative Evaluation
Design	<ul style="list-style-type: none"> - Interface specification: Metaphor Design, Dialog Design, Media Design - Presentation Design
Implementation	Coding and Formative Evaluation

4 Generalized model of deliverable artifacts (products) of the GUI development process.

Products of the ten GDPs are normalized based on the activities of each GDP, – showed above on from Table 2 to Table 11– and on the location of these activities into the generic phases showed in the second column in Table 12. The products of the ten GPD were normalized taking as a base the activities and their location within the generic phases (shown in figure 1) as well as the location of the deliverables products of each GDP within the activities of the GPD.

Table 12. Generalized model of deliverable artifacts to develop a GUI.

Macro-phase	Phase	Generalized Deliverables	Description
Definition	Requirements	Plan	This deliverable contains the project plan to be executed.

Development	Design	<p>Definition of behavioral objectives.</p> <p>Requirements Specification</p>	<p>The definition of behavioral objectives is a document that describes the user's model:</p> <ul style="list-style-type: none"> - Stakeholders' analysis and their cognitive, conductual, antropometric and actitudinal characteristics. - Task and procedure analysis of the user: task list, hierarquy of task, task description. - Context Analysis (Scenario Analysis): that shows wakness and streghness of the user to work (mails, much thaan enough time invertid, etc.) - Fucntional Analysis - Platform Analysis <p>Requirements specification is a document that defines:</p> <ol style="list-style-type: none"> 1. The conceptual and contextual model of user' system-functions: <ul style="list-style-type: none"> - data, - processes - System task analysis - Hierarquical decomposition of functions 2. Functional objectives and goals: <ul style="list-style-type: none"> - Functions list. - Functions clasification as principals, necesaries, desirables or unnecesaries. 3. Usability Objectives 4. Interactional Style (sketch + storyboard) 5. Evaluation Metrics 6. Formative Evaluation
		<p>Preliminary Design (sketched)</p>	<p>This deliberable includes:</p> <ol style="list-style-type: none"> 1. The GUI design: <ul style="list-style-type: none"> - Task Analysis = hierarquical analysis (HTA) of Task-Object-Rol-Agent-Event - Functional Analysis - Methapor's Analysis - Paper prototyping and storyboard 2. Definition of behavioral objectives of the interface: <ul style="list-style-type: none"> - User' description - Task to be performed, and the circumstances in which they should be performed - Definition of metrics of interest, such as learning time, errors, number of help requestings. Definition of criteria to be reached to each metric. 3. Presentation, review and verify 4. Structural description: <ul style="list-style-type: none"> - Dialog objetcs (inputs, validations, messages) - Interface representations (dtails of the form, content) 5. Sequence
	<p>Design</p>	<p>This deliverable includes:</p> <ol style="list-style-type: none"> 1. Developers' model of the task 2. Computing specifications: <ol style="list-style-type: none"> a) Functional descomposition (objects + operations + interfaces). This artifact maps the real-world work with computacional specifications assigned to objects. In it is defined how the objects sould to behave in the display, in a predecible, and consistent way, according to the conceptual model of the user. It artifact defines layout, appareance, control, behavior for each object. b) Presentation: 	

			<ul style="list-style-type: none"> - layout - appearance (graphical design) - control - comportamiento <p>b) Interface specification:</p> <ul style="list-style-type: none"> - Metaphor Design - Dialog Design - Media Design - Presentation Design
	Construction	First Prototype / Operative Prototype / Final System	It is referring to several system models, since a first prototype to the final system.
Evolution	Operation = Postdesign	<p>Operation Register</p> <p>Evaluation Documents</p>	<p>Operative Register is referring to registrar of operative metrics.</p> <p>Evaluation Documents is referring to register of analysis of such metrics resumed in assesment or decisions about the evaluation to evolving.</p>

5 Mapping with the international standard ISO-13407.

In this section, we map the activities-products of the process defined in the standard ISO-13407 “Human-centered design process for interactive systems” to the new proposed generic GDP.

5.1 Overview of ISO-13407

ISO-13407 is an international standard established in 1999. The standard provides guidance on human-centred design activities throughout the life cycle of computer-based interactive systems. The standard aims at “those managing design processes” and does not provide detailed coverage of methods and techniques. ISO 13407 describes user-centered design from four aspects:

Rationale. It describes the benefits of usable systems; as training, support and cost reduction, and more satisfied and productivity user.

Planning. It provides guidelines to introduce the activities of the process design into the SDLC.

Principles. It identifies four general principles that characterize user-centered design that can be followed over all design process of User Interfaces: (1) The active involvement of user and a clear understanding of user and task requirements; (2) an appropriate matching and allocation of functions between user and technology; (3) iteration of design process of the User Interfaces; and (4) Multi-disciplinary design.

Activities. It refers to the activities of the design process of User Interfaces showed in Fig. 2, and briefly described in Table 13. The model in Fig. 2 shows two circles involving the part where the usability guidelines and principles are applied to the design process of User Interfaces.

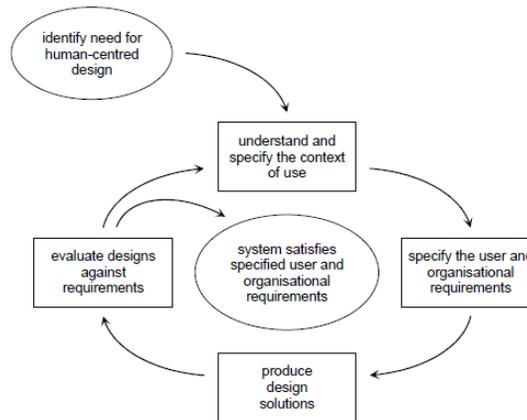


Figure 2. Activities of the standard ISO-13407 for the design process of User Interfaces.

Table 13. Description of Activities of the standard ISO-13407 for the design process of User Interfaces.

Activity	Description
Understand and Specify Context of Use.	Know the user, the environment of use, and the tasks that he or she uses the product for.
Specify the User and Organizational Requirements.	<ul style="list-style-type: none"> - Determine the success criteria of usability for the product in terms of user tasks, e.g. how quickly a typical user should be able to complete a task with the product. - Determine the design guidelines and constraints.
Produce Design Solutions.	Incorporate HCI knowledge (of visual design, interaction design, and usability) into design solutions.
Evaluate Designs against Requirements.	The usability of designs is evaluated against user tasks.

Table 14 shows the products of each activity (similar on how they are presented for H1 to H10 in Tables 2 to 11), but with the description of the artifact.

Table 14. Activities, and its product (deliverable artifacts) with their descriptions proposed in the standard ISO-13407 for the design process of User Interfaces.

Activity	Deliverable Artifact	Artifact definition
Understand and Specify Context of Use.	Listo of user and their relevant characteristics	The relevant characteristics of the user can include knowledge, skill, experience, education, training, physical attributes, habits, preferences and capabilities.
	List of Tasks the user perform; and their description.	The description should include the overall goals of the use of the system. Tasks should not be described solely in terms of the functions or features, but in terms of usability.
	Environment of Use	The environment includes the hardware, software and materials to be used. Their description can be in terms of a set of products, one or more of which can be the focus of human-

		centred specification or evaluation, or it can be in terms of a set of attributes or performance characteristics of the hardware, software and other materials. Relevant characteristics of the physical and social environment should also be described. These can include relevant standards, attributes of the wider technical environment, the physical, ambient, legislative and the social and cultural environment.
Specify the User and Organizational Requirements.	Requirements Analysis	<ul style="list-style-type: none"> - Required performance of the new system against operational and financial objectives; - Relevant statutory or legislative requirements, including safety and health; - Co-operation and communication between user and other relevant parties; - The user' jobs (including allocation of tasks, user' well-being, and motivation); - Task performance; - Work design and organization; - Management of change, including training and personnel to be involved; - Feasibility of operation and maintenance; - The human-computer interface and workstation design.
	Design Metrics	The specification of user and organizational requirements should provide measurable criteria against which the integrated design can be tested.

5.2 Mapping the activities and deliverable products of ISO-13407 with the generic GDP

We do not pretend to cover all points of view of the standard. We have carried out a mapping of activities and deliverable products from the process point of view. The mapping was carried out against the standard ISO-13407 and the mapping is only focused on its proposed activities and the products of these activities. This mapping is shown in the Table 15.

Table 15. Mapping of activities of the standard ISO-13407 with the proposed generic GDP.

Macro-Phase	Phase	Generalized Deliverables	Standard ISO Products	Standard ISO Activities
Definition	Requirements = Pre-design	Plan	Plan (implicit in the activity)	Planning (in an implicit way as an aspect defined on the standard)
		Definition of behavioral objectives.	List of user and their relevant characteristics List of Tasks the user perform; and their description. Environments of Use	Understand and Specify Context of Use.
		Requirements Specification	Requirements Analysis Design Metrics	Specify the User and Organizational Requirements.
		Preliminary Design (sketched)		

Development	Design	Design	Design (implicit for the activity)	Produce Design Solutions.
	Construction	First Prototype / Operative Prototype / Final System	Prototyping (in an implicit way for user-centered design abroad by the standard)	
Evolution	Operation = Postdesign	Operation Register Evaluation Documents	Evaluation (implicit for the activity)	Evaluate Designs against Requirements.

We can see that the standard ISO-13407 matches with the proposed GDP respect deliverable artifacts.

6 Conclusions and Future Works.

The issue about separation versus not separation of such processes is considered -in several studies- as a complex task; for example Zang [14] presents an integrated methodological proposal. In [12] several arguments are presented about the alternatives of keeping GUI activities in the SDLC against to have two separate processes. We claim that an integration of SDLC and GDP enables us to avoid duplication of tasks mainly in user analysis, task analysis and functional analysis. On the other hand, a well-defined process for developing GUIs enables us to simplify the GUI design of experts and support the GUI design of novices. We highlight the importance of integration of SDLC with GDP, for evolving toward a better process to reach successful software applications. Also, we claim this study would be useful as a base to propose a Model Driven Methodology as a future work, to develop a GDP.

This study paves the way towards an explicit GDP and a better understanding on how a well-constructed GUI can be developed. Also this study represents some progress towards a systematization of the GDP to enable the integration of it into an SDLC. This work focuses on the methodological part of the design process of GUI. Finally, as future work we are looking at: (a) a generic GDP that can be easily and explicitly integrated with either model of SDLC; (b) the integration of such generic GDP with the related generic model of deliverable artifacts; and finally (c) a complete and generic GDP that includes roles and tools to reach a well-constructed GUI.

References

1. Abd Majid, R., Md Noor, N. L., Wan Adnan, W. A., Mansor, S.: A Survey on HCI in the Software Development Life Cycle: from Practitioner's Perspective. ACM 978-1-60558-710-3/09/11. pp. 1-4. (2009)
2. Bailey, G.: Iterative Methodology and Designer Training in Human-Computer Interface Design. ACM INTERCHI'93. pp. 198-205. (1993)
3. Butler, A.: Usability Engineering Turns 10. ACM Interactions. pp. 59-75. (1996)
4. Gould, J., Lewis, C.: Designing for Usability: Key Principles and What Designers Think. Communications of the ACM Volume 28 Number 3. pp. 300-311. (1985)
5. Granollers, T., Lorés J., Cañas, J. J.: Diseño de Sistemas Interactivos Centrados en el Usuario. Editorial UOC. España. (2005)
6. Hartson, H.R., Hix, D.: Human-Computer Interface Development: Concepts and Systems for its Management. ACM Computing Surveys Volume 21 Number 1. pp. 5-92. (1989)
7. Lowgren, J., Stolterman, E.: Design Methodology and Design Practice, ACM Interactions. pp. 13-20. (1999)
8. Losada, B., López, D., Martínez, J.: Guía de actuación en el desarrollo de interfaces de usuario según la metodología centrada en el usuario INTEGRAM. Depto. de Lenguajes y Sistemas Informáticos, Facultad de Informática. Universidad del País Vasco. pp. 1-10. (2004)
9. Nielsen, J., Bellcore: The Usability Engineering Life Cycle. IEEE. pp. 11-22. (1992)
10. Oktaba, H., Ibarguengoitia, G.: Software Process Modeled with Objects: Static View. Computación y Sistemas Vol. 1 No. 4 CIC-IPN ISSN 1405-5546. pp. 228-238. (1998)
11. Rodríguez, L. C., Mora, M., Alvarez, F. J.: Process Models of SDLCs: Comparison and Evolution, Chapter on Handbook of Research on Modern Systems Analysis and Design Technologies and Applications. Minnesota State University. Mankato, USA. pp. 76-89. (2008)
12. Rodríguez-Martínez, L. C., Mora, Manuel, Muñoz, J., Mendoza-González R.: A Descriptive-Comparative Study of Activities of GUI Development Processes and their Evolution. CONISOFT 2012. Abril 25-27, 2012. pp. 17-24. (2012)
13. Rosson, M., Maass, S., Kellog, W.: Designing for Designers: Analysis of Design Practice in the Real World. ACM CHI+GI 87. pp. 137-142. (1987)
14. Zhang, P., Carey, J., Tremaine, M.: Integrating Human-Computer Interaction Development into the Systems Development Life Cycle: A Methodology. Communications of the Association for Information Systems Volume 15, Conference. June 14-16, 2005. pp. 512-543. (2005)

Desarrollo de aplicaciones en las nuevas plataformas móviles para la gestión de la seguridad del control de accesos.

Arturo Estudillo-Quilantán¹, Alberto-Israel Castellanos-Reyes¹, Erick Cacho-Rojas¹ y Ted Echeverría-Dionisio¹.

¹Instituto Tecnológico Superior de Coatzacoalcos, Coatzacoalcos, Ver., México
a.estudillo@live.com.mx

Paper received on 25/07/12, Accepted on 07/09/12.

Resumen. El presente proyecto es resultado del estudio e investigación de estas nuevas tecnologías para mejorar la seguridad de acceso a hogares y negocios, siendo su objetivo principal, incrementar la seguridad a través de la integración de herramientas computacionales, innovando en el uso de las tecnologías de la información y la comunicación mediante el empleo de aplicaciones para dispositivos móviles que operan bajo plataformas de última generación. De esta manera, implementa una aplicación smartphones que sustituye las llaves tradicionales por claves digitales que se envían mediante Bluetooth a la cerradura, para que ésta las evalúe identificándolas como permitidas o denegadas; un sistema de control encargado de configurar las llaves y aplicar restricciones de horarios, aspectos que se notifican a la cerradura electrónica usando Wi-Fi, que recibe y evalúa las llaves enviadas desde el dispositivo móvil, permitiéndole o no el acceso, y registrando el evento ocurrido para enviarlo al sistema de control.

Palabras Clave: Seguridad, Aplicaciones móviles, sistemas embebidos.

1 Introducción.

En la actualidad, uno de los temas que causa mayor interés social en nuestro país es la inseguridad. En los últimos años, los registros de incidencia de delitos, ha incrementado considerablemente en las diferentes entidades de la república, lo que a su vez ha obligado a las autoridades dirigir la mirada a esta situación, que cada vez es más alarmante. Estadísticamente en el 2010 fueron robadas 287.6 viviendas a diario a nivel nacional [1], de acuerdo con las estadísticas del Sistema Nacional de Seguridad Pública. Las cifras muestran que en las Procuradurías de Justicia Estatales y del Distrito Federal se denunciaron 94 mil 254 robos cuando las personas estaban ausentes de sus viviendas. Así el robo a casa habitación se convirtió en la tercera modalidad del ilícito con mayor incidencia en el país. Estos actos son planeados por estas personas la cual saben por dónde atacar, ya que algunos estudios han demos-

trado que, el 27% de los casos, se lleva a cabo brincando las bardas, mientras que el 24% forzar las puertas [2]. Lo cual nos indica que, una de las vulnerabilidades que más explotan los amantes de lo ajeno es la entrada principal o secundaria de nuestros hogares. Debido a lo anterior, la misma sociedad ha tenido la necesidad de buscar opciones que incrementen su seguridad personal y patrimonial, en un mercado donde existen numerosas alternativas, tanto para el hogar como en oficinas de diferentes instituciones; sin embargo suelen ser costosas y aun así son vulnerables.

El surgimiento de tecnologías emergentes y la consolidación de las ya existentes, permiten desarrollar aplicaciones que ofrezcan solución a problemáticas como la de la inseguridad. Las Tecnologías de la Información y la Comunicación (TIC) han facilitado el intercambio de información entre diferentes aplicaciones, permitiendo a los usuarios tener mayor disponibilidad de la misma. Estas tecnologías mantienen un vínculo entre ellas, lo que hace posible su interacción con sistemas heterogéneos como dispositivos electrónicos y móviles a través de redes, por lo que las TIC ofrecen una amplia gama de soluciones en automatización, domótica e inmótica.

Uno de los equipos de las TIC mayormente utilizados en la actualidad son los teléfonos móviles. Según datos arrojados por la COFETEL [3], para finales del 2011 en México habría más de 94.5 millones de usuarios de telefonía móvil, tal y como se observa en la figura 1.

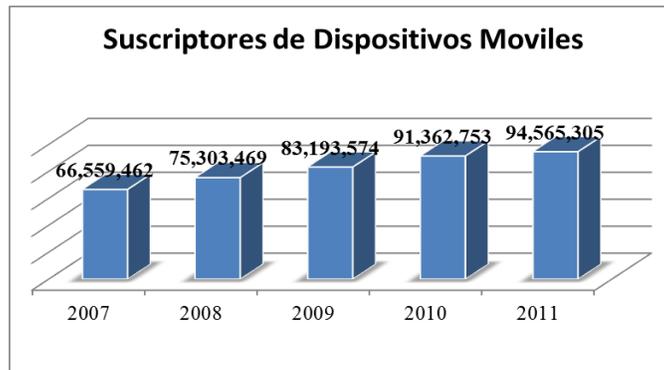


Figura 1. Usuarios de telefonía móvil según COFETEL[1].

De esta manera se aprecia que la telefonía móvil es ampliamente utilizada por más del 75% de la población, con una penetración en 2008, de 77.4 usuarios por cada 100 habitantes en todo el país [4]. Lo que ha propiciado que los desarrolladores de software hayan incursionado en el mundo de estos dispositivos, y que cada vez sea más frecuente emplearlos como portadores de sistemas complejos de gran diversidad, ya sea por comodidad, diversión o para efectos laborales.

Estos cuatro aspectos: Necesidad de seguridad en los hogares, cerraduras electrónicas robustas, administración de accesos mediante sistemas computacionales y el empleo masivo de celulares; nos permite concebir la idea del presente proyecto para solucionar el primer punto (seguridad) mediante el empleo de la tecnología existente (celulares, cerraduras electrónicas y administración computacional). De esta forma se desarrolla un sistema integral, que administre accesos mediante horarios

a partir de llaves encriptadas (claves) construidas en los celulares y controladas mediante software, que una vez enviadas a las cerraduras electrónicas verifiquen la coincidencia con aquellas que fueron dadas de alta por el sistema y suministradas únicamente por celulares igualmente registrados. En la figura 2 se puede apreciar en forma gráfica, la manera en que interactúan estos elementos del sistema en su conjunto.

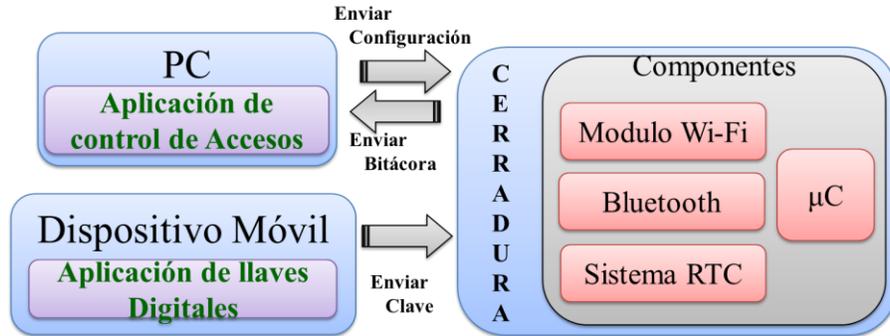


Figura 2. Diagrama de flujo de datos

De esta manera, los delincuentes no solo se enfrentarán a adivinar una clave; sino que para irrumpir sería necesario adquirir el dispositivo móvil, entrar al sistema (el cual también requerirá de una contraseña), enviar la llave (que sólo debe ser conocida por el usuario autorizado) a la cerradura, y por si fuera poco, usarla en el periodo en el que se ha permitido el acceso mediante la misma. Lo que se pretende a través de este proyecto, es reducir el riesgo de que nuestro patrimonio se vea comprometido y que todo el esfuerzo que hacemos día a día para hacernos de pertenencias que nos provean de comodidad, se vean afectadas por aquellos que gustan de arrebatar lo que no les pertenece.

2 Desarrollo.

El robusto sistema que se propone, integra la utilización de tres subsistemas que ofrecen un control apropiado de la seguridad de accesos, reduciendo la posibilidad de entrada a personas no autorizadas al incluir varios niveles de autenticación. La integración de estos subsistemas incrementa la seguridad al utilizar cerraduras electrónicas especiales, que reciben una clave digital vía Bluetooth, haciendo que no sea suficiente con ingresar un código o deslizar una tarjeta, sino que ahora se requiere la presencia de un celular previamente registrado (por la aplicación de escritorio) que ejecute una aplicación especial mediante la cual se envía una llave digital, que debe coincidir con las que el sistema registró en la memoria de la cerradura.

De esta forma existen niveles de seguridad adicionales: la presencia de celulares específicos, la utilización de aplicaciones especiales y la coincidencia con una llave encriptada para que la cerradura electrónica permita el acceso de las personas a la casa-habitación.

2.1. Aplicación de llaves digitales en dispositivos móviles.

Durante muchos años, la gran mayoría de los equipos celulares incluían una máquina virtual de Java, denominada KVM (Kilo Virtual Machine) incrustada de fábrica, la cual permitía a los teléfonos móviles ejecutar aplicaciones y juegos con un mínimo de recursos consumidos [5]. Esta pequeña máquina virtual corresponde al perfil de desarrollo MIDP (Mobile Information Device Profile, Perfil de Dispositivos de Información Móviles), por lo que cualquier aplicación para dispositivos móviles (MIDlet) desarrollada en java, puede funcionar en estos dispositivos.

Sin embargo, en los últimos años, el avance tecnológico y la demanda de equipos más sofisticados y con mayores capacidades, han permitido ofrecer un abanico más amplio de aplicaciones más cercanas al usuario para los bien llamados *smarphone* [6]. El lanzamiento de Android como nueva plataforma para el desarrollo de aplicaciones móviles, han causado gran expectación, y está teniendo una importante aceptación tanto por los usuarios como por las industrias.

Considerando la popularidad, la evolución del mercado de esta tecnología y la posible obsolescencia de las aplicaciones en java con perfil MIDlet, el durante la presente investigación se decidió migrar el desarrollo a la plataforma Android. De esta forma, se desarrolló una aplicación que nos permitiera enviar a través del puerto Bluetooth, la llave de acceso a la cerradura, al mismo tiempo que mantuviera los niveles de seguridad del usuario.

Para poder controlar el Bluetooth, el perfil MIDlet requiere el uso de la JSR (Java Specification Request) 82, un paquete que contiene el núcleo del API (Application Programming Interface, Interfaz de Programación de Aplicaciones) para Bluetooth con los métodos necesarios para esta tarea, sin embargo, esta no está disponible en todos los equipos, aun cuando estos tengan un puerto Bluetooth disponible. Por su parte, Android, a través de una Bluetooth API, permite la administración del puerto Bluetooth del *smarphone*, el cual codificará la llave de acceso y la enviará hacia la cerradura, a través de un flujo de datos serial por el protocolo RFCOMM. El proceso de conexión para el envío de esta llave cifrada involucra: a) encender y poner visible el puerto Bluetooth del equipo, b) realizar una búsqueda de aquellos dispositivos con puerto visible y con la clase de dispositivo acorde al de la cerradura, y mostrarlas en pantalla, c) seleccionar el dispositivo al cual se enviara la llave, d) solicitar la llave al usuario, y por ultimo e) establecer la conexión y enviar el flujo de datos serial; como se ejemplifica en la figura 3.

Para poder evitar la suplantación de identidad, al momento de generar la llave, no solo basta con que el usuario introduzca su palabra secreta, sino que además, el programa agrega a la cadena, datos que identifiquen el equipo de donde se está generando. Así, cuando el usuario envía su llave, esta no podrá ser reproducida desde otro teléfono al que no se le haya asignado, aun con el mismo password. Otra medida de seguridad con la que cuenta la aplicación, es un módulo de autenticación como se puede observar en la figura 4, donde el usuario deberá ingresar su nombre de usuario y contraseña para acceder, disminuyendo el riesgo de que alguien pueda generar una llave desde ese dispositivo.

damentando el código de la lógica de negocios, control y entidades en C#. En la figura 6 se puede observar el diagrama de casos de uso de este subsistema.



Figura 5. Módulo de cambio de credenciales.

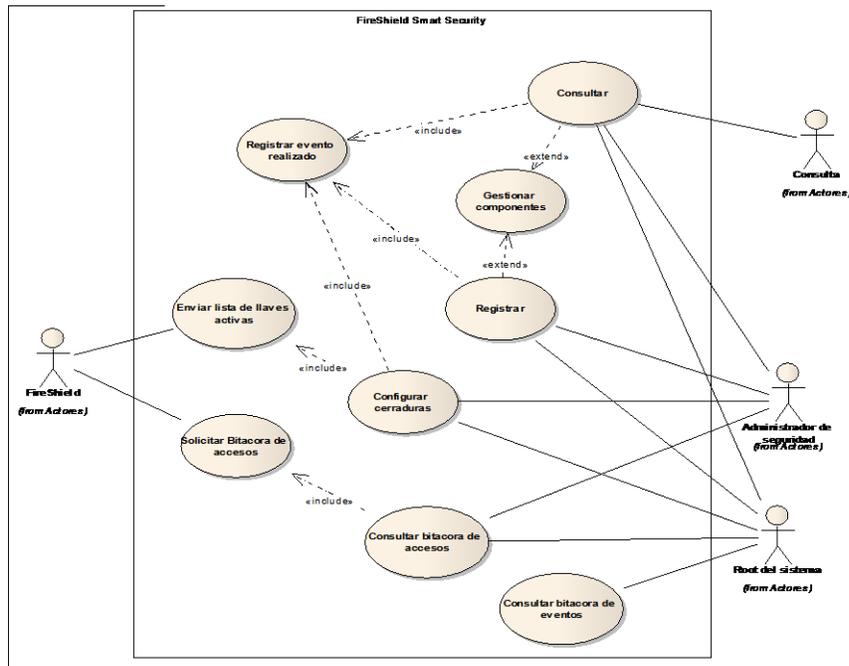


Figura 6. Diagrama de casos de uso de la Aplicación de Control.

En la figura 7 se muestra el módulo que permite capturar los datos de los usuarios del sistema. Estos usuarios dados de alta, se encargarán de registrar los equipos celulares que podrán acceder al recinto, previa autenticación a partir de atributos como su número, Bluetooth address y número de serie. Es necesario que cada vez que se agregue un celular se capturen también los datos de su dueño para consultas posteriores.

Figura 7. Módulo Registro de Usuarios

Esta aplicación tiene la capacidad de administrar diversas cerraduras, es decir que no se limita sólo a una, por lo que también cuenta con un módulo que permite agregar las cerraduras a las que podrá configurar con los horarios y equipos autorizados. Como se puede apreciar en la figura 8, de las cerraduras instaladas se registra su número de serie y DNS con el que cuenta para establecer la comunicación vía red inalámbrica.

Figura 8. Módulo Registro de Cerraduras

Una vez que existan en la base de datos los teléfonos celulares autorizados y las cerraduras existentes, se procede a capturar las llaves lógicas que serán recibidas por la cerradura desde el celular vía Bluetooth, permitiendo o denegando su entrada. Para dar de alta una llave, se debe elegir primero el celular que estará ligado a ella y la cerradura que podrá abrir, una vez hecha esta relación, se ingresa una clave alfanumérica de máximo 6 caracteres, así como la fecha inicial y de vencimiento para determinar el intervalo de tiempo en que será efectiva dicha clave. Con estos datos se construye la llave lógica que será almacenada en las cerraduras y con las cuales se verificará la posibilidad de acceso por parte de los solicitantes.

Uno de los aspectos más relevantes de la aplicación es la configuración de las cerraduras. Para ello es necesario elegir una de ellas del listado general y seleccionar la opción “configurar”, con lo que se detallará en una tabla todas las llaves que habrán de registrarse, modificarse, eliminarse o permanecer intactas en la cerradura. El estado de las mismas también es mostrado en la tabla, es decir se muestra si la llave es nueva, modificada o ya existente, en caso de que una llave previamente registrada en la cerradura no se muestre en la tabla, significaría que ha sido eliminada

y por ende se dará de baja. En la misma tabla se muestran los horarios de acceso permitidos para cada llave por día, así como la fecha de inicio y expiración. Una vez que el usuario elija la opción enviar, se inicia el proceso de conexión mediante la red inalámbrica con la cerradura a través de la búsqueda del DNS de la misma y su autenticación a través de un diálogo de presentación en el que se utiliza el mecanismo de reto-respuesta. Una vez efectuada de manera adecuada la autenticación de ambas partes, la aplicación envía una secuencia de instrucciones a ser efectuadas por la cerradura. Como parte importante de la aplicación, es posible también pedir a la cerradura, la bitácora de solicitudes de acceso concedidos y denegados almacenados en su memoria. De esta forma, es necesario indicar primero a la aplicación el nombre DNS con el que se identifica a la cerradura a nivel de usuario, para que una vez elegida la opción “actualizar”, se establezca la conexión con la cerradura seleccionada, y después de su respectiva autenticación, se le notifique mediante el comando apropiado, la requisición de la bitácora, la cual se recogerá en la aplicación y se mostrará en una tabla para ser analizada por el usuario. Esta bitácora muestra la fecha de la solicitud de acceso, la hora en que ésta se realizó, el nombre de la persona que se tiene registrado como dueño del equipo celular y si se le concedió o no el acceso. De manera continua, la aplicación registra todas las actividades realizadas por todos los usuarios dentro de la misma, para efectos de auditoría o revisión posterior exclusivamente por parte del root. Por último, como cualquier aplicación que utiliza información delicada, se cuenta, como se muestra en la figura 9, con la posibilidad de efectuar respaldos y restauraciones para maximizar la funcionalidad del sistema.



Figura 9. Módulo de Configuración y Respaldos

2.3 Cerradura electrónica.

Se desarrolló una interfaz receptora de la llave enviada por el celular, con la capacidad de almacenar las diferentes combinaciones válidas para cada uno de los usuarios, así como el registro de los eventos ocurridos. Los aspectos que se consideraron son que la cerradura debía recibir vía Wi-Fi la configuración desde el sistema central en el interior del domicilio, almacenar esa información y compararla con los datos que se reciban por el puerto Bluetooth. Utilizando un microcontrolador AT89C51 y multiplexando su puerto serial a los módulos Bluetooth MTS2BTSMI y Wi-Fi MT800SWM de Multitech, se recibe el flujo de datos provenientes del teléfono móvil o de la computadora, interpretando mediante el uso de interrupciones la

fuente de los datos. En la figura 10 se observan los módulos Bluetooth y Wi-Fi en la placa, junto al microcontrolador AT89C51.

Cuando la interrupción generada se produce por parte del módulo Bluetooth, la cerradura está recibiendo una llave de acceso, se almacena la cadena de caracteres de forma temporal y se compara con los registros existentes en la memoria. Si la llave es localizada, ahora deben evaluarse las restricciones de acceso y la caducidad de la misma. Para esto, la cerradura cuenta con un RTC (Real Time Clock, Reloj de Tiempo Real) DS1302 y un microcontrolador AT89C2051 que está programado de tal manera que sincroniza las señales que el reloj necesita para guardar y solicitar la fecha y hora para validar el acceso.

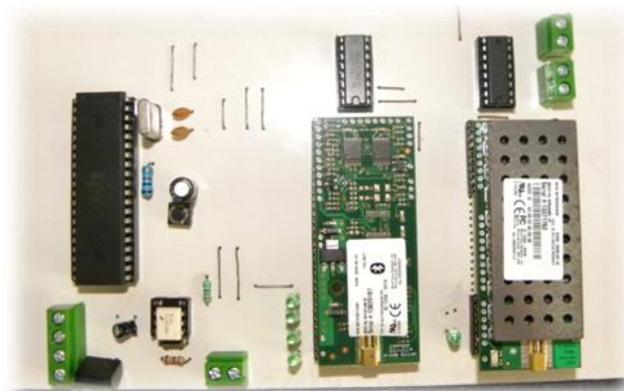


Figura 10. Módulos Bluetooth y Wi-Fi conectados al Microcontrolador AT89C51

De esta forma, el primer microcontrolador (identificado como maestro) solicita al segundo microcontrolador (identificado como esclavo) la fecha y la hora, respondiéndole este último, una cadena de caracteres que lleva la fecha y hora actual. Siguiendo un elaborado algoritmo, se efectúa la validación de las restricciones, verificando que se encuentre dentro del rango establecido por el sistema de control para ese día específico. Al terminar, el microcontrolador maestro puede resolver si le otorga el acceso o no, al usuario que lo ha solicitado, almacenando en el registro de eventos la llave recibida, la hora, la fecha y su resolución (falso o verdadero).

La elaboración del circuito de la etapa de potencia es fundamental para el funcionamiento de la cerradura al momento de que la clave sea validada. Para esto se cuenta con un MOC3010, un Triac MAC223, resistencias de 330 Ohms, 180 Ohms y 1.2 KOhms, capacitores cerámicos de 0.1 uF, un transformador de 12 VCA y por su puesto una contra eléctrica. Al pin número 1 del MOC3010 le corresponde al ánodo y es el que va a ir conectado directamente uno de los puertos del microcontrolador maestro de donde va a salir el bit en alto al momento en que se valide la clave, a su vez con el pin número 2, que le corresponde al cátodo del MOC3010 conectado a tierra, se provocará que los pines numero 4 y 6 se pongan en un nivel de alto activando la terminal 2 y el GATE del Triac MAC223 ocasionando un disparo que activa la contra. La figura 11 ejemplifica esta conexión. En caso de que el acceso sea denegado, el microcontrolador maestro encenderá un led, indicando que la llave no es válida.

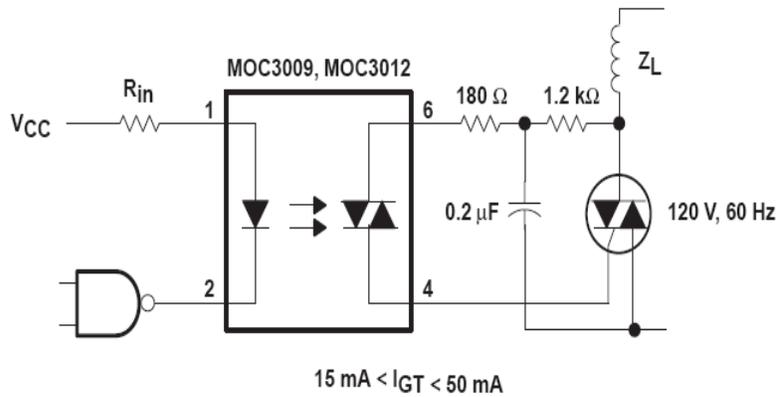


Figura 11. Sistema de la etapa de potencia.

3 Conclusiones.

El Sistema de Control de Acceso propuesto, es una respuesta a la demanda social existente en carácter de seguridad. La existencia de un delito genera una importante pérdida económica que alcanza a víctimas directas e indirectas, por lo que el objeto de la realización de este proyecto fue poder ofrecer una alternativa que permita ayudar a reducir estas mermas. Con este conjunto de aplicaciones propuesto, se busca dar solución a las necesidades básicas de seguridad que se detectan en hogares y negocios, teniendo entonces la encomienda de realizar un software adaptable al medio en que se desenvuelva.

El proyecto se encuentra en una etapa temprana de desarrollo, aun cuando los componentes de software están prácticamente terminados, la parte de la cerradura aun no ha sido finalmente desarrollada. Si bien los aspectos básicos de la comunicación entre las aplicaciones ya han sido resueltos, aun quedan pendientes algunos detalles como el mecanismo secundario de acceso, en caso de que la carga del dispositivo móvil o el mismo celular fallen; así como una mejora en el sistema de energía de reserva de la cerradura, en caso de que el suministro de energía falle, ya que la que actualmente cuenta, solo brinda 10 horas de respaldo en modo de espera.

Este proyecto innova en el uso de las tecnologías de la información y la comunicación en campos poco usuales, como la resolución de problemas sociales y específicamente la inseguridad, al mismo tiempo que combina diferentes áreas de las TIC, como redes inalámbricas y los sistemas embebidos. El aprovechamiento de todas las herramientas que ofrece el campo de la informática y la versatilidad de las comunicaciones a través de radiofrecuencias, en conjunto con el soporte de la electrónica, permite dar origen a una de las soluciones de software más creativas y revolucionarias en la región, al mismo tiempo que se explota la creciente dependencia de la sociedad en el uso de tecnologías de nueva generación, como lo es la vanguardia de los dispositivos móviles inteligentes, y la relación que estas guardan con sus actividades diarias.

Referencias.

1. L. Brito (2012,Junio). Robo a casa-habitacion repunto. *Criterio*. Disponible: <http://www.criteriohidalgo.com/notas.asp?id=15997>
2. V. Alvarez (2011,Junio). Los robos mas frecuentes. *Periodico correo*.
3. COFETEL. *Suscripciones a teléfonos celulares móviles. Serie Anual a partir de 1990*. México: COFETEL, 2012.
4. COFETEL. *Telefonia movil penetracion 1990-2010*. México: COFETEL, 2008.
5. S. Galvez y L. Ortega. *Java a tope: J2ME (Java 2 Micro Edition)*. España : S. Gálvez, 2003
6. J.Tomas Girones. *El gran libro de Android*. Mexico: Alfaomega, 2011.

Desarrollo y prueba de un Software de Apoyo para la Mejora de la Pronunciación en Alumnos con Déficit del Habla por Dislalia

José Manuel Soto-Corzo¹, José Juan Hernández-Mora², José Federico Ramírez-Cruz³, Blanca Estela Pedroza-Méndez⁴

^{1,2,3,4}Instituto Tecnológico de Apizaco,
Apizaco, Tlaxcala, México

¹jmsc_1984@live.com.mx, ²jjhmora@itapizaco.edu.mx,
³framirez@itapizaco.edu.mx, ⁴thelismedina@hotmail.com

Paper received on 24/07/12, Accepted on 12/09/12.

Resumen. Vivimos en una etapa de evolución informática orientada en gran parte hacia una evolución constante en el desarrollo de interfaces y métodos de interacción humano computadora. Dentro de esta evolución se observa una clara mejoría en la forma como las aplicaciones se adaptan a las necesidades de personas que presentan alguna deficiencia en sus capacidades. Desde nuestra perspectiva esta misma evolución debe converger no solo en la interacción, sino en el propósito incluyente del software mismo. De esta forma se planteó la necesidad de desarrollar software enfocado a la rehabilitación de personas déficit del habla por Dislalia como una herramienta de apoyo en su proceso de rehabilitación que les permita mejorar su capacidad de pronunciación por medio de actividades.

Palabras Clave: Dislalia, Rehabilitación, Reconocimiento Automático del habla, Ingeniería de Software.

1 Introducción

El término “trastornos del aprendizaje” se aplica a los problemas que plantean obstáculos al rendimiento académico o escolar. Estos problemas afectan a uno de cada diez niños, y se pueden detectar desde los cinco años de edad. Dependiendo de qué tan graves sean estos trastornos pueden afectar el aprendizaje del niño, y también sus relaciones con otras personas. Las dificultades escolares se pueden clasificar por su etiología en dificultades de origen primario, dificultades específicas y dificultades derivadas del entorno socio-familiar y cultural [1]. Dentro de las dificultades específicas, que se refieren a los trastornos del aprendizaje, se encuentra a la Dislalia, la cual se define como la dificultad para pronunciar un fonema determinado. Muchos de los niños que presentan este déficit del habla son

diagnosticados de manera tardía y su desarrollo cognitivo se ve limitado dependiendo del nivel del trastorno que presentan y del tiempo que tarde su rehabilitación. Algunos niños con dislalia pertenecen a sistemas educativos de atención especial debido a la presencia de este trastorno de aprendizaje. Existen productos comerciales enfocados a la rehabilitación de personas con déficit del habla, pero muchos no están habilitados para el idioma español y aquellos que si lo están, representan una inversión que muy pocas instituciones públicas de educación especial pueden adquirir. El reto consiste en proveer de aplicaciones de software libre diseñados específicamente para apoyar los procesos de enseñanza y/o rehabilitación de personas con déficit del habla. El sistema propuesto esta desarrollado para ser empleado en escuelas de educación especial de nivel primaria en el estado de Tlaxcala y cuenta con un repositorio de actividades a manera de juegos para que los alumnos puedan mejorar su capacidad de pronunciación. El sistema cuenta con un modulo de reconocimiento fonético que permite determinar si el alumno ha pronunciado correctamente los fonemas, silabas o palabras solicitadas a través del desarrollo de las actividades contenidas en el mismo sistema. Asi también cuenta con una base de datos embebida en el sistema y un visualizador de graficas de progreso.

2 Entornos de la investigación

En esta sección se presenta de manera breve una descripción de los aspectos relacionados con la investigación requerida para el desarrollo del software.

2.1 Dislalia

La dislalia se define como la dificultad para pronunciar un fonema determinado. Las posibles causas en este trastorno del habla son:

- a) **Retraso fonológico:** Relacionados a un retraso en el desarrollo del habla. El niño simplifica los sonidos pues no aprendió a producir los más complejos de su lengua. Su capacidad articulatoria no está afectada.
- b) **Trastorno fonético:** Cuando el niño no adquirió correctamente los patrones de movimiento necesarios para la producción de algunos sonidos del habla, por lo que no mueve los músculos que se encargan del habla, como debería, y por eso comete omisiones, sustituciones y distorsiones de algunos sonidos de la lengua.
- c) **Alteraciones físicas:** presencia de malformaciones físicas en los órganos que intervienen en el habla y que impiden pronunciar muchos sonidos. Las lesiones en el sistema nervioso pueden ocasionar alteraciones en el movimiento y en la coordinación de los músculos implicados en el habla. (ejemplo, labio leporino o hendido), pueden ocasionar una dislalia.

Un programa de rehabilitación para la dislalia consiste en los siguientes pasos:

1. Estimulación de la capacidad del niño para producir sonidos.

2. Estimulación de la coordinación de los movimientos necesarios para la pronunciación de sonidos, ejercicios labiales y linguales.
3. Realización de ejercicios para producir sonidos dentro de sílabas hasta automatizar el patrón muscular necesario para articular el sonido.
4. Empezar con palabras completas a través de juegos.
5. Una vez que el niño es capaz de pronunciar los sonidos difíciles, tratar que lo realice en conversaciones espontáneas (fuera de terapia).

2.2 Fonología y Fonética

La fonología es un sub campo de la lingüística. Mientras que la fonética estudia la naturaleza acústica y fisiológica de los sonidos, la fonología describe la manera en que los sonidos funcionan (en un lenguaje o en una lengua en general) en un nivel abstracto o mental [2]. Un sonido o fono se caracteriza por una serie de rasgos fonéticos y articulatorios, el número de dichos rasgos y la identificación de los mismos es tarea de la fonética. Un fono es cualquiera de las posibles realizaciones acústicas de un fonema. Los fonemas son unidades teóricas, postuladas para estudiar el nivel fonético-fonológico de una lengua humana. Entre los criterios para decidir, qué constituye o no un fonema se requiere que exista una función distintiva. Un fonema es una unidad fonológica diferenciadora, indivisible y abstracta.

2.3 Reconocimiento Automático del Habla

Dado que el habla se transmite a través del sonido y que se pueden extraer patrones de una señal de sonido, el reconocimiento automático del habla recae en el área del reconocimiento de patrones. Para nosotros los humanos es una tarea intuitiva pues forma parte del núcleo de nuestra inteligencia, sin embargo en el área de la computación representa uno de los problemas con más retos [3]. La forma de extraer estos patrones es a través de métodos estadísticos y probabilísticos. Dependiendo del propósito para el cual se requiera del reconocimiento automático del habla se han de emplear diferentes técnicas de extracción y clasificación de datos. Dado que la explicación detallada de estas técnicas está más allá del propósito de este documento se sugiere consultar la bibliografía citada para más información [3], [4], [5].

3. Trabajos y herramientas similares

En esta sección se hace mención de algunos trabajos y herramientas similares para poder proporcionar un panorama muy general del estado del arte en cuanto a software de apoyo para personas con déficit del habla se refiere. En cuanto a sistemas comerciales los más representativos son los siguientes.

Speech Viewer.- Se trata del software comercial más antiguo diseñado para tal propósito. Fue desarrollado por la IBM. El programa utiliza animaciones a partir de entradas de voz, la versión actual es el Speech Viewer III. Está dirigido a problemas de lenguaje, audición y habla. Contiene varios ejercicios que tienen una respuesta

inmediata por medio de la voz y es utilizado en personas de todas las edades que cuenten con una discapacidad que afecte el desarrollo de su lenguaje [6].

Sistema Visha.- Su diseño gráfico es poco atractivo, sin embargo el sistema Visha consta de un conjunto de componentes de hardware (memoria, tarjeta de sonido, etc.) que deben instalarse en el ordenador, y un conjunto de programas de software que permiten el análisis de los parámetros del habla, la síntesis y codificación visual de la señal acústica y el reconocimiento de los sonidos [7].

Sistema AVEL.- El sistema AVEL incluye una tarjeta analógica para la computadora. La señal es captada por un micrófono unido a la computadora y almacenada de esta manera. En este sistema se encuentran aplicaciones de desmutización, articulación vocálica correcta, ejercicios de ritmo, control de tono e intensidad [8].

Como trabajo de investigación y diseño de prototipo uno de los más citados para el idioma español es desarrollado por Liseth García Robles y Sergio Suárez Guerra [9]. En su propuesta se observa la utilización de diversas técnicas para la extracción de características de la señal acústica, como son, el análisis, oscilográfico, el análisis espectral y la codificación por predicción lineal. El sistema realiza dos funciones: la primera consiste en permitir el seguimiento y rehabilitación de pacientes con trastornos del habla por medio de la visualización de su señal de voz, previamente analizada con las técnicas antes descritas. La segunda, consiste en permitir grabar la voz de personas y por medio de sus análisis diagnosticar si el paciente tiene algún trastorno. El trabajo está orientado al reconocimiento del habla continua y no de fonemas o palabras aisladas. De estas herramientas de apoyo y software específico se observa que aquellos que son comerciales son una buena opción para terapeutas enfocados a la rehabilitación de personas con problemas del habla en general, sin embargo por cuestiones de precio o distribución no son accesibles para la mayoría de las instituciones de educación especial del estado de Tlaxcala. Los proyectos de investigación y prototipos no comerciales analizados, son buenos trabajos y aportan grandes ideas al desarrollo del sistema, pero su aplicación es muy general. La intención del sistema mostrado en este trabajo, se debe a la necesidad de un desarrollo orientado a aspectos específicos como la edad de los usuarios finales, su nivel de atención, etc.

4. Descripción del Sistema

El trabajo consiste en el desarrollo de una aplicación de computadora que sirva como herramienta de apoyo en los procesos de rehabilitación de alumnos de educación especial con déficit del habla por dislalia originada por retraso fonológico y/o trastorno fonético. El sistema está diseñado para ser administrado por un profesor que pueda llevar el control de los registros almacenados de los alumnos dados de alta en el sistema, así como observar y extraer información estadística del progreso de los alumnos. Los alumnos registrados en el sistema podrán llevar a cabo sesiones con actividades a manera de juegos que les permitan mejorar su capacidad de pronunciación de fonemas específicos. Cada alumno tendrá asignado un nivel de acuerdo a su capacidad de pronunciación. Esta escala de tres niveles es como se describe a continuación: el nivel A para alumnos con dificultad para pronunciar

fonemas de manera aislada; el nivel B para aquellos que no pueden pronunciar correctamente los fonemas dentro de sílabas; el nivel C es para quienes presentan dificultad para pronunciar los fonemas dentro de palabras completas. Dichos niveles plantean una escala de complejidad, por lo que el nivel deberá ser asignado al alumno por el profesor encargado de apoyarlo en su proceso de rehabilitación.

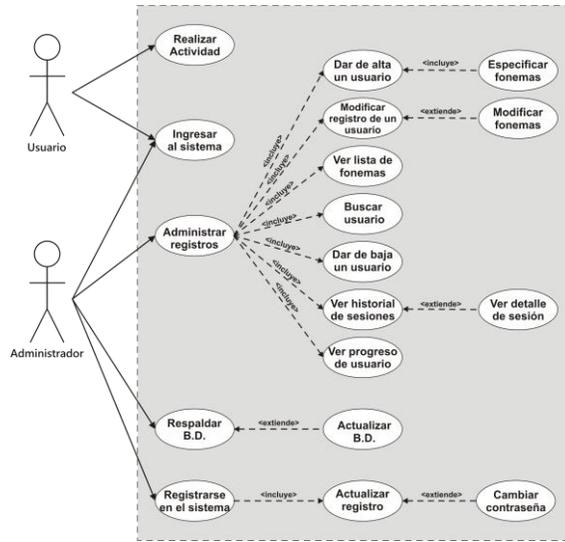


Figura 1. Diagrama de casos de uso del sistema.

En la Fig. 1 se pueden observar el diagrama de casos de uso. La mayoría de las funciones están contempladas para el administrador del sistema, sin embargo la más importante, corresponde a la realización de actividades por parte de los alumnos. La arquitectura general del sistema denominado *Speech View* se muestra en la Fig. 2.

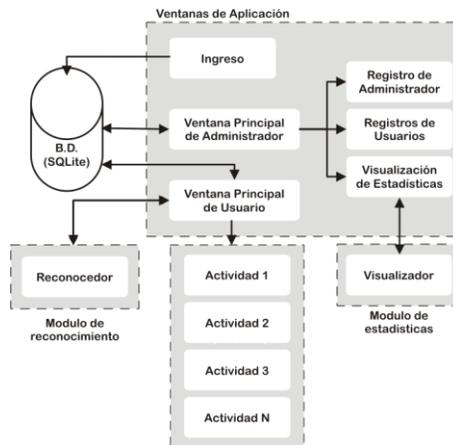


Figura 2. Arquitectura general del sistema.

Finalmente se presenta el diagrama de componentes del sistema en la Fig. 3.

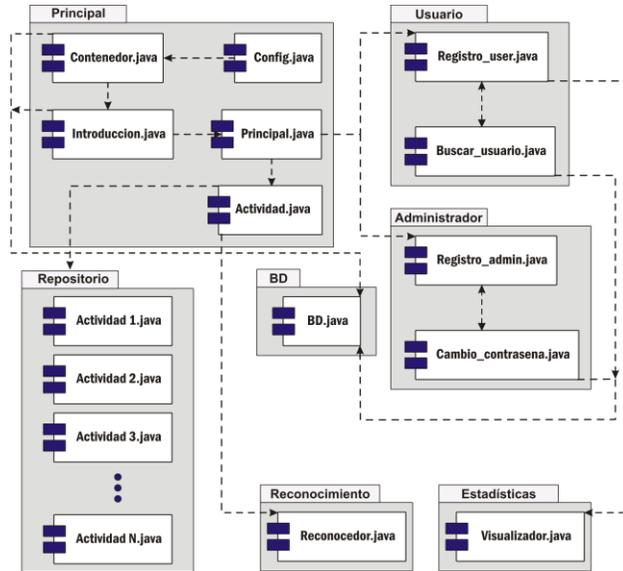


Figura 3. Diagrama de componentes del sistema.

A continuación se muestran imágenes de algunas pantallas del sistema.



Figura 4. a) Pantalla inicial; b) Ventana principal del sistema; c) Ventana de registro de administrador; d) Ventana para el registro de alumnos

Desde las pantallas mostradas en la Fig. 4 se puede ingresar al sistema (a), se puede llevar a cabo los escenarios de uso del administrador (b), se puede llevar el control de registro del administrador (c) y de los alumnos (d) respectivamente.

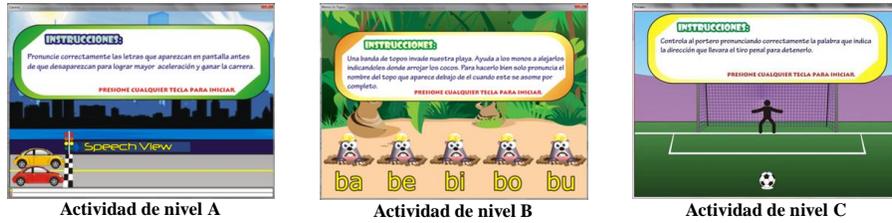


Figura 5. Ventanas iniciales de las actividades de pronunciación.

Las imágenes mostradas en la Fig. 5 corresponden a las actividades de pronunciación diseñadas e implementadas en el sistema.

5. Desarrollo del sistema

El sistema fue desarrollado empleando una metodología basada en prototipos, por medio de un patrón de desarrollo conocido como MVC (Modelo- Vista – Controlador) y empleando las siguientes herramientas de desarrollo:

Tabla 1. Herramientas de desarrollo y requerimientos asociados.

Herramienta	Requerimiento de desarrollo
Java SE Development Kit 1.7	Aplicación multiplataforma de software libre
Java Speech API 1.0	Reconocimiento automático del habla
Paquete Swing de Java	Desarrollo agíl de la interfaz de usuario
SDK de JavaFX 2.0	Elaboración de actividades de pronunciación animadas
SQLite 3.7	Resguardo de datos de los usuarios
JFreeChart 1.0.13	Generación de gráficas de progreso de los usuarios

A continuación se describe brevemente cada herramienta de desarrollo.

Java SE Development Kit.- La edición estándar del kit de desarrollo de Java en su versión 1.7 es un entorno de desarrollo para construir aplicaciones, applets y componentes usando el lenguaje de programación Java [10].

Java Speech API 1.0.- Java Speech Recognition define un estándar, fácil de usar, así como una plataforma que combina el software para el desarrollo de interfaces con el estado del arte de la tecnología en reconocimiento del habla [11].

Paquete Swing de Java.- El paquete Swing es parte de la JFC (*Java Foundation Classes*) en la plataforma Java. La JFC provee facilidades para ayudar a la gente a construir GUIs [12].

SDK de JavaFX 2.0.- El SDK de JavaFX permite crear y desplegar RIA's (*Rich Internet Applications*) que se comporten de forma consistente en múltiples

plataformas. Construida sobre la tecnología Java, la plataforma JavaFX proporciona un rico conjunto de gráficos y API's multimedia para hardware de alto rendimiento, aceleración de gráficos y los motores multimedia [13].

SQLite 3.7.- SQLite es una biblioteca que implementa un sistema autónomo, sin servidor, sin configuración, así como un motor de base de datos SQL. SQLite lee y escribe directamente en archivos de disco normal, por tanto, una base de datos SQL completa con varias tablas, índices, triggers y vistas, estará contenida en un archivo de disco único que puede ser autocontenido por la aplicación [14].

JFreeChart 1.0.13.- JFreeChart es una biblioteca gráfica de Java 100% libre, que hace que sea fácil para los desarrolladores poder mostrar gráficos de calidad profesional en sus aplicaciones [15].

6. Pruebas y resultados

Además de las pruebas unitarias para cada módulo de código durante la fase de desarrollo, fueron realizadas pruebas para evaluar el desempeño del módulo de reconocimiento del habla y el desempeño del sistema como herramienta de apoyo a procesos de rehabilitación de personas con déficit del habla por dislalia.

6.1 Pruebas al módulo de Reconocimiento Automático del Habla

Se llevaron a cabo 29 sesiones de pronunciación con tres niños en un rango de seis a ocho años de edad, todos ellos sin problemas de pronunciación, puesto que lo que se buscó fue validar el porcentaje de reconocimiento correcto del módulo. Cada uno de ellos realizó actividades de un nivel distinto. Fueron 6 sesiones de 5 actividades para el nivel A, 20 sesiones de 3 actividades para el nivel B y 3 sesiones de 3 actividades para el nivel C. A continuación se presentan los resultados y observaciones obtenidos. Los datos en la Tabla 2 muestran la tasa de reconocimiento general obtenido de las pruebas realizadas para cada tipo de elemento. Dentro de esta, las vocales y consonantes forman parte de las pruebas realizadas por medio de la actividad de nivel A, las sílabas por medio de la actividad de nivel B y finalmente las palabras corresponden a pruebas por medio de la actividad de nivel C.

Tabla 2. Tasas de reconocimiento generales por tipo de elemento

Tipo de elemento	Tasa de reconocimiento
Vocales	90.76 %
Consonantes	96.20 %
Sílabas	88.15 %
Palabras	92.33 %

El porcentaje de reconocimiento general del sistema, con base en los datos de la Tabla 2, es de 91.86 %. Se observa que los porcentajes de reconocimiento varían de acuerdo a diversos factores como lo son el ruido ambiental, la calidad del dispositivo de entrada de audio empleado, etc.

Otra observación importante, es que debido a que el motor de reconocimiento discrimina la salida que genera con base en los elementos de la gramática especificada, tiende a presentar errores de reconocimiento cuando dicha gramática contiene elementos cuyo sonido es demasiado similar. El sonido que representa un fonema esta clasificado con base en los factores que intervienen en su formación [16], por tanto entre mayor sea la similitud en la forma como son generados, mayor será la dificultad para diferenciarlos por parte del módulo de reconocimiento. Un ejemplo de este hecho, se observa rápidamente en las pruebas realizadas con las vocales, donde el sonido de la /e/ y la /i/ tienden a generar un error de reconocimiento por parte del sistema.

6.2 Pruebas al sistema como herramienta de apoyo

Para las pruebas de evaluación al sistema como herramienta de apoyo al proceso de rehabilitación de personas con dislalia, se llevaron a cabo sesiones con cada uno de los tipos de actividades que presenta el sistema. Se realizaron 9 sesiones de pronunciación con tres niños en un rango de edad entre cuatro a seis años, específicamente una niña de 4 años, y dos niños de 5 y 6 años respectivamente. Cada uno de ellos llevo a cabo actividades de un nivel distinto. Al igual que en las otras pruebas, un niño efectuó actividades de nivel A, otro niño llevo a cabo actividades de nivel B y el otro niño ejecutó actividades de nivel C. Cada uno de los niños realizo 3 sesiones con 5 actividades para cada sesión.

Es apropiado mencionar que las actividades de nivel A y B son prácticamente de refuerzo para poder llevar a cabo las actividades de nivel C, puesto que es en el nivel C donde realmente se realizan las actividades que permiten practicar la correcta práctica de los fonemas o silabas que presentan dificultad al usuario dentro de palabras muy específicas. Por ello se buscó que los niños empleados en estas pruebas al sistema presentaran las características necesarias para poder determinar si el sistema cumple con su objetivo principal a través de la ejecución de las actividades de pronunciación.

La niña asignada a actividades de nivel A, tiene un conocimiento intermedio del alfabeto, razón por la cual se buscó observar si por medio del uso del sistema puede reforzar la asociación mental entre un símbolo del abecedario y su correspondiente fonema, los elementos seleccionados para esta prueba fueron: b, v, i, y, c. El niño asignado al nivel B, posee un correcto conocimiento del alfabeto y los grupos silábicos que se pueden formar con las consonantes, en este caso se buscó observar la efectividad del sistema para practicar y mejorar el aprendizaje de la correcta pronunciación del grupo silábico conformado por la consonante /g/. Se eligió este grupo silábico por tener la peculiaridad de que en las silabas: ga, go, y gu, el sonido de la consonante es suave, mientras que en las silabas: ge y gi, la consonante presenta un sonido fuerte. El niño asignado al nivel C, presenta dislalia evolutiva por sustitución de sonidos entre /r/ y /r/. La frecuencia de sustitución en los sonidos es moderada, sin embargo se trata de un caso que nos permite observar el uso del sistema.

Se observa que en todos los casos se presenta una tendencia positiva en cuanto a las evaluaciones obtenidas entre cada sesión. En las pruebas de nivel A se observó

una mejora promedio del 13% entre cada sesión, para las pruebas de nivel B se observó una mejora promedio del 6% entre cada sesión, finalmente en las pruebas de nivel C se observó una mejora promedio del 6% entre sesiones. La descripción tabular de estas observaciones, se muestra en la Tabla 3.

Tabla 3: Tabla de evaluación entre sesiones de prueba.

Nivel	Evaluación entre sesiones.
A	+ 13%
B	+ 6%
C	+ 6%

A pesar de que los resultados obtenidos son positivos, será necesario observar durante un tiempo mayor los resultados que produce en la capacidad de pronunciación, la aplicación del sistema en las aulas de educación especial para poder realizar una evaluación más completa del sistema. También es necesario recordar que el sistema por si mismo nunca podrá sustituir el apoyo de expertos en el tratamiento de la dislalia, y que el propósito principal del sistema es servir únicamente como herramienta de apoyo.

7 Conclusiones

De acuerdo a las Tablas 2 y 3, se concluye que el software diseñado e implementado puede ayudar progresivamente a niños con déficit del habla por Dislalia en sus procesos de enseñanza y/o rehabilitación. Sin embargo su nivel de eficiencia es directamente proporcional al buen desempeño del módulo de reconocimiento automático del habla que el sistema use y a la capacidad de las actividades desarrolladas que están tengan de captar y mantener la atención del usuario.

Dentro de las herramientas para el reconocimiento automático del habla consultadas, la selección del API de Java Speech como herramienta de desarrollo se justifica por su fácil integración con todas las demás herramientas de desarrollo y su buen desempeño como se observa en las pruebas hechas al módulo de reconocimiento. En cuanto al desarrollo de actividades de pronunciación atractivas a los usuarios se observó que la opción más recomendable para el desarrollo de actividades es por medio de animaciones. La opción más viable fue el SDK de JavaFX, cuya versión 2.0 se presenta como una herramienta completa y robusta para dicho fin.

El desarrollo de un software con fines de apoyo a procesos de rehabilitación de la dislalia, requiere de un amplio soporte de documentación y expertos en el área, pues no es posible plantear actividades de apoyo si no se cuenta con un fundamento teórico que respalde su propósito específico. Lo anterior se concluye de toda la información consultada y recabada sobre este trastorno del aprendizaje y los temas relacionados con el mismo, además de la asesoría proporcionada por los profesores de la escuela de educación especial No. 2 de Tlaxcala. Finalmente se concluye que la metodología de desarrollo basado en prototipos es una opción viable para la

generación de software de apoyo a proceso de rehabilitación, puesto que el entendimiento completo del sistema, tanto por parte de los usuarios finales, como de los desarrolladores fue más fácil de obtener mediante una constante validación de los procesos por medio de prototipos.

Referencias

1. Magaña, R., Lazaro, P.: Trastornos específicos del aprendizaje. *Sociedad Española Española de Pediatría Extrahospitalaria y Atención Primaria (SEPEAP)*. Mayo de (2010).
2. Gómez, P., Bernal, P., Bobadilla, J.: Reconocimiento de voz y fonética acústica. Alfaomega Ra-Ma, España, (2000).
3. Huang, X., Acero, A., Hsiao-Wuen J.: Spoken Language Processing, A Guide to Theory, Algorithm, and System Development. Prentice Hall International Inc, 133 -200. Upper Saddle River, New Jersey, U.S.A. (1994).
4. Lawrence, R., Biing-Hwang, J.: Fundamentals of Speech Recognition. Prentice Hall International Inc, 69 -139. Upper Saddle River, New Jersey, U.S.A. (1993).
5. Jurafsky, D., Martin, J.: Speech and Language Processing. Prentice Hall International Inc. 233- 281. Engewood Cliffs, New Jersey, U.S.A, (1999).
6. CECAPROIN.: SPEECHVIEWER III - Visualizador Fonético. (2010). <http://www.adaptat.com/productos/Comunicacion/speeви.htm>.
7. Consuelo, B.: VISHA (Visualización del habla) (2010). <http://www.uv.es/bellohc/pdf/visha.pdf>
8. Juaresti, J.: Tarjetas reconocedoras de voz para la intervención logopédica. II Congreso de CIIEE 2000, (2000).
9. García, L., Suárez, S.: Análisis de la voz, aplicaciones en sistemas de entrenamiento y rehabilitación. VIII Congreso Internacional de Informática en la Educación. (2002).
10. Oracle, Corp.: Java SE Overview - at a Glance. <http://www.oracle.com/technetwork/java/javase/overview/index.html>
11. Sun Microsystem, Inc.: Java Speech API Programmer's Guide. 35-36. (1998).
12. Calderón, M., Davis, E.: Swing, la solución actual de Java para crear GUIs. (2012). <http://www.dcc.uchile.cl/~Imateu/CC60H/Trabajos/edavis/swing.html>
13. Oracle, Corp.: What is JavaFX? (2012). <http://docs.oracle.com/javafx/2.0/overview/jfxpub-overview.htm>
14. SQLite, Consortium.: About SQLite. (2012). <http://www.sqlite.org/about.html>
15. JFree.org.: Welcome To JFreeChart! (2011). <http://www.jfree.org/jfreechart/>
16. García, L.: Transformada Wavelet Aplicada a la Transformación en Señales de Voz. PhD thesis, Departament de Teoría del Señal i Comunicacions, U. P. Catalunya, España (1998).

BlueNews: modelo para la propagación de información por medio de la tecnología Bluetooth

Oscar D. Ozaine, Alberto Urbina, Arnoldo Díaz-Ramírez, Heber S. Henández y Juan F. Ibáñez

Departamento de Sistemas y Computación
Instituto Tecnológico de Mexicali
{oscar.ozaine, aurbinaes}@gmail.com, {adiaz, heberhdz, pacois}@itmexicali.edu.mx
Paper received on 23/07/12, Accepted on 06/09/12.

Resumen. La alta disponibilidad de dispositivos móviles y el bajo consumo de energía de la tecnología *Bluetooth* ha motivado la proliferación de aplicaciones de cómputo ubicuo. En este artículo se propone *BlueNews*, un modelo para la propagación de información (e.g., noticias) entre usuarios, basado en la tecnología Bluetooth. El modelo propuesto tiene dos modos de operación: modo *Cliente/Servidor* y modo *Servent*. En el primer modo, la aplicación Servidor busca periódicamente dispositivos móviles con el servicio BlueNews instalado (Clientes), para actualizarlos con la información mas reciente que se ha publicado. En el modo Servent, los dispositivos pueden operar como clientes y servidores, intercambiando información entre sí. De esta manera, un dispositivo que no ha tenido conexión al servidor pueda recibir la información mas reciente desde otro dispositivo cliente. Se presentan también los detalles de la implementación de un prototipo con el modelo propuesto.

Palabras Clave: Cómputo Ubicuo, Redes Ad hoc, Bluetooth.

1 Introducción.

Las computadoras personales y los dispositivos estacionarios de comunicación son cada vez menos usados y en algunos casos se han vuelto obsoletos. Esto se debe a que estamos entrando a una nueva era en el área de la computación: la era del *cómputo ubicuo*. El nombre de cómputo ubicuo fue utilizado por primera vez por Mark Weiser [13] para definir la existencia de dispositivos de cómputo distribuidos en un entorno físico, desplegados de manera *invisible* para los usuarios y que sin embargo, llevan a cabo tareas en beneficio de los usuarios.

Con la proliferación y disponibilidad de dispositivos móviles (e.g., teléfonos móviles, tabletas, etc), los usuarios pueden intercambiar información almacenada en sus dispositivos o disponible en la Web. Sin embargo, esta capacidad puede verse afectada por la estructura de la red de comunicación y su dependencia de un punto de acceso centralizado. En este escenario, las redes inalámbricas de propósito es-

pecífico o *Mobile Ad hoc Networks* (MANETs) tienen gran relevancia, ya que proporcionan un método de interconexión más flexible entre dispositivos [7]. Existen diversos estándares de comunicación con los que pueden implementarse redes MANET. Dos de ellos destacan: el estándar IEEE 802.11 para WLANs y la especificación Bluetooth [2] para comunicación inalámbrica de corto alcance. La velocidad de enlace, el rango de comunicación y el poder de transmisión de Bluetooth fueron elegidos para soportar bajo costo y eficiencia en el consumo de energía [9].

En este artículo se presenta *BlueNews*, un modelo para la propagación de mensajes por medio de dispositivos móviles, basado en el uso de Bluetooth. Con *BlueNews*, una organización puede difundir rápidamente información de interés para sus integrantes utilizando el paradigma de cómputo ubicuo. Además, en el artículo se presentan aspectos de implementación de un prototipo utilizando Android.

Este documento está organizado de la siguiente manera. En la Sección 2 se describen los conceptos básicos de las MANETs. Los aspectos más relevantes de tecnología Bluetooth se presentan en la Sección 3. En la Sección 4 se describe el modelo *BlueNews*, así como los protocolos para cada modo de operación. En la sección 5 se discute la implementación de un prototipo del modelo propuesto, utilizando dispositivos con el sistema operativo Android. Finalmente, en la Sección 6 se presentan las conclusiones y el trabajo futuro.

2 MANETs.

Una MANET, también conocida como red *Ad hoc*, es una colección de nodos inalámbricos que pueden organizarse dinámicamente en una topología arbitraria y temporal, para formar una red sin necesidad de usar una infraestructura predefinida. En la Fig. 1 (parte superior) puede observarse una red con infraestructura (e.g., Wi-Fi), que depende de la existencia de un nodo central o punto de acceso. En contraste, una MANET permite la comunicación entre los dispositivos sin la necesidad de un punto de acceso (Fig. 1, parte inferior). En una MANET, los nodos se pueden comunicar utilizando otros nodos como encaminadores si el destino está fuera del rango de comunicación.

La proliferación de dispositivos con capacidades de comunicación inalámbrica ha motivado el desarrollo de aplicaciones móviles (*mobile apps*), así como la necesidad de comunicar dispositivos entre sí de manera sencilla y eficiente. En este contexto, las MANETs tienen gran relevancia, ya que permiten la creación de una red sin la necesidad de una infraestructura predefinida.

Para implementar una red *Ad hoc* se han propuesto diversas tecnologías, entre las que destacan las definidas en los estándares Bluetooth, UWB Ultra-Wideband, HiperLAN/1, HiperLAN/2 y el IEEE 802.11 [7]. La elección de la tecnología de comunicación debe considerar diversos factores, como el costo del hardware, el consumo de energía, el rango de cobertura de la señal, por mencionar algunos. La tecnología Bluetooth representa una de las opciones más interesantes para la implementación de MANETs.

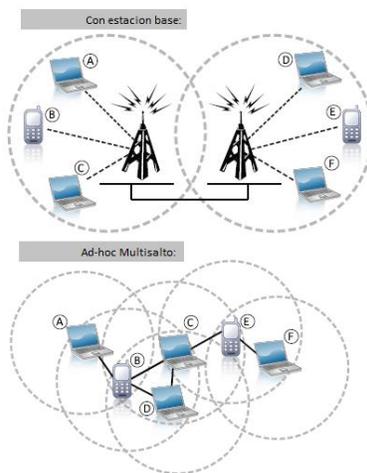


Figura 1. Comparación entre una red con infraestructura y una MANET.

3 Tecnología Bluetooth.

El estándar 802.15.1 [4], también conocido como Bluetooth, se creó para resolver un simple problema: reemplazar los cables usados en dispositivos móviles con comunicación de radio. Es un estándar abierto de conectividad inalámbrica para la transferencia de voz y datos entre dispositivos de comunicación y las computadoras personales, permitiendo a los usuarios la creación de *redes de área personal* (PANs) y de redes *Ad hoc*.

La tecnología Bluetooth opera en la banda sin licencia *Industrial, Científica y Médica* (ISM), que va desde los 2.4 a 2.485 GHz, usando el espectro de propagación, salto de frecuencias y la señal full-dúplex a un rango nominal de 1600 saltos/segundo. La banda ISM 2.4 GHz está disponible y sin licencia en la mayoría de los países. Su rango nominal es de 10 cm a 10 m en teoría, pero puede ser extendido a 100 m a través del incremento de poder de transmisión (con una mayor distorsión). Para reducir la interferencia al mínimo, Bluetooth utiliza saltos de frecuencia adaptados (AFH), que fueron diseñados para detectar otros dispositivos en el espectro y para evitar interferencias cuando muchos dispositivos usan el mismo canal. Esto se logra por saltos de señal entre 79 frecuencias en intervalos de 1 MHz.

Bluetooth soporta hasta 3 canales asíncronos de voz de 64 Kbyte/s cada uno, y datos asíncronos hasta 723.2 Kbyte/s (asimétricos) o 433.9 Kbyte/s (simétricos) [5]. También es un módulo de radio de bajo consumo, que consume hasta 97% menos energía que un teléfono móvil. Además, es *inteligente*, ya que cuando el tráfico de datos disminuye, el transmisor adopta modos de bajo consumo de energía.

La topología de comunicación Bluetooth se basa en la creación de *piconets*, que son subredes formadas por un máximo 8 dispositivos y en las que siempre debe existir un nodo *maestro* y hasta 7 *esclavos*. También se pueden formar *scatternets*, que son la unión de varias piconets por medio del dispositivo maestro, que es el único que se puede comunicar con otros piconets.

Las especificaciones definen todas las capas de la *Pila de Protocolos Bluetooth*. Las tres capas bajas: *Radio*, *Banda Base* y el *Manejador de Enlaces* están agrupados en un subsistema denominado *Controlador Bluetooth* y donde se encuentra la capa L2CAP. La capa SDP y las capas más altas se conocen como *anfitrión Bluetooth*. Esta agrupación de las capas requiere una interfaz de comunicación entre el controlador Bluetooth y el anfitrión Bluetooth. A esta interfaz se le conoce como HCI (*Host Controller Interface*). Las especificaciones Bluetooth hacen posible la compatibilidad entre diferentes sistemas Bluetooth por medio de la definición de protocolos de mensajes, que son intercambiados entre las capas equivalentes. Además, determinan una interfaz común entre los controladores y los anfitriones Bluetooth para hacer compatible los diferentes subsistemas.

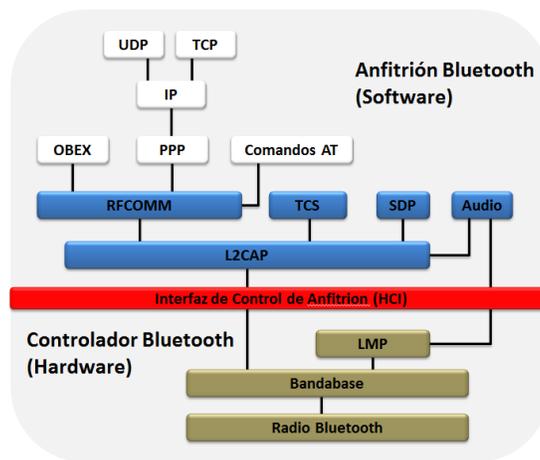


Figura 2. Pila de Protocolos Bluetooth.

Una característica atractiva es que los fabricantes pueden implementar sus propias capas de protocolos propietarios. Por lo tanto, la especificación abierta de Bluetooth incrementa el número de aplicaciones que pueden ser beneficiadas por sus capacidades. Debido a las muchas ventajas de este protocolo, se han desarrollado muchas *Interfaces de Programación de Aplicaciones* (APIs) para construir aplicaciones.

BlueZ [3] es la *Pila de Protocolos Bluetooth* oficial de Linux. Se desarrolló inicialmente por Qualcomm y se distribuye bajo la *Licencia Pública General* (GPL). BlueZ es parte del núcleo oficial de Linux desde la versión 2.4. BlueZ es una gran herramienta para desarrollar aplicaciones de Bluetooth eficientes [12]. La API de BlueZ fue elegida para implementar BlueNews debido a su gran desempeño. Android provee soporte a la pila de Bluetooth desde su versión 2.0 (Android SDK 5) y permite que una aplicación acceda a la funcionalidad Bluetooth por medio de la API de Bluetooth de Android, con el paquete *android.bluetooth*. Android SDK no proporciona funcionalidad para el protocolo L2CAP, que es requerido por la especificación Bluetooth HID. Sin embargo, para utilizar el protocolo L2CAP, que si es pro-

porcionado por BlueZ, es posible que una aplicación Android se comunique directamente con la pila de protocolos BlueZ a través del *Kit de Desarrollo Nativo* (NDK) [1].

4 BlueNews.

BlueNews es un modelo basado en Bluetooth, que permite la difusión de información de interés de un conjunto de usuarios de una organización. El modelo contempla que las aplicaciones puedan operar en dos modos. En el modo *Cliente/Servidor*, existe un servidor que contiene y administra la información que se desea transmitir. Cada mensaje incluye parámetros como la fecha de inicio y final de propagación del mensaje, prioridad y perfil de usuario al que va dirigido, entre otros. Los dispositivos clientes que estén en el rango de alcance del servidor recibirán los mensajes recientes de manera *invisible* para el usuario.

El modo de operación *Servent*, utiliza el paradigma de comunicación *peer-to-peer* (P2P) [11], en el que los nodos pueden operar como clientes y servidores. En este modo, los dispositivos pueden intercambiar información vigente sin necesidad de conectarse al servidor. La Fig. 3 muestra la arquitectura del modelo BlueNews.

Para ilustrar los beneficios del modo *Servent*, puede considerarse el caso de una universidad. Si un estudiante o profesor no asistió al campus universitario el día en que una noticia importante fue publicada, tiene la posibilidad de recibirla si en algún sitio su dispositivo móvil entra en el rango de alcance de otro dispositivo con el servicio BlueNews y que contenga la información actualizada.

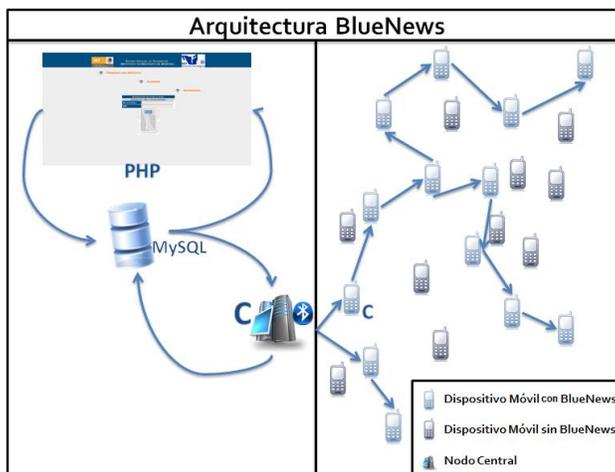


Figura 3. Arquitectura del modelo *BlueNews*.

4.1 Modo de operación Cliente/Servidor.

El algoritmo del modelo BlueNews cuando opera en modo *Cliente/Servidor*, como puede observarse en la Fig. 4, consiste de siete pasos:

1. El nodo central inicia la búsqueda de dispositivos (clientes) en el entorno.
2. Ya que encuentra un dispositivo remoto, se definen los roles *Cliente/Servidor*.
3. Se realiza la conexión L2CAP al dispositivo remoto.
4. El nodo central verifica si el dispositivo remoto tiene el servicio BlueNews. Si el dispositivo remoto no tiene el servicio BlueNews, ir a paso 7.
5. El dispositivo remoto envía al servidor el último mensaje que recibió.
6. Si existen nuevos mensajes, el servidor los envía al dispositivo remoto.
7. Termina la conexión L2CAP.



Figura 4. Diagrama de flujo del modo Cliente/Servidor.

4.2 Modo de Operación Servent.

El algoritmo utilizado por el modo *Servent* se muestra en la Fig. 5. El modo *Servent* considera la existencia de dos *hilos (threads)* de ejecución. El primer hilo de ejecución, denominado *hilo de investigación*, busca periódicamente en el entorno dispositivos con la aplicación BlueNews. El algoritmo de este hilo consta de seis pasos:

1. El nodo *Servent* inicia la búsqueda de dispositivos en el entorno.
2. Ya que encuentra un dispositivo remoto, define los roles *Cliente/Servidor* (e.g., el nodo con el mayor valor de su dirección MAC adquiere el rol de servidor).
3. Cuando los roles están definidos, se realiza la conexión al dispositivo remoto que encontró.
4. Se verifica si el dispositivo remoto tiene el servicio BlueNews. Si el dispositivo remoto no tiene el servicio BlueNews, ir a paso 6.
5. Se lleva a cabo la actualización.

6. Termina la conexión L2CAP.

El segundo hilo de ejecución, denominado hilo *listen*, espera por una conexión de algún dispositivo con la aplicación BlueNews, y consta de 6 pasos:

1. El nodo *Servent* siempre se encuentra en espera de nuevas conexiones entrantes.
2. Al momento de recibir una conexión, se definen los roles Cliente/Servidor.
3. Se lleva a cabo la conexión L2CAP al dispositivo remoto.
4. El dispositivo remoto envía una verificación de servicio, al que éste responde que si tiene el servicio BlueNews.
5. Se lleva a cabo la actualización.
6. Termina la conexión L2CAP.

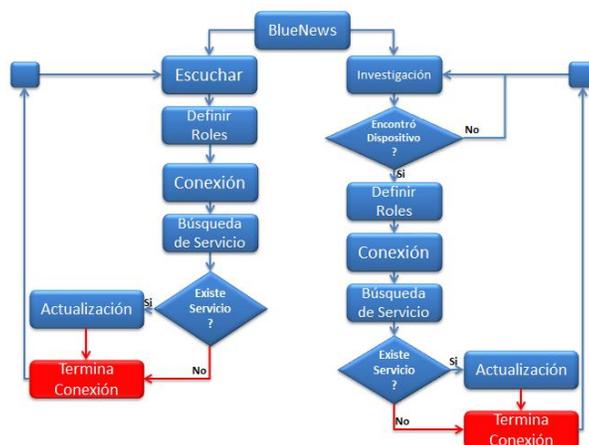


Figura 5. Diagrama de flujo del modo *Servent*.

Cuando un dispositivo Bluetooth ofrece un servicio, actúa como servidor. En esta situación es necesario especificar el tipo de servicio que se está ofreciendo y registrarlo. Los dispositivos Bluetooth utilizan el *Protocolo de Descubrimiento de Servicios* (SDP), en donde un cliente puede descubrir los servicios disponibles en otros dispositivos Bluetooth, así como los atributos de los servicios.

En contraste, el modelo BlueNews no registra un servicio Bluetooth. En su lugar, utiliza una aproximación similar al protocolo Gnutella [6]. En BlueNews, una vez que el *Servent* se ha identificado y ha establecido una conexión con el nodo central u otro *Servent* remoto, envía al *Servent* remoto una cadena de caracteres como petición de servicio: “BlueNews”. El *Servent* con el servicio BlueNews debe responder con “BlueNewsOk”. Este esquema reduce el consumo de recursos y tiempo ya que evita múltiples conexiones SDP para búsqueda de servicios.

Existen varios protocolos P2P, entre los que destacan Gnutella [10] y Freenet. En [8] se presentan los resultados de una evaluación comparativa de estas dos arquitecturas. Gnutella se desarrolló para el descubrimiento de nodos en la red y el intercambio de ficheros. Freenet tiene una topología de red similar a Gnutella, excepto

que su propósito principal es crear un sistema de almacenamiento de información global seguro. Para nuestro modelo se eligió implementar una aproximación del protocolo Gnutella ya que está orientado más al intercambio de información que al almacenamiento. Además, permite que la aplicación funcione de una manera más eficiente al hacer la búsqueda de servicios.

Para determinar el rol del *Servent* (servidor o cliente), la dirección MAC de cada dispositivo Bluetooth encontrada es utilizada. El *Servent* que tenga la dirección con menor valor actúa como cliente y debe hacer una petición de conexión al *Servent servidor*. Además, cada dispositivo debe ser capaz de atender múltiples conexiones entrantes y salientes. Esto puede hacerse implementando una aplicación multi-hilos.

Después de la petición de conexión, el *Servent cliente* verifica si la dirección remota pertenece a un dispositivo que ya se haya registrado en una lista *sin servicio BlueNews*. Esto se hace con el propósito de evitar conexiones repetidas con dispositivos que no tienen el servicio BlueNews. Si la dirección remota no es rechazada, una conexión L2CAP al *Servent* remoto es creada.

Ya que la conexión L2CAP se hizo con éxito, el *Servent servidor* envía el último número de mensaje insertado en su base de datos hacia el *Servent cliente*. Una vez que el cliente lo recibe, hace una resta del número de mensaje más reciente menos el número de mensaje más reciente del servidor. Con el resultado, que se envía al servidor, se toma una decisión. Si el resultado de la resta es igual a 0, termina la conexión L2CAP. Si es menor a 0, el *Servent servidor* actualiza al *Servent cliente* y termina la conexión L2CAP. Si es mayor a 0, el *Servent cliente* actualiza al *Servent servidor* y termina la conexión L2CAP.

Se eligió como protocolo de comunicación a L2CAP sobre RFCOMM debido a que con L2CAP la conexión se realiza de manera transparente al usuario, ya que trabaja en una capa por debajo de RFCOMM. El uso del protocolo L2CAP permite una mayor velocidad de conexión, evita la intervención del usuario, tanto como emparejamiento con otros dispositivos como para aceptar o denegar la conexión a estos. Es importante recordar que uno de los objetivos del modelo propuesto es que el usuario pueda llevar a cabo el intercambio de información con los demás dispositivos de manera automática, es decir, sin la intervención del usuario.

5 Implementación del Prototipo.

Se implementó un prototipo para comprobar la funcionalidad del modelo propuesto. La aplicación *Servidor* BlueNews se implementó en una computadora personal equipada con el sistema operativo Linux (Ubuntu 12.04). Se utilizó el lenguaje de programación C para el desarrollo de la aplicación y el API de BlueZ para la comunicación Bluetooth. Las aplicaciones *Cliente* y *Servent* se implementaron en dispositivos móviles con el sistema operativo Android 4.0. El lenguaje de programación utilizado fue Java y para el manejo de los aspectos de comunicación Bluetooth se utilizó el paquete *android.bluetooth*, así como la herramienta NDK para poder utilizar el protocolo L2CAP en Android. Para obtener un mejor desempeño, las aplicaciones *Servidor*, *Cliente* y *Servent* utilizan múltiples hilos.

Las aplicaciones *Cliente* y *Servent* tienen dos módulos principales, denominados *backend* y *frontend*. El módulo *backend* lleva a cabo las funciones relacionadas con

la recepción y envío de mensajes por medio de Bluetooth. El *frontend* es el encargado de mostrar los mensajes al usuario, configurar el funcionamiento de la aplicación, así como definir los perfiles de usuario, como puede observarse en la Fig. 6.

Las pruebas que se hicieron con el prototipo implementado mostraron que el modelo funciona correctamente en sus dos modos de operación: *Cliente/Servidor* y *Servent*. La Fig. 7 muestra un ejemplo del uso del prototipo.

Para el control y almacenamiento de los mensajes, la aplicación *Servent* utilizó una base de datos SQLite. Para que los usuarios autorizados puedan generar nuevos mensajes, además de la aplicación Android, se desarrolló una aplicación Web utilizando el lenguaje de programación PHP.



Figura 6. Módulo de configuración de la aplicación Cliente.



Figura 7. Ejemplo del módulo *frontend* de la aplicación Cliente.

6 Conclusiones y Trabajo Futuro.

En este artículo se presentó BlueNews, un modelo basado en la tecnología Bluetooth que permite la propagación de información utilizando el paradigma de cómputo ubicuo. El modelo propuesto tiene dos modos de operación: Cliente/Servidor y Servent, que permiten a los usuarios recibir información importante de manera automática y en diversos escenarios.

Para validar el modelo propuesto se implementó un prototipo utilizando herramientas de código abierto, entre las que destacan el API de BlueZ y el sistema operativo Android. La evaluación del prototipo mostró que el modelo propuesto funciona correctamente y que opera de manera invisible a los usuarios. En el artículo se presentan algunos aspectos importantes de la implementación del prototipo.

Como trabajo futuro se pretende implementar BlueNews en diferentes plataformas para dispositivos móviles.

Referencias

1. Android NDK official web site: <http://developer.android.com/tools/sdk/ndk/index.html>.
2. Bluetooth official web site. <http://www.bluetooth.com>.
3. Bluez official web site. <http://www.bluez.org>.
4. The IEEE 802.15.1 Standard. , June 2002..
5. Bluetooth SIG. Specification of the Bluetooth System. , 2001.
6. Woosik Kim Ho-Hyun Park. A Gnutella-based P2P System Using Cross-Layer Design for MANET. , 2007.
7. Chlamtac, M. Conti, J. Liu. Mobile ad hoc networking: imperatives and challenges. *Ad Hoc Networks*, 1(1):13 — 64, 2003.
8. Jem E. Berkes. Decentralized Peer-to-Peer Network Architecture: Gnutella and Freenet. 2003.
9. Pravin Bhagwat. Bluetooth: Technology for Short-Range Wireless Apps. *IEEE Inter*, 5(3):96—103, 2001.
10. M. Ripeanu. Peer-to-peer architecture case study: Gnutella network. *Proceedings of the 1st IEEE International Conference on Peer to Peer Computing*:99—100, 2001.
11. Stephanos Androutsellis-Theotokis, Diomidis Spinellis. A Survey of Peer-to-Peer Content Distribution Technologies. *ACM Computing Surveys*, 36(4):335—371, 2002.
12. Sukey Nakasima, Francisco Reyna, Arnoldo Díaz-Ramírez, Carlos T. Calafate. First Experiences With BlueZ. *Research in Computer Science*, 39:97—114, June 2008.
13. Mark Weiser. The computer for the 21st century. *Scientific American*, 265(3):94—104, 1991.

Uso de Ontologías y otras Fuentes de Datos para la Generación de Currículos Vitae.

Kenenias Bethuel Pérez-Betanzos¹, Beatriz Alejandra Olivares-Zepahua²,
Giner Alor-Hernandez³, Ana María Chávez -Trejo⁴

^{1,2,3,4}Instituto Tecnológico de Orizaba
Orizaba, Veracruz, México

¹kenepz@acm.com, ²bolivares@ito-depi.edu.mx,
³galor@ito-depi.edu.mx, ⁴achavez@ito-depi.edu.mx
Paper received on 24/07/12, Accepted on 05/09/12.

Resumen. En el presente trabajo se describe un prototipo para la recuperación de información de diferentes fuentes de información (ontologías y bases de datos) y la generación de Currículos Vitae en documentos PDF. En particular, se presenta el diseño de la arquitectura mostrando su vista lógica, vista de desarrollo y los casos de uso o escenarios; se presenta también la implementación de un prototipo desarrollado en lenguaje Java que usa el marco de trabajo Java ServerFaces (JSF) para la construcción de las interfaces de usuario, la API iText para la generación de Currículos Vitae (CV) en documentos PDF, los lenguajes de consulta SQL y SPARQL para la consulta a las fuentes de información, Jena y JDBC para el acceso a las fuentes de información. El propósito general de este trabajo es mostrar los beneficios del uso de las ontologías en los sistemas de recuperación de información.

Palabras Clave: Web Semántica, SPARQL, SQL, Fuentes de Información, Ontologías, Bases de Datos, Java

1 Introducción

La Web Semántica es la evolución de la Web actual en la que el contenido es procesable automáticamente a escala global. Su objetivo principal es permitir, tanto a humanos como a máquinas, encontrar, compartir y combinar información de manera sencilla y automatizada, es decir, tener una Web más útil [1].

El núcleo de la Web Semántica son las ontologías, una ontología es una especificación formal y explícita de una conceptualización compartida [2]. Una ontología está formada por una taxonomía que define clases y sus relaciones, y un conjunto de reglas de inferencia que tienen un conocimiento compartido y común para cualquier dominio entre aplicaciones y personas. Entre los beneficios que se tienen en el uso de las ontologías con respecto a otras fuentes de datos como las bases de datos relacionales son interoperabilidad, búsqueda, reutilización y reestructuración [3].

En el presente artículo se muestran la arquitectura y un prototipo de una aplicación Web para la generación de formatos de Currículos Vitae haciendo uso de una base de datos y una ontología de productividad de la Maestría en Sistemas Computacionales del Instituto Tecnológico de Orizaba, el prototipo fue desarrollado en el lenguaje de Java y usa un editor de formatos para mostrar, adecuar y guardar cada formato de Currículo Vitae para su exportación a documentos PDF.

2 Problema

Actualmente en la División de Estudios de Postgrado e Investigación (DEPI) del Instituto Tecnológico de Orizaba (ITO), los profesores de la Maestría en Sistemas Computacionales entre sus actividades administrativas necesitan actualizar constantemente sus Currículos Vitae (CV). Estos Currículos Vitae se presentan bajo distintos formatos de acuerdo a la dependencia que lo solicita como PROMEP, CONACYT y DGEST entre otras. Estas actividades de actualizar, modificar y adecuar cada formato de CV conllevan un aumento en la carga de trabajo de los profesores ya que es un distractor respecto a las tareas prioritarias propias de sus cargos.

Los formatos de CV difieren entre sí de acuerdo a la dependencia solicitante y, en ocasiones, a la plaza del investigador (tiempo completo o tiempo parcial). Sin embargo, el CV principal se mantiene en el sistema del Consejo Nacional de Ciencia y Tecnología (CONACYT) y es exportable a un documento PDF pero dicho sistema no incluye consultas sobre la información, de tal forma que recuperar el CV en un formato distinto implica necesariamente la recaptura de información.

3 Propuesta

Para dar solución al problema mencionado anteriormente, se propone aprovechar el uso de las tecnologías Web 2.0 y Web semántica, para desarrollar una aplicación Web basada en una ontología de productividad con la cuál se genera el CV de cada profesor de la Maestría en Sistemas Computacionales del ITO bajo distintos formatos exportables a documentos PDF y Word, los datos de los CV se recuperan de las siguientes fuentes de información: 1) la ontología de productividad de la Maestría en Sistemas Computacionales del ITO, 2) la base de datos de Recursos Humanos del ITO y 3) una ontología que mantiene los datos de cada CV que no se encuentran incluidos en las dos fuentes antes mencionadas.

El principal propósito de este trabajo, es mostrar los beneficios del uso de las ontologías con respecto a otras fuentes de datos en los sistemas de recuperación de información, ya que como se mencionó anteriormente los formatos de CV contienen muchos datos similares que solo varían por sinónimos o posición dentro del documento o algunos datos que no son comunes en todos ellos, por lo que el uso de la ontología nos ayuda a tener una clara especificación de los datos, agregando a las propiedades sinónimos o anotaciones teniendo así más reusabilidad y fiabilidad en los datos entre otros beneficios[4].

4 Descripción de la aplicación.

El desarrollo de esta aplicación Web se basa en un modelo de desarrollo de prototipos incrementales [5], este modelo nos ayuda a evaluar el diseño y la implementación, las funcionalidades, así como tener una interacción con la aplicación para mejorar la usabilidad de cada prototipo.

En esta aplicación Web se usa una ontología de productividad de la Maestría en Sistemas Computacionales del ITO, que recupera la productividad de los profesores de la misma respecto a información de proyectos de investigación, artículos publicados y libros entre otros; esta ontología de productividad se usa como fuente de datos primaria para la aplicación, se usa también como fuente de datos la base de datos de Recursos Humanos del ITO, esta base de datos almacena los datos personales de los profesores de la Maestría en Sistemas Computacionales del ITO como nombre, dirección y RFC entre otros; para completar los demás datos faltantes de los formatos de los Currículos Vitae se desarrolló una ontología que mantiene los datos de los profesores que no forman parte ni de la productividad ni de los datos personales, tales como distinciones, estancias de investigación y participación en congresos entre otros; esta última ontología extrae sus datos de los CV de CONACYT que se encuentran en documentos PDF.

A continuación se muestra en la figura 1 una parte de las clases que se crearon para la ontología basada en los datos analizados e identificados de los distintos formatos de CV, dicha ontología se modeló por medio de la herramienta Protégé[6] con el lenguaje OWL [7] (Web Ontology Language, Lenguaje de Ontología Web).

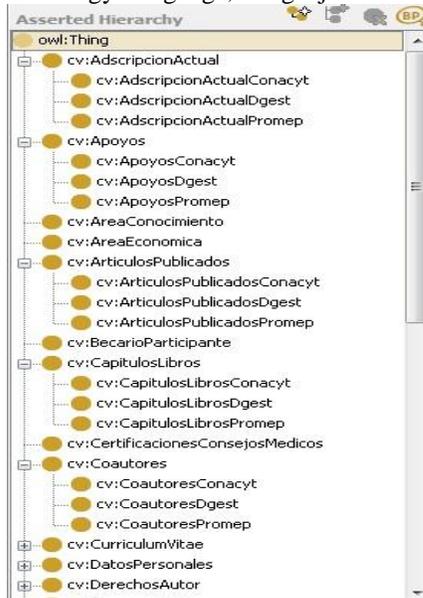


Figura 1. Estructura de las clases de la ontología basada en el CV de CONACYT.

El dominio de esta ontología consiste de los conceptos y sus relaciones con los subconceptos de los datos y secciones identificadas en los CV de CONACYT, DGEST y PROMEP. Las relaciones de los conceptos y sus subconceptos forman un árbol de la taxonomía para CV. Cada una de las clases y subclases de la ontología tienen un conjunto de propiedades para la relación entre los diferentes individuos [8].

Esta aplicación usa el lenguaje Java y el marco de trabajo Java Server Faces (JSF) [9] para construir las interfaces de usuario, la validación de la captura, control de eventos y manejo de estados entre otros.

En la figura 2 se presenta el Diagrama de Componentes, que describe los componentes de software, cada uno de los cuales proporciona una interfaz para la comunicación e interacción entre los mismos. Los componentes mostrados en este diagrama son los que hasta el momento se han identificado, conforme se avance en los prototipos es probable que surjan nuevos componentes. En este Diagrama de Componentes se tienen 8 componentes: 1) Componente CKE ditor que mantiene el control de las funcionalidades del editor de formatos para los CV, 2) Componente Elemento Edición controla el manejo de las plantillas para los formatos de los CV, 3) Componente iTextPDF controla la funcionalidad para la generación de documentos PDF, 4) Componente PrimesFaces controla el manejo y la funcionalidad para AJAX y RIAs, 5) Componente JSF BackingBeans contiene Java Beans que obtienen valores de controles e implementan métodos listener, 6) Componente ModelBeans contiene todas las clases que representan la información de cada CV, 7) Componente API Jena controla y maneja la funcionalidad para consulta, recuperación y modificación de los datos en las ontologías, 8) Componente API JDBC controla y maneja la funcionalidad para la consulta, recuperación y modificación de los datos en la base de datos.

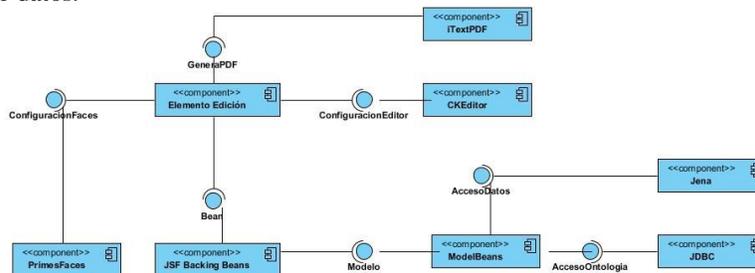


Figura 2. Vista de Desarrollo – Diagrama de Componentes.

5 Arquitectura de la aplicación.

La arquitectura de la aplicación se desarrolló bajo el modelo de 4 +1 vistas [10], este modelo ayuda a tener una relación con el diseño y la implementación de la estructura de la aplicación. En este trabajo se muestran los Casos de Uso o Escenarios que son una abstracción de los requisitos funcionales más importantes de la aplicación, la Vista Lógica que es una abstracción tomada del dominio del problema en la forma de objetos o clases de objetos y la Vista de Desarrollo que está cen-

trada en la organización de los módulos de software a desarrollar; la Vista de Procesos y la Vista Física no se incluyen en este documento debido a los cambios que experimentan con la aplicación del modelo de prototipos incrementales. A continuación se muestran las 3 vistas de la arquitectura mencionadas.

5.1 Casos de Uso o Escenarios

El diagrama de Casos de Uso en UML [11] para la aplicación se realizó con la herramienta Visual Paradigm[12], y se muestra en la figura 3. En este diagrama está incluida la funcionalidad de la aplicación en su conjunto, independientemente de la prioridad de los distintos prototipos. Existen dos actores, el administrador y el profesor, con 16 casos de uso que se describen a continuación:

1. Agregar Datos CV (Ontología) – Agrega datos de los Currículos Vitae a la Ontología, solo para el actor administrador.
2. Agregar Datos CV (BD) – Agrega datos a la base de datos de Recursos Humanos, solo para el actor administrador.
3. Modificar Plantilla CV – Modifica la plantilla de CV y guarda los cambios realizados, solo para actor administrador.
4. Deshabilitar Plantilla CV –Deshabilita la plantilla de CV, solo para el actor administrador.
5. Crear Tipo Formato CV – Crea un nuevo formato de CV, solo para el actor administrador.
6. Crear Plantilla CV – Crea una nueva plantilla de CV, solo para el actor administrador.
7. Mostrar Detalles CV –Muestra los datos del CV seleccionado por el usuario, solo para el actor administrador.
8. Modificar Tipo Formato CV – Modifica un formato de CV, solo para el actor administrador.
9. Eliminar Tipo Formato CV – Elimina un formato de CV, solo para el actor administrador.
10. Registrar Usuario – Agrega un nuevo usuario, solo para el actor administrador.
11. Modificar Usuario – Modifica los datos de un usuario y los guarda, solo para el usuario administrador.
12. Eliminar Usuario – Elimina a un usuario, solo para el actor administrador.
13. Iniciar Sesión – Ingreso a la aplicación, para el actor administrador y profesor.
14. Validar Usuario – Válida el ingreso a la aplicación para los usuarios.
15. Mostrar Detalles Usuario – Muestra los datos de cada usuario, solo para el actor profesor.
16. Generar CV – Genera el CV del usuario ingresado, solo para el actor profesor.

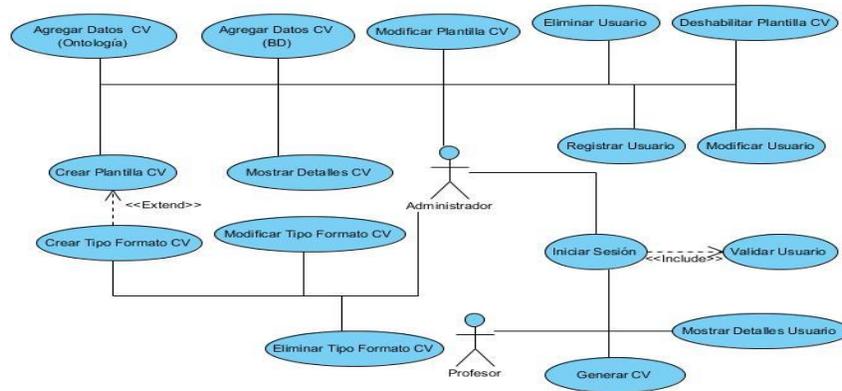


Figura 3. Diagrama de casos de uso

5.2 Vista de desarrollo

En la figura 4 se muestra la vista de desarrollo representada por un Diagrama de Paquetes que utiliza un estilo arquitectónico Modelo-Vista-Controlador [13].

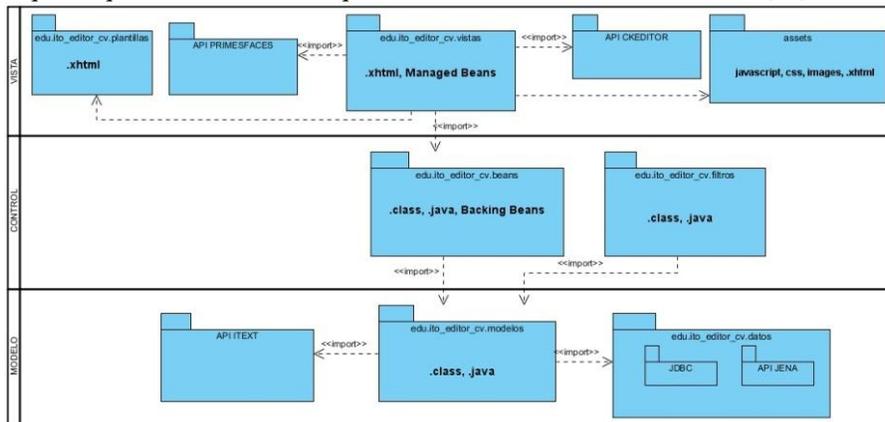


Figura 4. Vista de desarrollo – Diagrama de Paquetes

En la parte de la Vista se tienen 5 paquetes, 1) el paquete Plantillas es el encargado de mantener los formatos genéricos de CV, 2) el paquete API PrimesFaces [14] contiene toda la funcionalidad para el manejo de RIAs (Aplicaciones Enriquecidas de Internet) y de AJAX (Asynchronous JavaScript And XML, JavaScript asíncrono y XML), 3) el paquete Vista contiene todas las páginas en XHTML, 4) el paquete API CKEditor [15] contiene la funcionalidad para el manejo del editor de formatos que modificará, adecuará y guardará los cambios de cada uno de los formatos de Currículos Vitae, 5) el paquete Assets almacena todos los archivos JavaScript, hojas de estilo e imágenes que se usan en la aplicación.

para los CV, guarda los cambios hechos en cada plantilla de CV, y por último exporta el formato de CV previamente llenado hacia documentos PFD y Word.

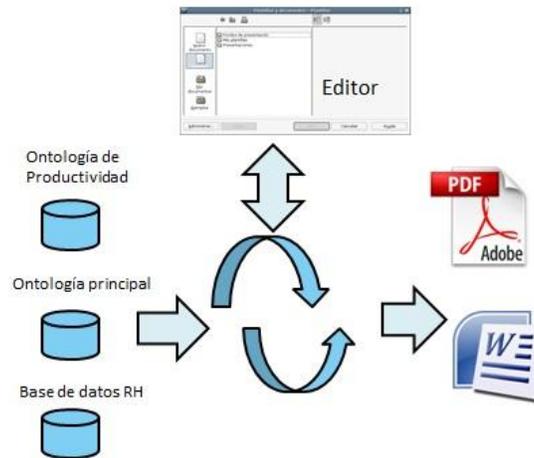


Figura 6.Esquema de la aplicación

El primer prototipo desarrollado genera un formato de CV genérico que no pertenece a los formatos oficiales de CONACYT, PROMEP o DGEST y que funciona como prueba de concepto para verificar la correcta relación entre las distintas tecnologías descritas en la Vista de Desarrollo mencionada en la sección anterior. El prototipo consiste en una pequeña interfaz que solicita el RFC del profesor y con base en él se genera el formato de CV para dicho profesor, esta interfaz se muestra en la figura 7.



Figura 7. Interfaz prototipo para la generación de CV por medio del RFC.

Con el marco de trabajo JSF se construyeron las interfaces de usuario y se manejan los eventos de la aplicación con los JSF BackingBeans, la recuperación de datos se obtiene a partir de las fuentes antes mencionadas a través de JDBC y la API Jena por medio de los lenguajes de consulta SQL[19] y SPARQL [20], los datos recuperados son gestionados por medio de los JB, con los datos recuperados de las fuentes se genera el formato de prueba de un CV genérico, con la ayuda de la API iText que realiza la exportación a documentos PDF, en la figura 8 se muestra una sección del formato de prueba del CV genérico. Actualmente se está trabajando en el desarrollo de los siguientes prototipos para incorporar las demás funcionalidades presentadas en la arquitectura de la aplicación.

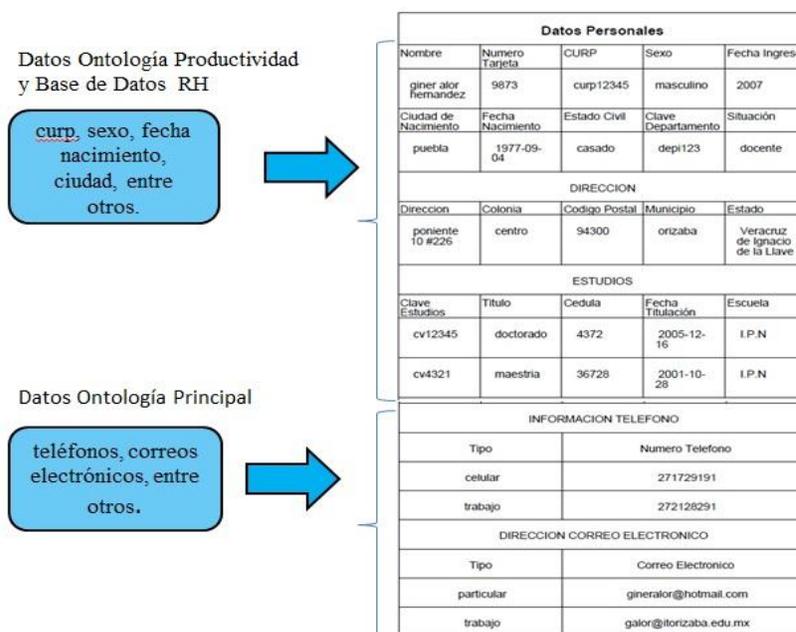


Figura 8. Formato de CV en documento PDF generado con iText

7 Conclusiones.

En este trabajo se presentó una arquitectura para una aplicación Web que genera CVs, esta arquitectura se encuentra basada en el modelo de 4 +1 vistas, cada una de estas vistas manejan los distintos requisitos funcionales y no funcionales separadamente, así como los diferentes intereses (componentes, conectores, contenedores entre otros) de los involucrados en el desarrollo como son: programadores, analistas, diseñadores, entre otros. Esta arquitectura se apoya en un desarrollo iterativo para lograr un mejor diseño, agregar nuevas funcionalidades a medida que se generan los diferentes prototipos incrementales, tener una notación común por medio de UML y manejar la escalabilidad en el desarrollo de la aplicación.

Los resultados obtenidos hasta el momento se reflejan en el desarrollo de una aplicación que está basada en un modelo de prototipos incrementales. El primer prototipo, mostrado en este trabajo, es desarrollado en el lenguaje Java con el marco de trabajo JSF, que maneja la recuperación de datos de cada una de las fuentes de información por medio de JDBC y la API de Jena a través de los lenguajes de consultas SQL y SPARQL respectivamente, los datos recuperados se usan para generar un formato de CV genérico en documentos PDF por medio de la API iText, este prototipo se usó para mostrar la interacción y la compatibilidad entre las diferentes tecnologías antes mencionadas.

8 Trabajos a futuro.

Como trabajos a futuros se tiene el complementar la arquitectura agregando la vista de procesos y la vista física, así como refinar las vistas de desarrollo y lógica con base en los hallazgos que pudieran surgir durante el desarrollo de los prototipos faltantes.

Se continuará con el desarrollo de los prototipos incrementales para agregar las siguientes funcionalidades: generar los CVs de CONACYT, DGEST y PROMEP en documentos PDF y documentos Word, desarrollar las plantillas para el manejo de cada uno de los CVs, implementar el editor de plantillas para modificar, adecuar y guardar los cambios en cada una de las plantillas de CVs y diseñar todas las interfaces de usuario para la aplicación.

Referencias

1. Guía Breve de la Web Semántica. [En línea]. Disponible: <http://www.w3c.es/divulgacion/guiasbreves/websemantica>.
2. J. Contreras y J. A. Martínez Comeche.: Tutorial de Ontologías, Universidad Complutense de Madrid, España, 2001.
3. L-van Tang, Hongyan Lu, Baojun Qiu, Meimei Li, Jianjun Wang, Lei Wang, Bin Zhou, Dongqing Yang, Shiwei Tang.: WISE: Prototype for Ontology Driven Development of Web Information Systems, Peking University, Beijing, China.
4. Urvi Shah, Tim Finin, Anupam Joshi.: Information Retrieval On The Semantic Web, University of Maryland, Baltimore County, EUA.
5. Jacobson, I., Booch, G., Rumbaugh, J.: El Proceso Unificado de Desarrollo, Addison Wesley, 2000
6. Horridge Matthew, Knublauch Holger, Rector Alan, Stevens Roberts.: A Practical Guide To Building OWL Ontologies Using Protégé and CO-ODE Tools Edition 1.3, The University of Manchester, March 24, 2011, EUA.
7. McGuinness Deborah L. and Harmelen Frank Van.: OWL Web Ontology Language, [En línea]. Disponible: <http://www.w3.org/TR/owl-features>
8. Maedche and Staab.: Ontology Learning for the Semantic Web, University of Karlsruhe, 2011.
9. “JavaServer Faces Technology - Documentation”, [En línea]. Disponible: <http://www.oracle.com/technetwork/java/javaee/documentation/index-137726.html>
10. Kruchten Philippe.: Planos Arquitectónicos: El Modelo de 4 + 1 Vistas de la Arquitectura del Software, Marzo 2006.
11. Booch Grady, Rumbaugh James, Jacobson Ivar.: Unified Modeling Language User Guide, 2nd Edition, Edit. Addison Wesley, 2005.
12. Visual Paradigm User’s Guide. [En línea]. Disponible: <http://www.visual-paradigm.com/support/documents/vpumluserguide.jsp>
13. Bass Len, Clements Paul, Ken Bass.: Software Architecture in Practice, Edit. Addison Wesley, December 30, 1997.
14. PrimesFaces – Documentation [En línea]. Disponible: <http://primefaces.org/documentation.html>
15. CKSource Docs -The Official Documentation Site [En línea]. Disponible: <http://docs.cksource.com/>
16. API Documentation iText [En línea]. Disponible: <http://api.itextpdf.com/>
17. Oracle documentation: Getting Started with Java DB, June 19, 2012.

18. JeongDongwon, Shin Heeyoung, Baik Doo-Kwon and Jeong Young-Sik.: An Efficient Web Ontology Storage Considering Hierarchical Knowledge for Jena-based Applications, Kunsan National University, Korea, March 17, 2009.
19. K. Kevin: SQL in a Nutshell: A Desktop Quick Reference, Edit.O'Really Media
20. H. Steve: SPARQL query processing with conventional relational database systems, University of Southampton, UK.

Redes Sociales para el Seguimiento de Egresados

Briceida González-Jiménez¹, Gerardo Contreras-Vega²
María Karen Cortés-Verdín³

^{1,2,3} Universidad Veracruzana, Facultad de Estadística e Informática,
Xalapa, Ver., México

¹briceida9@gmail.com, ²gcontreras@uv.mx, ³kcortes@uv.mx

Paper received on 23/07/12, Accepted on 05/09/12.

Resumen. Las redes sociales en Internet constituyen hoy en día un medio de comunicación altamente utilizado y su poder ha sido explotado de manera ventajosa por diversas organizaciones en variados dominios. Por su parte, los sistemas de seguimiento de egresados tienen el objetivo de proporcionar información acerca del desempeño laboral de los egresados con el fin de retroalimentar a su programa educativo de egreso. El principal problema de un sistema de seguimiento de egresados es contactar al egresado ya que para ello se requiere de un directorio actualizado de los datos del egresado. En este caso la Facultad de Estadística e Informática decidió implementar una red social dentro de su sistema de seguimiento de egresados, como la solución al problema del directorio actualizado de egresados y, al mismo tiempo, como un medio de comunicación permanente con la comunidad de egresados. Este desarrollo se encuentra basado en un CMS *Open Source* que ha facilitado el desarrollo y además, ha permitido sentar las bases para la evolución de la red social para el seguimiento de egresados. En este artículo se presenta el desarrollo de dicha red.

Palabras Clave: Redes sociales, Seguimiento de Egresados, Facebook, Twitter, Open Source, CMS, gestor de contenidos.

1 Introducción.

El seguimiento de egresados es una estrategia evaluativa que posibilita conocer la situación profesional o laboral de un egresado de una carrera profesional con la finalidad del mejoramiento institucional [1]. Debido a la necesidad de la Facultad de Estadística e Informática (FEI) de la Universidad Veracruzana, de implementar un Sistema de Seguimiento de Egresados bajo la guía de la metodología propuesta por la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) [2] (llamada Esquema Básico para el Seguimiento de Egresados) la FEI decide iniciar el proyecto para la implementación un Sistema de Seguimiento de Egresados.

Durante este proyecto surgió la necesidad de elaborar un directorio de egresados actualizado. La falta de un directorio actualizado de egresados ha causado algunos

problemas en el pasado cuando la FEI ha tenido la necesidad de contactarlos. Dado la importancia que en los últimos tiempos han cobrado las redes sociales en Internet, se pensó en implementar una red social que permitiera contar con el directorio de egresados actualizado.

Las redes sociales se han convertido en una forma de sostener comunicación con otros usuarios y han ganado popularidad en Internet. Las redes sociales además de la comunicación pueden tener distintas funcionalidades, como el Marketing, el entretenimiento, fines laborales y difusión de temas. Además algunas universidades han hecho uso de este tipo de herramienta web para apoyar a sus estudiantes en sus estudios.

Se eligió trabajar con redes sociales ya que es un medio de comunicación que tiene gran éxito, 6 de cada 10 mexicanos accede a una red social, de los cuales el 78% de esos usuarios de Internet se encuentra entre los 6 y 34 años [3]. Además, se realizó una encuesta a estudiantes y egresados de la FEI (que comprende dos programas educativos: Informática y Ciencias y Técnicas Estadísticas). A partir de una muestra de población de 50 personas, el 94% afirmó utilizar Facebook y un 46% Twitter. Por otra parte, las redes sociales cuentan con herramientas que facilitan la interacción entre usuarios mediante el envío de mensajes, compartiendo audio, video y fotografías. Para el desarrollo de la red social se utilizó un gestor de contenidos o CMS, el cual debía contar con una licencia de software libre que permitiera las adecuaciones necesarias y de esta forma ajustarse a las necesidades de la FEI.

Actualmente existen tecnologías como los Gestores de Contenido (*Content Manager System*, CMS en sus siglas en inglés) que permiten desarrollar una red social de manera eficiente, optimizando el tiempo de desarrollo y reduciendo los costos. Además algunos de ellos son de licencia *Open Source* (código abierto); el código abierto permite ser modificado y distribuido libremente, además está creado en colaboración con desarrolladores de diversas partes del mundo, quienes se encargan de revisar la calidad y funcionalidad del software.

En el presente artículo se presenta el desarrollo de la red social para el seguimiento de egresados de la Facultad de Estadística e Informática y para tal efecto se encuentra organizado de la siguiente manera: en primer lugar se presentan los antecedentes. En segundo lugar se explica de forma el desarrollo de la red social. En el tercer apartado se aborda el desarrollo del plugin para obtener datos de los egresados a partir de sus redes sociales por medio de su registro a la red. En el cuarto apartado se aborda el tema de los resultados, donde se explica cómo se puso a prueba la red social de seguimiento de egresados y las impresiones de los egresados. Finalmente se encontraron las conclusiones del trabajo realizado así como la descripción de los trabajos futuros.

2 Antecedentes.

El estudio de seguimiento de egresados es una vía para determinar las exigencias laborales actuales, así como situación profesional del egresado. Además, el estudio de seguimiento de egresados permite ofrecer mayor calidad en la educación ya que hace posible determinar actualizaciones en los planes académicos. El estudio de seguimiento de egresados es un aspecto importante considerado por esquemas de

acreditación como pueden ser los del Consejo Nacional de Acreditación en Informática y Computación (CONAIC) y de los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES). De igual forma es tomado en cuenta por la ANUIES. La Facultad de Estadística e Informática (FEI) cuenta con dos programas educativos el de la Licenciatura en Informática y la Licenciatura en Ciencias y Técnicas Estadísticas y recientemente pasó por el proceso de acreditación por parte de CONAIC. Una de las recomendaciones que hizo fue la de realizar un seguimiento de egresados, convirtiéndose así en uno de los objetivos del plan de trabajo de la FEI.

La FEI ha realizado diversas estrategias a lo largo de los años para obtener información acerca del desempeño laboral de sus egresados. Estas estrategias no han sido sistemáticas ni se ha implementado un sistema propio de seguimiento de egresados. Por ejemplo, cuando se ha requerido información acerca de los egresados, se ha recurrido a profesores que tengan contacto con el egresado, comunicándose con ellos mediante correo electrónico o teléfono para posteriormente aplicar una encuesta. Por otra parte, tampoco se ha sistematizado el procesamiento de las encuestas ni se han almacenado sus resultados de manera que puedan apoyar la toma de decisiones posteriores.

Uno de los principales problemas de un sistema de seguimiento de egresados es mantener un directorio actualizado de egresados, que permita contactar al egresado para conocer primordialmente las impresiones que ha tenido con la carrera cursada y su estado profesional actual. Ante esta situación se planteó el realizar una red social de egresados de tal manera que se pueda mantener una comunicación constante entre la FEI y los egresados y así obtener un directorio con los datos actualizados de los egresados.

En la actualidad distintas universidades hacen uso de las redes sociales con diversos propósitos. Por ejemplo:

- Universidad de Sevilla: creó la red US2.0 (<http://web20.us.es/>), es una red social exclusiva de los estudiantes y maestros de la universidad con la finalidad de planificar eventos, exponer dudas, compartir artículos y sitios web de interés.
- Universidad de Guanajuato: lanzó la red social Buzzler (<http://redsocia.ugto.mx/>), esta red es de uso académico para publicación de textos académicos y mensajería instantánea.
- Universidad Autónoma de Barcelona: crea la red social llamada Patatabrava (<http://www.patatabrava.com/>) de contenido universitario, libros, tienda en línea, oferta laboral e inclusive apuntes.
- Universidad de Vigo: tiene un red social bajo el mismo nombre pero de carácter público, permite al usuario crear blogs, foros, subir videos e interacción con otras redes sociales (<http://redsocia.uvigo.es/?code=4f18e2e7e7adc>).

3 Red social de seguimiento de egresados.

El objetivo de utilizar una herramienta CMS de tipo *Open Source* es debido a que permite acortar el tiempo de desarrollo, además el tipo de licencia permite co-

nocer el código y desarrollar aplicaciones como los plugins para añadir nuevas funcionalidades o bien cambiar o adecuar la interfaz.

Una de las primeras actividades en el desarrollo de la red social de seguimiento de egresados fue realizar un análisis comparativo entre gestores de contenido de software libre, con el fin de seleccionar el CMS más adecuado para la implementación de la red. Se llegó a la conclusión que la mejor herramienta para desarrollar este tipo de tarea era el gestor de contenidos Elgg [5]. Elgg fue creado por Ben Wedmuller y David Tosh, lleva por nombre el de una ciudad en Suiza que significa “alce”. Elgg crece y mejora gracias al desarrollo comunitario de programadores en todo el mundo. En agosto de 2008 recibió el premio como la mejor plataforma social de trabajo [6].

Elgg fue desarrollado bajo el patrón de diseño Modelo Vista Controlador (MVC) [7]. El modelo es un conjunto de clases que define la forma en la que funciona la plataforma y almacena en la base de datos. La vista es aquello que se muestra en la interfaz del usuario. El controlador es un intermedio entre la vista y el modelo, pues recibe las órdenes del usuario a través de la vista y recibe la respuesta del modelo para enviarla a la vista.

Una vez que se eligió la herramienta a emplear, fue necesario adaptarla de acuerdo a las necesidades que se tenían. Una de estas necesidades es que tuviera funciones similares a las redes sociales más conocidas y utilizadas por los usuarios Facebook y Twitter, por ejemplo: publicación fotos, publicar un cambio de estado, enviar mensajes privados, compartir archivos, contar con un perfil para cada usuario, creación de grupos privados para dialogar sobre un tema de interés, además de que pudiera compartir contenidos con otras redes sociales.

Los plugins son pequeñas aplicaciones que se agregan o se eliminan hasta obtener todas las funcionalidades deseadas en Elgg [8]. Debido a que Elgg se desarrolló bajo el patrón de diseño MVC, permite hacer cambios principalmente mediante la aplicación de plugins en su mayoría sin necesidad de modificar el núcleo de Elgg. Es posible descargar e instalar plugins de la página oficial y probar su funcionamiento o bien crear plugins propios en caso de no encontrarse alguno que cumpla con la función deseada. Para el desarrollo de plugins en Elgg se tienen las siguientes ventajas:

- Contar con documentación en línea.
- Verificar como están desarrollados otros plugins.
- Disponer de una comunidad de desarrolladores de plugins los cuales ayudan a resolver problemas y dudas.
- La reutilización de códigos de otros plugins generando nuevos sin dejar de considerar la idea de origen de cada autor.

De acuerdo con las necesidades que tenía la FEI se comenzó a implementar los plugins necesarios para la captación de la atención e información del egresado, por lo que se utilizaron los siguientes plugins:

- *Profile* permite al usuario tener una página de perfil con sus datos.
- *The Wire* o el muro, el cual provee el servicio de *microblogging* en la red.
- *Messages* elegido para permitir al egresado enviar mensajes privados.

- *Groups* para la generación de grupos de usuarios, esta funcionalidad fue elegida para proporcionar espacios donde se comenten temas de interés entre egresados y maestros.
- *Friends* fue elegido con la finalidad de que los egresados puedan relacionarse entre sí.
- *Tydropics*, es un plugin que no viene dentro de la instalación básica de Elgg pero fue seleccionado a partir de la necesidad de poder compartir fotografías.
- *profile_manager* permite modificar el registro del usuario y el perfil adhoc a las necesidades de la FEI.
- *Twitter* permite al egresado mostrar sus publicaciones de Twitter en su perfil de la red.
- *Twitterservice* es un plugin que permite al egresado enlazar sus publicaciones en el muro con su cuenta de Twitter.
- *Uservalidationbyemail* envía al correo de los egresados una solicitud de validación una vez registrado.
- *Search* éste plugin permite la búsqueda dentro de la red como usuarios, publicaciones y etiquetas.

4 Plugin conector con redes sociales.

Una API (del inglés *Application Programming Interface*) es un conjunto de métodos y clases que sirven como una biblioteca o como se le conoce comúnmente como librería. Facebook y Twitter cuentan con esta API abierta con la que cualquiera que tenga una cuenta verificada le es posible acceder a ella [9].

Previamente al desarrollo del plugin se probó con plugins existentes que permitieran incorporar a Elgg la funcionalidad de registrar al usuario con sus cuentas de Facebook y Twitter. Sin embargo se encontró que la API de Facebook se actualizó por razones de seguridad por lo que los métodos y clases ocupadas en el desarrollo de esos plugins eran inservibles. En otro caso, se debía incorporar un plugin para conectar con Facebook y en el tercero para conectar con Twitter, el inconveniente al instalar ambos Elgg falla.

Es por ello que se procedió a diseñar un plugin llamado Fbconnect que permita al usuario realizar su registro con los datos que se encuentran en Facebook o Twitter y editar su perfil en la red social creada con Elgg con los datos almacenados en Facebook. Para desarrollarlo se utilizó la Graph API de Facebook [10] y Twitter [11], además el plugin desarrollado debía ser compatible con el MVC de Elgg para que pudiera ser funcional.

Con el uso del plugin pueden obtenerse de la cuenta del egresado: su correo electrónico, nombre completo, en caso de encontrarse en su perfil el lugar de ubicación, lugares en los que ha estudiado y lugar donde labora. Debido a que el usuario puede cambiar sus datos de perfil se pensó en dividirlo en dos casos de uso que se describen a continuación y pueden verse en la Figura 4.1 abajo:

- Registrarse con Facebook y Twitter: el usuario utiliza su cuenta y datos de Facebook o Twitter para registrarse en la red social de egresados.

- Llenar perfil con Facebook: para este caso de uso es necesario que el usuario se encuentre registrado en la red social de egresados, sin embargo es indistinto si el usuario se registró por medio del caso de uso anterior o no. En este caso de uso se obtendrán los datos del perfil de Facebook.

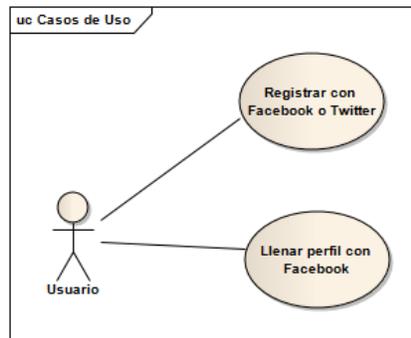


Figura 4.1. Casos de uso de Fbconnect.

Además de seguir la estructura estándar de desarrollo propuesta por Elgg, lo que se conoce como plugin *skeleton* [12], es decir que los plugins deben contar con ciertas carpetas y archivos para que Elgg pueda reconocerlos. La estructura del plugin puede verse en el diagrama de despliegue en la Figura 4.2 en la siguiente página.

Para el desarrollo del plugin es necesario desarrollar el plugin bajo el MVC de Elgg, para este caso específico funciona como se describe en las Secciones 4.1 a 4.3 a continuación.

4.1 Modelo Fbconnect.

Para el caso del modelo se utilizaron como base las clases y métodos de la API de Facebook y también la API de Twitter. Para utilizar dichos métodos y clases es necesario descargar los archivos de la API que pueden descargarse en las páginas oficiales para desarrolladores. Para el caso de Facebook la página oficial *Facebook Developers* (<http://developers.facebook.com/>) y en el caso de Twitter en la página *Twitter Developers* (<https://dev.twitter.com/>). También se requiere de registrar la aplicación que se va a desarrollar o que utilizará la API, en la misma página oficial. Al finalizar el registro de la aplicación se otorgan dos llaves que permiten hacer comunicarse con Facebook o Twitter.

4.2 Vista Fbconnect.

Al desarrollar la vista del plugin se emplearon los métodos proporcionados en la documentación de Elgg [13]. Para este caso específico se utilizó la vista extendida (*extend_view*), dicha función recibe como parámetros el espacio donde se espera agregar la nueva vista y la ruta en el plugin donde se encuentra la vista que se desea

aumentar, por lo regular esta función se inicializa en el archivo start.php del plugin. Para Fbconnect se crearon 3 vistas: una que permite al usuario ingresar a las redes sociales correspondientes la cual se nombró login.php, una segunda encargada de mostrar los datos en el formulario de registro llamada register.php y una tercera que tiene una función similar a la anterior sólo que permite mostrar los datos en la edición de perfil del usuario que tiene el nombre de edit.php. Para que Elgg pueda leer correctamente estas vistas los archivos de las vistas creadas se han colocado en la ruta de vistas default del plugin así Elgg podrá llamarlas cuando se inicialice el plugin.

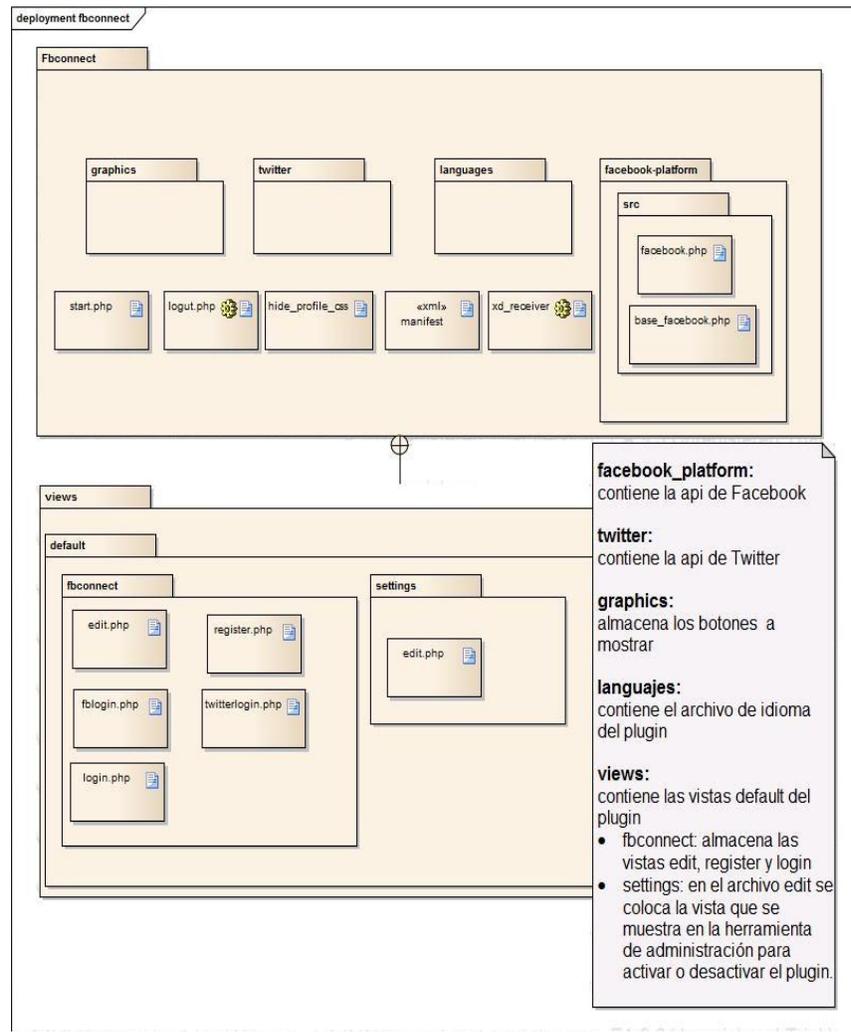


Figura 4.2. Diagrama de despliegue de Fbconnect.

4.3 Controlador Fbconnect.

El controlador [14] en Elgg es el encargado de las acciones que puede realizar el usuario, para este plugin se utilizó “*register_event_handler*” que para Elgg es la función encargada de inicializar los plugins y de realizar eventos especiales. En fbconnect se hizo uso de *init* y *pagesetup* los cuales se explican a continuación:

- Procedimientos de configuración (Init): se implementa en el caso de agregar un nuevo menú o submenú y para el inicio de un plugin. Primero se envía el registro de acción *init* con la función “fbconnect_init” la cual se encarga de inicializar el plugin, de la siguiente manera:
`register_elgg_event_handler('init', 'system', 'fbconnect_init');`
- Configuración de página (Pagesetup): generalmente se emplea como apoyo para la vista. Después envía el registro de acción *pagesetup* con la función “fbconnect_forms” la cual extiende la vista *register.php* o *edit.php* una vez que el usuario ingresó en su cuenta de Facebook, resultando entonces la siguiente función:
`register_elgg_event_handler ('pagesetup','system','fbconnect_forms');`

5 Pruebas.

Una vez terminada la red social se decidió poner a prueba la red en primera instancia para determinar la opinión de los egresados. La red social se subió a un servidor web que contaba con los recursos necesarios (PHP, MySQL). La red social de egresados fue puesta a prueba de el 27 de Diciembre de 2011 hasta el 22 de Enero de 2012. Para lograr que los egresados probaran la red social se enviaron invitaciones personalizadas por correo a aquellos egresados que anteriormente habían confirmado su cuenta de correo. En total se inscribieron a la red 63 egresados quienes realizaron actividades distintas como actualizar su perfil, leer un boletín de noticias de la FEI, subir fotos, crear grupos y enviar mensajes.

La primera prueba fue de usabilidad para conocer si el usuario podía utilizar la red social correctamente, por lo que se tomó una muestra de 10 usuarios y se les pidió que realizaran 10 actividades y que posteriormente contestaran un cuestionario para conocer sus impresiones. Como resultado se obtuvieron algunas observaciones como problemas para subir varias fotos y que el tamaño de la barra principal no les fue agradable. Con respecto a las sugerencias opinaron que les agradaría un cambio de colores de la interfaz, que hubiera actividades como juegos, un chat y concursos. Además el 91% de los egresados que respondió la encuesta dijo que si la recomendaría a otros egresados.

Para el caso del uso del plugin Fbconnect los egresados indicaron que les gustaría que el botón estuviera en español y detectaron problemas para cargar datos en perfil si tenían su cuenta de Facebook abierta. Gracias a que Facebook lleva una bitácora del uso de cada aplicación, fue posible conocer la cantidad de usuarios que usó el plugin Fbconnect. El número máximo de veces que fue utilizado en un día fueron 6 veces. Los resultados pueden verse en la Figura 5.1 en la siguiente página.



Figura 5.1. Estadística de uso de Fbconnect.

Esta primera etapa de pruebas permitió conocer la opinión de los egresados. Gracias a ella se pudieron determinar mejoras para una siguiente versión. Entre estas mejoras están: 1) mejorar la presentación de la interfaz para hacerla más atractiva pero al mismo tiempo más institucional, 2) incluir actividades para los miembros de la red, tales como juegos y concursos, e 3) incluir un chat. Pruebas de seguridad y rendimiento no se realizaron de manera formal durante esta primera etapa pues el objetivo primordial en ésta era saber si la red era del agrado de los egresados. Sin haberse planeado, se contó con la participación de estudiantes de ambos programas educativos de la FEI en esta primera etapa de pruebas y su retroalimentación ha sido tomada en cuenta. De esta manera, se asegura la continuidad en el uso de la red ya que estos estudiantes, al egresar, seguirán siendo miembros de la comunidad de la red social.

En lo que respecta al seguimiento de egresados y al estudio de su desempeño laboral, no se tienen pruebas ni resultados aún. Esta primera etapa de desarrollo lo que busca solucionar es el problema del directorio actualizado. Mediante la participación de esta primera muestra de la población de los egresados se puede concluir que las expectativas son buenas y el contar ya con la participación de estudiantes hace posible plantearse mejores expectativas para el corto plazo.

Actualmente se trabaja en la siguiente versión que incluye las mejoras ya mencionadas. Al mismo tiempo, se trabaja en la encuesta a egresados y su automatización para lo cual ya se ha diseñado el instrumento de la encuesta conforme a la metodología de la ANUIES.

6 Conclusiones.

El seguimiento de egresados permite el análisis de la situación laboral actual del egresado, o si éste ha decidido continuar con estudios de formación o capacitación. Esta información hace posible evaluar el plan de estudios y realizar las actualizaciones correspondientes.

Las redes sociales actualmente son muy populares como un medio de comunicación y es por ello que se planteó como una solución para el directorio de los egresa-

dos actualizado del el Sistema de Seguimiento de Egresados de la Facultad de Estadística e Informática.

El desarrollo de la red se hizo empleando un gestor de contenidos (CMS) de software libre llamado Elgg. La Elgg funciona mediante plugins que pueden ser desarrollados o descargados en su página oficial.

Elgg fue adaptado mediante la instalación de los plugins. Además, fue necesario desarrollar un plugin que conectara con redes sociales comúnmente utilizadas (Facebook y Twitter). Posteriormente se realizó una primera etapa de pruebas a la red social de egresados, durante la cual se obtuvieron opiniones y mejoras. Finalmente podemos concluir que:

1. La red social de seguimiento de egresados puede dar solución a los problemas como establecer contacto con el egresado y su falta de tiempo del egresado para contestar encuestas.
2. La red social permite fomentar los lazos de comunicación entre los egresados y la FEI, favoreciendo el intercambio de experiencias y conocimientos.
3. Por lo anterior, se puede obtener información del desempeño de los egresados, para la mejora de los programas educativos y planes de estudio.
4. Implementar un medio de interacción con las redes sociales Facebook y Twitter permite utilizar a las redes sociales como Facebook y Twitter como una especie de banco de datos de datos, además de servir como un medio de difusión y mercadotecnia si se emplean las técnicas correctas.

En lo que respecta a trabajos futuros, ya se ha mencionado que actualmente se trabaja en una nueva versión que incluye las mejoras ya mencionadas. Además, se realizarán las pruebas formales de seguridad y rendimiento. En este momento ya se ha liberado una versión beta a la misma comunidad que realizó la primera etapa de pruebas y los comentarios hasta el momento han sido muy positivos. Incluso se han agregado nuevos miembros a la red, tanto egresados como estudiantes. Como ya se dijo, se trabaja en la encuesta a egresados y su automatización que contempla el instrumento de la encuesta conforme a la metodología de la ANUIES.

Referencias

1. Barrón, (2003) "Los estudios de seguimiento de egresados en el período 1992-2002", en Educación, trabajo, ciencia y tecnología, Colección la Investigación Educativa en México 1992-2002, libro 6, México: COMIE
2. ANUIES. (1998). Esquema Básico para Estudios de Egresados en Educación Superior. Dirección de Servicios Editoriales. México. Recuperado Marzo 5, 2011. Disponible en: http://www.anuies.mx/servicios/p_anuies/publicaciones/libros/lib10/0.htm
3. AMIPICI, (2011, Mayo). Hábitos de los usuarios de Internet en México. Recuperado Julio 15, 2011. Disponible en: <http://www.e-2box.com/datos/Habitos2011AMIPCIPrensacomprimida-0010959001305646317OB.pdf>
4. Fresán, (1998). Esquema Básico para Estudios de Egresados en Educación Superior. Dirección de Servicios Editoriales. México. Recuperado Marzo 5, 2011. De: http://www.anuies.mx/servicios/p_anuies/publicaciones/libros/lib10/000.htm

5. Elgg docs (2011, Octubre) ¿Qué es Elgg? Recuperado Octubre 3, 2011. Disponible en: http://docs.elgg.org/wiki/Cos%C3%A8_Elgg
6. InfoWorld (2008, Agosto) InfoWorld announces our 2008 Best of Open Source Awards. Recuperado Octubre 31, 2011. Disponible en: <http://www.infoworld.com/d/open-source/infoworld-announces-our-2008-best-open-source-awards-065>
7. Elgg MVC (2011, Octubre) Núcleo Elgg. Recuperado Octubre 3, 2011. Disponible en: <http://docs.elgg.org/wiki/Engine>
8. Elgg plugins (2011, Octubre) ¿Qué es plugin? Recuperado Octubre 3, 2011. Disponible en: <http://community.elgg.org/plugins>
9. API (2012, Enero) Definición de API. Recuperado Enero 4, 2012. Disponible en: http://es.wikipedia.org/wiki/API_de_Windows
10. Facebook *Developers Documentation*. (2011, Marzo). Recuperado Marzo 5, 2011. Disponible en: <http://developers.facebook.com/>
11. Twitter API Wiki, (2011, Marzo). *API Documentation*. [Wiki]. Recuperado Marzo 6, 2011, Disponible en: <http://apiwiki.twitter.com/w/page/22554648/FrontPage>
12. Elgg Plugin Skeleton (2011, Abril). Estructura de plugin para Elgg. Recuperado Abril, 4, 2011. Disponible en: http://docs.elgg.org/wiki/Plugin_skeleton
13. Vista Elgg, (2011, Abril). Elgg *Documentation*. Recuperado Abril, 4, 2011. Disponible en: <http://docs.elgg.org/wiki/Engine/Views>
14. Eventos Elgg (2011, Abril). Elgg *Event*. Recuperado Abril, 4, 2011. Disponible en: http://docs.elgg.org/wiki/Elgg_Events

Sistema Web de Escenarios de Aprendizaje Virtuales para las Matemáticas de Primer Grado de Nivel Secundaria

Gaspar García-González¹, Carmen Cerón-Garnica² y Etelvina Archundia-Sierra³

^{1,2,3}Benemérita Universidad Autónoma de Puebla
Puebla, Pue., México

¹gaspar.garcia@gmail.com, ²mceron@cs.buap.mx, ³etelvina@cs.buap.mx
Paper received on 25/07/12, Accepted on 05/09/12.

Resumen. El propósito de este trabajo es presentar el impacto de las Tecnologías de Información y Comunicación en la educación, mediante el diseño y desarrollo de un sistema web como plataforma de contenidos educativos y la utilización de objetos de aprendizaje como escenarios de aprendizaje para la materia de matemáticas de primer año de educación secundaria. Los objetos de aprendizaje incluyen multimedia y realidad aumentada (RA) para facilitar actividades formativas para el aprendizaje personalizado y colaborativo bajo el enfoque pedagógico constructivista y de competencias en las matemáticas. El sistema web se modeló bajo UML y se desarrolló con HTML5, PHP, MYSQL y Flash en una arquitectura cliente-servidor de tres capas. Por último se presenta los resultados del nivel de uso de satisfacción del sistema al pilotarse con un grupo de alumnos de este nivel educativo.

Palabras Clave: Sistema Web, TIC's, Sistema Gestor de Contenidos, Objetos de Aprendizaje y Realidad Aumentada.

1 Introducción.

Actualmente el uso de las Tecnologías de Información y Comunicación en el campo de la educación son usadas como herramientas para apoyar procesos educativos centrados en el alumno. La utilización de la Web como medio de comunicación e interacción es utilizada en distintos modelos de educación basada en Web (*Web-Based Education*) [2] como son el e-learning¹ y b-learning² siendo una gran ventaja

¹ **e-learning:** Aprendizaje asistido por las tecnologías de la información y comunicación (TIC's). Facilita la creación, adopción y distribución de contenidos, la adaptación del ritmo de aprendizaje, la disponibilidad de las herramientas de aprendizaje independientemente de límites horarios o geográficos y el intercambio de opiniones y aportes a través de la plataforma.

para que los estudiantes tengan acceso a una gran cantidad de contenidos, información y actividades de aprendizaje de acuerdo a sus necesidades e intereses personales.

El modelo de aprendizaje semi-presencial, comúnmente conocido como *b-learning*² hace uso de plataformas tales como Sistemas de Gestor de Contenidos *LCMS*³ que facilitan el aprendizaje distribuido y colaborativo mediante las cuales es posible la realización y administración de nuevos escenarios de aprendizaje (e-actividades formativas) individuales y colaborativos, además permiten registrar usuarios y almacenar sus datos, organizar contenidos, y suelen incluir también herramientas de comunicación al servicio de los usuarios, tales como blogs, wikis y chats.

El sistema web se diseñó como un LCMS que proporciona distintas opciones para los procesos administrativos para configurar cursos, matricular alumnos, registrar profesores, asignar cursos a un alumno, llevar estadísticas y evaluaciones para dar seguimiento al desarrollo de competencias de la asignatura. Este sistema se desarrolló en una arquitectura cliente-servidor y se utilizó HTML5, PHP y JavaScript. Los contenidos educativos del sistema Web se diseñaron mediante *objetos de aprendizaje (OA) o Learning Objects (LO) estandarizados (SCORM y LOM)* representando distintos escenarios de aprendizaje los cuales integraron multimedia y realidad aumentada para visualizar y enriquecer los contenidos utilizando códigos QR y Flash. Estos OA permiten a los estudiantes experimentar y desarrollar las competencias necesarias en la resolución de problemas.

El alcance del proyecto fue dirigido a los alumnos de primer año de educación secundaria, y pretende facilitar mayor interacción entre profesores y alumnos mediante las herramientas de comunicación sincrónica y asincrónica como el chat, foro, blog y wiki que proporciona el sistema.

2 Desarrollo.

2.1 Contexto Educativo.

El *Sistema Web b-learning para Matemáticas de Educación Secundaria* es un LCMS utilizado específicamente para la enseñanza de matemáticas correspondientes al primer nivel de educación secundaria. Los contenidos de aprendizaje para este sistema web se presentaron bajo un enfoque pedagógico constructivista y de competencias que permitieron el desarrollo de los saberes: el saber que, saber ser, saber hacer y el saber convivir permitiendo un aprendizaje personalizado y colaborativo

² ***blended learning***: consiste en un proceso docente semipresencial; esto significa que un curso dictado en este formato incluirá tanto clases presenciales como actividades de e-learning.

³ ***Learning Content Management System***: Su objetivo es gestionar el contenido u objetos de aprendizaje. La mayor parte de los sistemas LCMS incorporan funcionalidades de LMS. Permiten crear, almacenar, reutilizar, gestionar y distribuir contenidos de aprendizaje a partir de un repositorio central de objetos de aprendizaje.

para toda la vida. La ONU define las competencias “como el conjunto de comportamientos socioafectivos y habilidades cognoscitivas, psicológicas, sensoriales y motoras que permiten llevar a cabo adecuadamente un desempeño, una función, una actividad o tarea” [3]. La teoría del aprendizaje constructivista es una de las principales teorías a desarrollar e implantar en los entornos de enseñanza-aprendizaje basados en los modelos b-learning, estos modelos se centran en la combinación de estrategias pedagógicas, propias y específicas, de los modelos presenciales y estrategias de los modelos formativos sustentados en las tecnologías Web donde el sistema representa distintos escenarios de actividades como son juegos, exploraciones y evaluaciones [10].

Los sistemas educativos b-learning son herramientas eficaces para una enseñanza activa, basada en el descubrimiento, la interacción y la experimentación. Su aporte principal reside en su contribución a la realización de una pedagogía activa. El estudio de las matemáticas busca que los niños y jóvenes desarrollen una forma de pensamiento que les permita expresar matemáticamente situaciones que se presentan en diversos entornos socioculturales, así como utilizar técnicas adecuadas para reconocer, plantear y resolver problemas; al mismo tiempo, se busca que asuman una actitud positiva hacia el estudio de esta disciplina y de colaboración y crítica. La actitud positiva hacia las matemáticas consiste en despertar y desarrollar en los alumnos la curiosidad y el interés por investigar y resolver problemas, la creatividad para formular conjeturas, la flexibilidad para modificar su propio punto de vista y la autonomía intelectual para enfrentarse a situaciones desconocidas; asimismo, consiste en asumir una postura de confianza en su capacidad de aprender a aprender por tal motivo se buscó la estrategia de escenarios de aprendizaje.

Un escenario de aprendizaje es aquel que permite al estudiante experimentar un conjunto de actividades, disponiendo de diversos recursos y estrategias para interactuar con los contenidos de una unidad de aprendizaje [7].

Los contenidos matemáticos que se estudian en la educación secundaria se han organizado en tres ejes: *Sentido numérico y pensamiento algebraico*; *Forma, espacio y medida* y *Manejo de la información*. Los ejes están organizados en cinco bloques, en cada uno hay temas y subtemas de los tres ejes descritos y que organizamos en unidades de aprendizaje y en talleres de aprendizaje. Esta organización tiene dos propósitos fundamentales; por una parte, se trata de que los profesores y sus alumnos puedan establecer metas parciales a lo largo del año escolar y, por la otra, se pretende garantizar el estudio simultáneo de los tres ejes durante el curso. Los contenidos que se han organizado en secuencias didácticas lo cual significa que permite la construcción de conocimientos significados y de la aplicación de herramientas matemáticas por parte de los alumnos, con base en la resolución de problemas, de tal manera que los alumnos vayan teniendo acceso gradualmente a contenidos cada vez más complejos y a la vez puedan establecer conexiones entre lo que ya saben y lo que están por aprender.

2.2 Modelado y Funciones del Sistema.

El sistema web se modeló mediante UML [1] y se implementó en una arquitectura cliente-servidor de tres capas [4] facilitando el desarrollo y permitiendo diferen-

tes vistas para los usuarios del sistema: Administrador, Profesor y Alumno, los cuales tiene distintas actividades como se visualiza en los casos de uso del sistema en la figura 1.

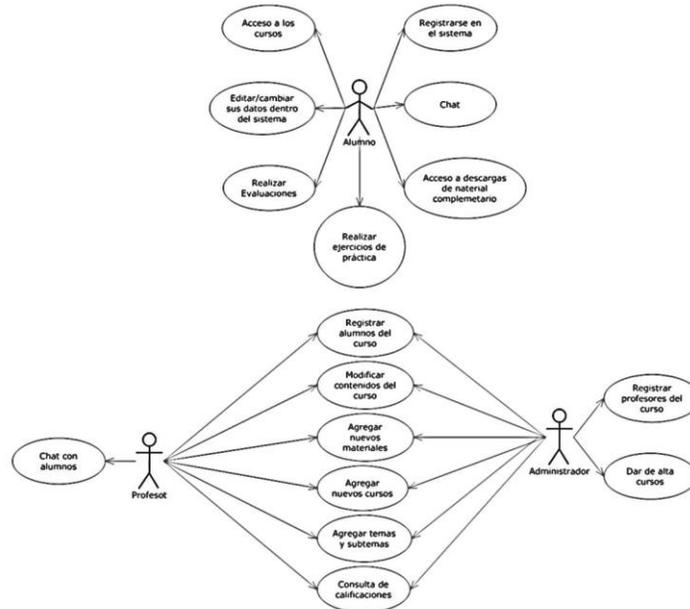


Figura 1 Diagrama UML que describe casos del Profesor, Alumno y Administrador.

2.2.1 Descripción y Casos de Uso del Sistema.

Administrador: Es la persona encargada de la elaboración, actualización y mantenimiento del sitio web.

- Registro de alumnos y de profesores.
- Agregar, consultar, modificar y eliminar materiales del curso y objetos de aprendizaje.
- Dar de alta y agregar nuevos cursos.
- Agregar, consultar y eliminar unidades de aprendizaje y talleres.
- Consulta de calificaciones/evaluaciones de alumnos.

Profesor: Es la persona encargada de guiar y llevar un control sobre el aprendizaje de los alumnos:

- Registrarse como Profesor, Registrar alumnos.
- Modificar los contenidos del curso.
- Agregar nuevos cursos, unidades de aprendizaje y talleres.
- Agregar nuevos materiales, objetos de aprendizaje y reutilizarlos.
- Consulta de los resultados de las evaluaciones de los alumnos.
- Interactuar mediante herramientas asincrónicas y sincrónicas: blog, chat y wiki.

JavaScript, se implementó una pizarra colaborativa que se ejecuta del lado del servidor, y permite dinámicamente la interacción de múltiples usuarios en un entorno colaborativo mediante el uso de websockets [4,5].

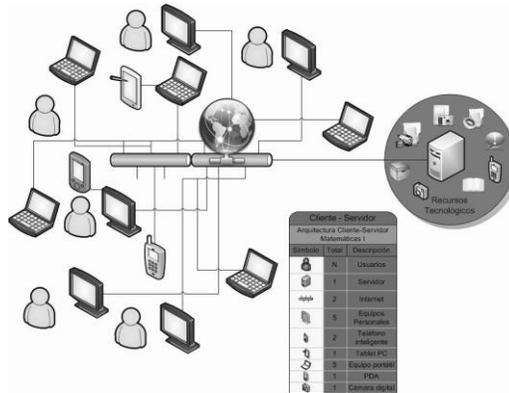


Figura 3 Diseño de Arquitectura Cliente – Servidor del Sitio Web.

2.5 Diseño de Contenidos mediante Objetos de Aprendizaje.

Para la construcción de los contenidos educativos del sistema se basó en el desarrollo de Objetos de Aprendizaje (OA). Según la IEEE los OA son definidos como “una entidad, digital o no digital, que puede ser utilizada, reutilizada y referenciada durante el aprendizaje apoyado con tecnología” [8]. La aplicación de los OA contribuyó a representar distintos escenarios de aprendizaje para cada unidad de aprendizaje y formar una secuencia didáctica. Los objetos fueron estandarizados bajo el formato de IMS (Instructional Management System) estándar para especificaciones de los contenidos y la especificación de LOM (Learning Objects Metadata) se utilizó para describir al objeto en su ámbito de aplicación [8]. Los OA cumplen con la especificación de meta-data lo cual permite identificar y clasificar los objetos de aprendizaje para generar el escenario de aprendizaje requerido por el usuario. Los objetos de aprendizaje tienen una plantilla para generar el escenario que incluyen: Conocimiento/Contenido, Actividad y Evaluación. Para su desarrollo de los OA con multimedia se utilizó el software de la Adobe Master Collection CS5, entre los cuales se incluyen Adobe Photoshop, Flash Profesional, Adobe Illustrator, Adobe Dreamweaver, InDesign, Adobe Flex, así como también un gran número de librerías *js* y de *actionsript* 2.0 y 3.0 para integrar animaciones, videos, audio, juegos y actividades de aprendizaje[9].

2.5.1 Diseño de Objetos de Aprendizaje con Realidad Aumentada.

La realidad aumentada [7] consiste en combinar datos virtuales – información, elementos multimedia, elementos tridimensionales, etc. – con lo que se ve en el mundo real. En este contexto, la realidad aumentada puede ayudar a los estudiantes a contextualizar su aprendizaje. La creación de contenidos altamente interactivos

basados en realidad aumentada apoya el proceso de aprendizaje, mayor interacción con los recursos, para lo cual usamos una tarjeta QR diseñada en illustrator y después con base a esa imagen se crea un patrón de marca con la ayuda de maker generator que es una aplicación AIR de Adobe, permitiendo la interacción con el objeto de aprendizaje.

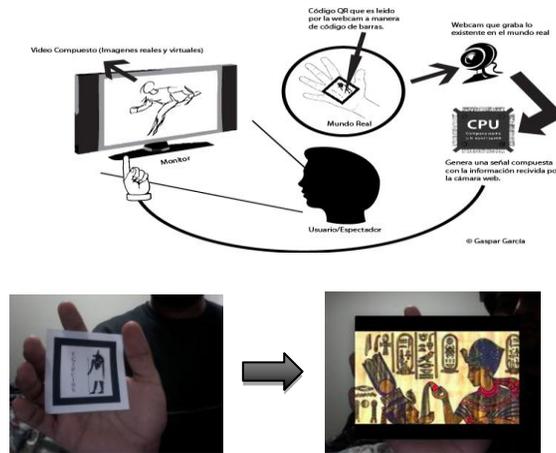


Figura 4. Esquema de Generación de la Realidad Aumentada para los objetos de aprendizaje.



Figura 5. Pantalla Principal de Navegación del Sistema.

2.6 Diseño de las Interfaces del Sistema Web.

Las interfaces del Sistema Web fueron realizadas principalmente mediante el uso software de adobe (Adobe Master Collection CS3), como animaciones, imágenes, videos, escritura de código php y textos; como se muestra en la figura 5.

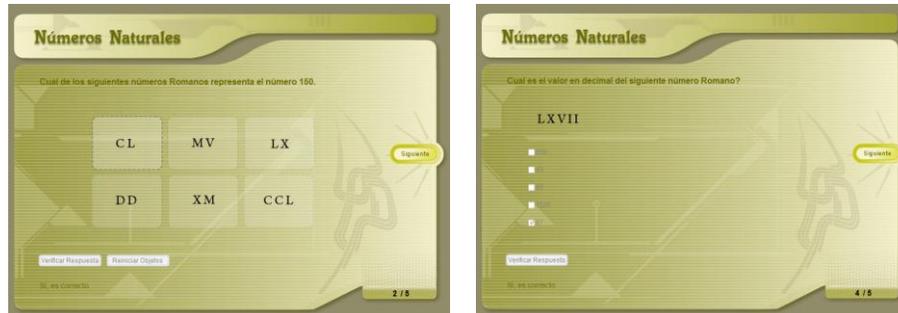


Figura 6. Presentación de Actividades.

3 Resultados.

En el presente trabajo se logró lo siguiente:

- Se desarrolló un Sistema Web como un LCMS para la gestión del curso de matemáticas de primer grado cuyos contenidos fueron elaborados de acuerdo al Plan de Estudios de Educación Secundaria 2006, siendo un software para apoyar los procesos educativos requeridos por la Reforma de educación Básica en México.
- El sistema web permite la generación de escenarios de aprendizaje para fomentar el aprendizaje colaborativo e individual mediante el uso de objetos de aprendizaje que incluyen multimedia y realidad aumentada.
- Se ofrece tanto a alumnos como a profesores una nueva alternativa y dinámica de aprendizaje, orientada siempre al desarrollo integral de sus capacidades y estrategias de estudio para construir un aprendizaje más significativo mediante un razonamiento lógico.
- El sistema fue piloteado por un grupo de alumnos de primero de secundaria de 35 alumnos donde interactuaron con el sistema de forma presencial y no presencial y evaluaron el sistema conforme al test de Usabilidad llamado UsalabCrea [11] con sus categorías respectivas como se muestra en Tabla 1.
- La categoría con mayor porcentaje fue la de Recuerdo Visual con el 94% y de acuerdo a las entrevistas argumentaron; que debido a los contenidos que presentaron multimedia y realidad aumentada permitieron mayor interactividad y uso de la información presentada en los OA.

Tabla 1. Resultados de Prueba de Usabilidad UsalabCrea.

Categorías UsalabCrea	SI	NO
Recuerdo-visual	94%	6%
Organización de Contenidos de Aprendizaje	90%	10%
Eficiencia -Funcionalidad	92%	8%

4 Conclusiones.

El principal aporte del uso del *Sistema Web B-learning para Matemáticas de Educación Secundaria* es brindar una herramienta que permita crear escenarios y ambientes de aprendizaje significativos para apoyar el proceso de enseñanza – aprendizaje. El hacer uso de las tecnologías de información y comunicación (TIC) permite mejorar los procesos educativos ofreciendo contenidos y materiales educativos de calidad. Se presentó el desarrollo de un sistema donde se diseñaron contenidos educativos basados en los objetos de aprendizaje reutilizables cumpliendo especificaciones bajo el estándar de SCORM requeridos por un LCM, que proyecta una mejor adaptación y flexibilidad en brindar ambientes de aprendizajes web personalizados de acuerdo a las necesidades educativas de los estudiantes donde se promueve el constructivismo sociocultural y el desarrollo de competencias del alumno construyendo su propio aprendizaje mediado por el docente y el uso de las TIC. Las tecnologías utilizadas en este sistema web incluyeron objetos de aprendizaje, PHP, hojas de estilo (CSS3), JavaScript y Actionscript.

En nuestro trabajo futuro, vamos a seguir investigando las diferentes posibilidades para proyectar mayor interacción a los objetos de aprendizaje y aplicar nuevas tecnologías de la web semántica a los OA, para una mejor clasificación y almacenamiento en bases de datos y facilitar la reutilización en el uso de ontologías que permitan ofrecer una estructura de objetos aprendizaje, dando un nuevo sentido pedagógico al desarrollo de objetos de aprendizaje inteligentes que puedan apoyar al usuario realizando diferentes tareas.

Una de las principales perspectivas es poder elaborar otros sistemas web para apoyar los procesos de aprendizaje en la educación basada en competencias que requiere la Reforma Educativa para contribuir a mejorar la calidad de la educación en México.

Referencias

1. Arlow J, Neustadt, UML 2, Anaya Multimedia, España (2001)
2. Berners-Lee T., Miller E., The Semantic Web lifts off, ERCIM News No. 51. Consultado [en línea] http://www.ercim.org/publication/Ercim_News/enw51/berners-lee.html marzo-2009.
3. Cabero J.: Tecnología educativa diseño y utilización de medios en la enseñanza, Paidós, España. (2003)
4. Fraude. J Ingeniería de software: una perspectiva orientada a objetos, Ra-ma, España. (2001)

5. Larman C.: UML y patrones Introducción al Análisis y Diseño Orientado a Objetos. Pearson Education, España. (2002)
6. Pavón J.: Creación de un portal con PHP y MySQL, Alfaomega, España, (2004)
7. Sherman, W.; Craig, A. Understanding Virtual Reality. Morgan Kaufmann. San Francisco, CA. (2003)
8. Pithamber R. Polsani.: Use and Abuse of Reusable Learning Objects. Journal of Digital Information, Volumen 3, número 4 Article No. 164, (2003)
9. Prendes Espinosa, María Paz. "Objetos de aprendizaje para la enseñanza". Dykinson.
10. Quesada R, Sánchez J.: Calificación y Diagnóstico del aprendizaje por Computadora. Limusa, México. (2000)
11. Archundia E: Tesis doctoral: Metodología para propiciar el aprendizaje de la programación creativa APROGC. Universidad Popular Autónoma del Estado de Puebla, México. (2009)

Tecnología móvil en la educación: Educ@-móvil

Karina Verastegui Barranco¹ y Paola Nayeli Cortez Herrera²

¹ Universidad de Londres Campus Querétaro, México
karina.verastegui@e-tecnosys.com

² Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas, IPN. México
pcortez@ipn.mx

Paper received on 03/07/12, Accepted on 11/09/12.

Resumen. Actualmente la deserción escolar es un problema que afecta el ámbito educativo en México. Para combatirlo se ha impulsado lo que se conoce como educación a distancia, en donde el alumno puede continuar sus estudios sin estar presente dentro del aula. Ante el incremento en la cantidad de tecnología móvil que las personas poseen y la cobertura de sitios con Internet gratuito, surge la idea de proponer al teléfono celular como un canal de comunicación ante el cual un alumno pueda continuar sus estudios. Este trabajo propone un sistema Web que pueda ser accedido por el alumno a través de un teléfono celular, apoyando con ello que el alumno continúe sus estudios en cualquier lugar. El principal objetivo es utilizar la tecnología móvil con un fin específico y orientado a cuestiones educativas.

Palabras Clave: educación a distancia, tecnología móvil, teléfono celular

1 Introducción

Hoy en día es común encontrar que los alumnos dejen de asistir a clases debido a diversos factores económicos, sociales, e inclusive personales. Esto va ocasionando que el alumno pierda la secuencia del temario y por lo tanto no comprenda los nuevos temas, cuando se vuelve a reincorporar a la escuela. Por lo tanto esto conlleva a que el estudiante vaya descuidando sus actividades académicas y en un futuro abandone la escuela.

En México la deserción en el nivel medio superior llega casi a un 16% [9]. Y esto es alarmante debido a como menciona Cabral [2], la educación no sólo es un indicador del nivel de desarrollo socioeconómico, sino una de las fuentes del proceso de desarrollo a nivel mundial y de elevación de la calidad de vida de la sociedad.

Ante esta problemática es necesario proponer metodologías que coadyuven a que el alumno continúe sus estudios y el hecho de no estar presente en el aula de clases no sea impedimento para su formación educativa. Para ello ha surgido una modalidad conocida como educación a distancia.

Este trabajo tiene como finalidad presentar una propuesta que permita hacer uso de la tecnología móvil (celular) para apoyar a la educación a distancia, a fin de coadyuvar a que los jóvenes no abandonen sus estudios.

A continuación se menciona la estructura con la cual ha sido organizado este artículo. En el marco de referencia se encuentran algunos trabajos relacionados con los conceptos de educación a distancia y tecnología móvil, posteriormente se hace el análisis del caso de estudio y se presenta la metodología y resultados obtenidos, finalmente se tienen las conclusiones.

2 Marco de Referencia

En esta sección se detallan brevemente los conceptos considerados en este trabajo, además de mencionar algunos trabajos relacionados con la propuesta planteada.

2.1 Educación a distancia

La educación a distancia es considerada por algunos autores como un método, otros mencionan que se trata de un proceso y algunos comentan que se trata de diseñar material de alta calidad para ser compartido a través de diversos medios [2].

Hilary Perraton menciona que es un proceso educativo en el cual una proporción significativa de aprendizaje se lleva a cabo por alguien no presente en el espacio y/o tiempo por parte del alumno [10]. Tomando como base esta última definición es posible proponer canales de comunicación por los cuales se pueda establecer un vínculo entre el alumno y sus estudios.

2.2 Tecnología Móvil

Hoy en día el teléfono móvil, permite la comunicación de señales de voz digitalizadas, se sirve de Internet para conectarse con cualquier red pública o privada, tiene capacidad para almacenar toda la información personal que se necesite, interactúa con el entorno, ayuda a conocer las coordenadas de la posición que se ocupa en un momento dado en el planeta, se puede emplear como radio y cámara fotográfica y permite jugar, trabajar, negociar, comprar, vender, informar y estar informado. Es de gran utilidad para implantar soluciones de movilidad en las empresas y para que el ciudadano disfrute de una comunicación inteligente [1].

De acuerdo a [5], la telefonía móvil, a la que ya accede la mayor parte de la población de los países desarrollados y en vías de desarrollo, va a convertirse en un vehículo imprescindible para la formación de los individuos que se mueven constantemente de un lugar a otro, conectados a través de su teléfono móvil. Las tecnologías WAP, GPRS y UMTS van a inundar todos los espacios y la educación no quedará al margen.

Esto lleva a considerar el uso de teléfonos móviles en el campo de la educación, principalmente en la educación a distancia.

2.3 Investigaciones relacionadas

En el ámbito de la educación existen diversas investigaciones las cuales hacen

uso de las Tecnologías de Información (TIC's) para permitir el intercambio de información, propiciando que el proceso de aprendizaje continúe fuera del aula. En [6] se desarrolla un trabajo en el cual se utiliza la red social Facebook para permitir el intercambio de información y establecer comunicación con alumnos a nivel universitario. Los autores mencionan que esta red social al ser popular entre personas con un rango de edad de 18 a 25 años permite que el alumno se familiarice más rápidamente con la plataforma Web. Sin embargo esta plataforma no está orientada a cuestiones educativas y eso al final puede influir para que los alumnos terminen distrayéndose al utilizarla.

Al momento de encontrar investigaciones que utilicen tecnología móvil en el proceso enseñanza-aprendizaje se encuentra un análisis realizado por Sharples [11] en donde se menciona que una computadora no es capaz de conocer si existen malos entendidos entre los estudiantes, y que los sistemas inteligentes de tutorías existentes aún se encuentran enfocados de forma individual, y no son capaces de atender necesidades colectivas tal como ocurre en un salón de clases. En ese trabajo se menciona que un celular proporciona un nuevo canal de comunicación por el cual los alumnos pueden estar en contacto. Al respecto se mencionan las características que tiene el usar este tipo de dispositivos en la educación:

1. Altamente portátil, de modo que está disponible dondequiera que el usuario necesita aprender.
2. Útil, adecuado para las necesidades diarias de comunicación, de referencia, el trabajo y el aprendizaje.
3. Intuitivo de usar por personas sin experiencia previa de la tecnología.

Sin embargo el utilizar las TIC's no es sinónimo de calidad e innovación en la enseñanza como lo menciona Díaz [4], puesto que se necesita diseñar nuevas teorías del diseño educativo a fin de asegurar que se puede lograr un aprendizaje significativo en el alumno. La autora enuncia la importancia de un perfil innovador por parte del docente, pues éste necesita tener conocimiento de las nuevas tecnologías y de la compatibilidad existente entre la tecnología y pedagogía. Además resalta el rol que ocupa el docente en la educación a distancia, no basta con colocar material en un sitio Web se debe de saber cómo utilizar ese material y sobretodo que sea de utilidad.

En [8] se menciona que las personas pueden estudiar en cualquier tiempo y lugar, con un sistema de aprendizaje a través de dispositivos móviles. El hecho de emplear tecnología móvil en el proceso enseñanza-aprendizaje da por resultado dos esquemas de aprendizaje: sincrónico y asincrónico. Comunmente al usar dispositivos móviles se emplea un esquema de aprendizaje asincrónico, pues no necesariamente el profesor y el alumno se encuentran conectados al mismo tiempo. Los autores a través de un teléfono celular hacen uso de una aplicación didáctica que permite realizar simulaciones.

En el ámbito comercial también se encuentran plataformas de índole educativo que hacen uso de la tecnología móvil tal es el caso del sistema denominado Cúrsame [3] el cual permite establecer cursos en línea. La plataforma funciona como un aula virtual en donde el alumno puede consultar las presentaciones o videos proporcionados por el profesor y este último puede calificar las tareas y observar las participaciones de los alumnos a través de foros. La ventaja que ofrece este sistema es que

posee un denominado “aprendizaje social” en el cual a través de un corto periodo de tiempo el sistema va recomendando el mejor material a los profesores para mejorar sus clases, y a los alumnos para obtener un mejor aprendizaje. Dependiendo de las mejoras que tenga el sistema va desde un costo gratuito hasta una cotización personalizada a instituciones educativas o empresas. En [7] se menciona que el aprendizaje móvil integra recursos digitales con auténticos contextos de aprendizaje, lo que da por resultado que se produzca un aprendizaje más efectivo.

Otra aplicación de los teléfonos móviles en la educación se presenta en el proyecto EducMI Móvil [13], el cual ofrece entrar a los cursos donde esté inscrito el alumno, leer mensajes pendientes, ver a los usuarios que se encuentran en línea y enviar mensajes a sus compañeros. Este proyecto se encuentra en desarrollo en la Universidad de Colima.

Cabe mencionar que el uso de las TIC’s no solo se aplica a educación a distancia, sino actualmente se pueden utilizar como complementos en educación de modalidades presenciales, solo que, estas ayudan a complementar las sesiones presenciales en el aula.

Teniendo como referencia estas definiciones, investigaciones relacionadas y aplicaciones, a continuación se analiza el caso de estudio y se menciona la metodología utilizada en esta propuesta.

3 Análisis del caso de estudio

Este trabajo propone el implementar un espacio a través del cual se lleven a cabo las sesiones de clases y el control de los trabajos de los participantes. De tal forma que su uso permite que el tutor pueda calificar y retroalimentar las actividades de los participantes de manera rápida y sencilla. Logrando con ello que un alumno continúe su proceso de enseñanza-aprendizaje fuera del aula.

A continuación se analizan los perfiles que intervienen en el sistema propuesto

Tutor: es la figura que se encarga de introducir, acompañar, asesorar y orientar al estudiante en su formación profesional [12]. Es un docente con los conocimientos específicos sobre los entornos virtuales de aprendizaje.

Alumno: la persona que bajo la guía de un tutor es responsable de su proceso enseñanza-aprendizaje. El cual no puede estar de forma presencial en un salón de clases.

El caso de estudio está orientado a los estudiantes de la carrera de informática de la Universidad de Londres (UDL) campus Querétaro, en las asignaturas de Ingeniería de Software y Soluciones de Software. La razón de la propuesta está basada en el hecho de que la mayor parte de los estudiantes (50%) son activos laboralmente y en ocasiones se les complica acudir en forma presencial a la clase. Además estas materias al ser en su mayoría teóricas, provocan que para algunos alumnos se vuelvan tediosas. La UDL tiene como propósito proporcionar tutorías a los alumnos fuera del horario de clase, y como usualmente los profesores trabajan en otras instituciones se

complica esta situación. Así que el sistema es propuesto para brindar asesorías fuera del horario de clase, pensando en apoyar tanto al profesor como al alumno.

Habiendo descrito a los usuarios que interactúan con el sistema y el caso de estudio, a continuación se detalla la metodología empleada.

4 Metodología

La forma en cómo se comporta el sistema se muestra en la figura 1, en primer lugar el profesor es el responsable de los permisos al sitio Web, así como del material que se tenga disponible para los alumnos. Posteriormente los alumnos a través de un teléfono celular (dispositivo móvil) podrán acceder al material de la clase y continuar con su proceso de aprendizaje fuera del aula.

Funcionamiento del sistema propuesto

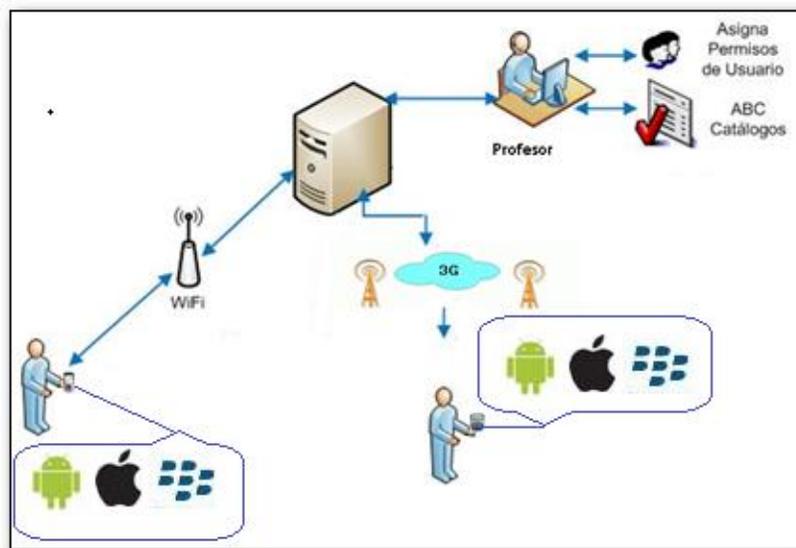


Figura 1. Presenta el esquema metodológico en el cual el sistema propuesto funciona

Esta propuesta está centrada en apoyar el aprendizaje del alumno, pues no importa en dónde se encuentre ubicado éste, al final con el simple hecho de disponer de un dispositivo móvil puede acceder al material de la clase y recibir su retroalimentación respectiva. Además se resalta la interoperabilidad del sistema, pues no importa el teléfono celular que posea alumno ya que basta con que tenga conectividad a Internet para que acceda al material de la clase.

El sistema ha sido dividido en dos módulos, uno correspondiente al sistema Web y el otro a la aplicación móvil. Sin embargo se cuenta con un módulo intermedio al cual se le ha denominado *Educ@-móvil*, a través de este módulo se establecen todas las condiciones y la programación que soporta al sistema.

Módulo 1. Web

Tiene la funcionalidad básica de acceso al sistema y población de datos por medio de catálogos por ejemplo (Alumnos, Tutores, Administradores, Perfiles, Materias).

Funcionalidad 1. Control de acceso al sistema.- Mediante una pantalla de Log in, el tutor captura un usuario y contraseña, y el sistema valida que estén registrados en la base de datos y que tenga los permisos suficientes.

Funcionalidad 2. Catálogos. -Tiene el objetivo de poblar el sistema por medio de una recolección de datos clasificada por alumnos y tutores, con sus respectivos permisos y materias.

Educ@-móvil

Este módulo es un punto intermedio entre el módulo Web y el módulo móvil, entre sus funciones tiene: guardar clase, iniciar clase, asesoría, blog.

Funcionalidad 1. Guardar una clase.- Por medio de esta pantalla el tutor podrá guardar en el sitio la presentación de la clase. Aquí también se resalta el hecho de que se puede adjuntar un video al sistema para que el alumno pueda observarlo.

Funcionalidad 2. Iniciar una clase.- Presenta la pantalla a través la cual el alumno podrá iniciar una clase. Una vez que el alumno ingresa al sistema es capaz de seleccionar la materia a la cual desea acceder y descargar el material respectivo, o en su defecto lee la retroalimentación que ha recibido por parte del tutor.

Funcionalidad 3. Asesoría.- En esta función se podrá realizar la asesoría personal. A través de un chat el tutor y el alumno pueden establecer contacto para resolver dudas.

Funcionalidad 4. Blog.- Aquí se genera un espacio de interacción entre el profesor y el alumno

Módulo 2. Móvil

Consiste en el acceso al sitio por medio de un dispositivo móvil, a través de un acceso a internet o bien una aplicación instalada.

Funcionalidad 1. Control de acceso al sistema.- Mediante una pantalla de Log in, el alumno captura un usuario y contraseña, y el sistema valida que estén registrados en la base de datos y que tenga los permisos suficientes para ingresar.

Funcionalidad 2. Iniciar Clase.- Presenta una pantalla en donde el alumno seleccionará su clase. En el caso de algunos dispositivos móviles como son iPod, iPhone, o con sistema operativo android podrán visualizar inclusive video.

5 Resultados

Después de haber detallado los componentes de la metodología a continuación se presentan imágenes del funcionamiento del sistema propuesto:

Pantalla de bienvenida al sitio Web



Figura 2. Pantalla principal para ingresar al sistema, aquí tanto el profesor como el alumno inician sesión dentro del sistema

Una vez que se ingresa al sistema (figura 2) se pueden desarrollar todas las funcionalidades pertenecientes al modulo Web.

Lo correspondiente al modulo Móvil es presentado en la figura 3, del lado izquierdo se presentan la visualización de cuando el alumno ingresa al sistema y revisa las materias a las cuales tiene acceso, del lado derecho aparece la información cuando ingresa a una materia en particular.

De igual forma se resalta el hecho de que el sistema despliega la cantidad de mensajes que tienen los alumnos por parte de los profesores. Esto permite que cada alumno lleve su propio ritmo, en cuestión de cómo va leyendo los mensajes.

Al momento de utilizar el sistema se obtuvieron resultados favorables, pues en promedio unos 60 alumnos accedieron al sistema y se mostraron satisfechos con los resultados obtenidos. La figura 4, presenta la mejoría de los promedios en los alumnos al momento de utilizar el sistema.

El aprendizaje a través del uso de dispositivos móviles permite que la comunicación entre alumno y el tutor continúe a través de otros canales de comunicación, ya que este es un medio de permitir la continuación de los estudios a pesar de no estar de forma presencial en el aula.

6 Conclusiones

La educación a distancia puede ocurrir simultáneamente en diferentes lugares, basada en el uso de las telecomunicaciones. Debido a que México posee índices de deserción escolar considerables, este trabajo está enfocado en apoyar la educación a distancia a fin de que los jóvenes continúen con su formación escolar. En esta ocasión se presenta una propuesta que hace uso de la tecnología móvil con el fin de que sea utilizado el teléfono celular y no importando en dónde se encuentre el estudiante,

este continúe con su formación educativa. Además también se busca al igual que sucede en un salón de clases, que se establezca retroalimentación por parte del tutor a fin de que este último note el avance del alumno. El sistema funcionando tuvo gran aceptación entre los estudiantes que lo utilizaron, pues su aprovechamiento no decayó por no estar de forma presencial en el aula. El propósito es dejar de ser consumidores de tecnología, para aplicarla y redirigirla atendiendo a las necesidades actuales en el país.

Sistema en funcionamiento con el teléfono celular



Figura 3. En esta pantalla se muestra la forma en que el alumno accede al sistema a través del teléfono celular.,

Graficas con el promedio de los alumnos antes y después de usar el sistema propuesto

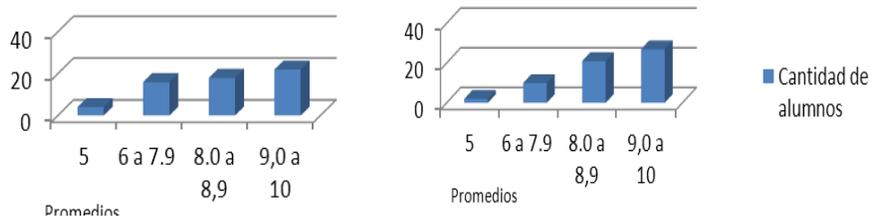


Figura 4. En la gráfica del lado izquierdo se presenta los promedios del grupo (60 alumnos) antes de usar el sistema de tutorías, del lado derecho aparecen los promedios después de usar el sistema de tutorías.

Referencias

1. Arroyo, Luis, (2003). Tecnología Móvil. ISBN: 8441515824. Anaya Multimedia-Anaya Interactiva

2. Cabral, B. (2011). La educación a distancia vista desde la perspectiva bibliotecológica. Universidad Nacional Autónoma de México ISBN. 978-607-02-1425-7. Extraído el 10 de Marzo del 2012 desde http://132.248.242.3/~publica/archivos/libros/educacion_distancia_bibliotecologica.pdf
3. Cúrsame, Sistema de Aprendizaje + Red Social. Extraído el 12 de Marzo del 2012 desde: <http://cursa.me/>
4. Díaz, F. (2008). Educación y nuevas tecnologías de la información y la comunicación: ¿Hacia un paradigma educativo innovador? Sinéctica No.30, revista virtual de educación. Extraído el 7 de Marzo del 2012 desde <http://portal.iteso.mx/portal/page/portal/Sinectica/Revista/fridadb>
5. García Aretia Lorenzo. El impacto de las tecnologías en educación a distancia. Extraído desde el 2 de Junio del 2012 desde http://ipes.anep.edu.uy/documentos/libre_asis/materiales/apr_tec.pdf
6. Güzin S., Koçak, Y. "Modeling educational usage of Facebook" Computers & Education Volume 55, Issue 2, September 2010
7. Hwang G; Kuo M; Wu P.; Huang Y; Zhuang Y. An Investigation on Students' Cognitive Load and Learning Achievements for Participating in a Local Culture Mobile Learning Activity Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE), 2010 6th IEEE International Conference on Digital Object Identifier: 10.1109/WMUTE.2010.28 Publication Year: 2010 , Page(s): 27 - 33
8. Kamata M, Shibusawa S, Yonekura T. e-Learning System "Mobile Web-Com" for Mobile Phones. Computational Intelligence and Security, 2006 International Conference on Volume:1 Digital Object Identifier: 10.1109/ICCIAS.2006.294265 Publication Year: 2006 , Page(s): 886 – 890
9. Sánchez, J. (2010). Deserción escolar afecta competitividad: SEP. El Universal. Obtenido el 26 Marzo del 2012, desde <http://www.eluniversal.com.mx/notas/723195.html>
10. Schlooser, C. Anderson M. (1994, Enero). Distance Education: Review of the literature. Iowa Distance Education Alliance. Extraído el 12 de Marzo del 2012 desde: <http://eric.ed.gov/PDFS/ED382159.pdf>
11. Sharples, M. (2002). Disruptive devices: mobile technology for conversational learning. International Journal of Continuing Engineering Education and Lifelong Learning 2002 Vol. 12, No.5/6 pp. 504-520. Extraído el 7 de Marzo del 2012 desde <http://www.tlu.ee/~kmeta/haridustehnoloogiaTLU/mobilesharples.pdf>
12. Subsecretaría de Educación Superior. Perfil del Tutor. Extraído el 8 de Marzo del 2012 desde http://www.abiertayadistancia.sep.gob.mx/index.php?option=com_content&view=article&id=114&Itemid=103
13. Universidad de Colima. Sistema para la gestión del aprendizaje en línea. Extraído desde el 29 de Mayo del 2012 desde <http://educ.ucol.mx/educmi/index.asp>

Descubrimiento y Visualización de Conocimiento en la Herramienta HACOPDS usando Tecnología de Web Semántica

Luis Alberto Nava-Ordoñez¹, César Enrique Rose-Gómez² y Ana María Chávez-Trejo³

^{1,3}División de Estudios de Posgrado e Investigación
Instituto Tecnológico de Orizaba, Orizaba, Ver., México
¹albert_nava007@hotmail.com, ³achavez@itorizaba.edu.mx
²Instituto Tecnológico de Hermosillo, Hermosillo, Son., México
crose@ith.mx

Paper received on 01/07/12, Accepted on 08/08/12.

Resumen. Dentro del proceso de desarrollo de software es fundamental la gestión de conocimiento, ayuda a los miembros del equipo de desarrollo a entender el conocimiento que existe en el dominio de la aplicación y proporciona información para futuros desarrollos, asimismo proporciona soluciones de problemas ya resueltos anteriormente con lo cual se reduce de manera significativa el desarrollo del mismo, esto se refleja en el oportuno cumplimiento del calendario y uso eficiente de los recursos. En este trabajo se describe un módulo para el descubrimiento y visualización de conocimiento que se integrará a la herramienta HACOPDS, la cual administra el conocimiento orientado al proceso de desarrollo de software.

Palabras Clave: Gestión de Conocimiento, Web Semántica, Ontología, Razonamiento.

1 Introducción.

La construcción de un sistema o producto de software, requiere de una serie de pasos, llamados procesos o subprocesos. Las actividades del proceso de desarrollo se realizan por ingenieros de software, básicamente son cuatro actividades fundamentales en el proceso de desarrollo de software: la especificación, el diseño e implementación, la validación y la evolución del software. Estas actividades generan una gran cantidad de información y conocimiento, lo cual ha motivado en la ingeniería de software a incluir un subproceso para la Gestión del Conocimiento. Este subproceso de gestión del conocimiento de la organización tiene como objetivo entre otros, mantener disponible y administrar la base de conocimientos, que contiene la información y los productos generados por la organización. Hoy en día las empresas que desarrollan software gestionan el conocimiento generado en todas las actividades del proceso de desarrollo de software, algunas muy formalmente, aunque no

todas gestionan el conocimiento de acuerdo a las 4 fases de creación del conocimiento (Obtención, uso/explotación, creación/descubrimiento y difusión del conocimiento). La fase que menos importancia tiene en las empresas, es la fase de creación y descubrimiento de conocimiento, ésta es de gran importancia dentro del desarrollo de un producto de Software, ayuda a la entrega oportuna del producto, presenta información de resolución de problemas, lecciones aprendidas necesarias para nuevos integrantes de los equipos de desarrollo, entre algunas. Para descubrir conocimiento se tiene que analizar una gran cantidad de información, lo que consume tiempo a una persona o grupo de personas, dicha información va aumentando conforme pasa el tiempo, lo que complica el uso de dicha información, esto reditúa en pérdida de tiempo, esfuerzo, recursos humanos, dentro de la organización.

Lo expuesto anteriormente indica que es necesario tener un sistema que permita la gestión de conocimiento para que apoye en este tipo de procesos. Por tal motivo, en el Instituto Tecnológico de Orizaba se desarrolla la herramienta HACOPDS (Herramienta de Administración de Conocimiento para el Proceso de Desarrollo de Software) que permite llevar a cabo la administración del conocimiento orientado a procesos de desarrollo de software.

Este trabajo presenta los avances en el diseño e implementación del módulo de descubrimiento de conocimiento y su visualización, su diseño plantea el uso de tecnologías de la Web Semántica. Este módulo se integrará a la herramienta HACOPDS lo que ha permitido alcanzar el objetivo de la fase de creación y descubrimiento de conocimiento en la teoría de Gestión de conocimiento.

2 Gestión del Conocimiento en Ingeniería de Software.

El Conocimiento según la Real Academia Española [1] se define como *acción y efecto de conocer*, y conocer como *averiguar por el ejercicio de las facultades intelectuales la naturaleza, cualidades y relaciones de las cosas*. El conocimiento se subdivide como conocimiento explícito (es decir artefactos, documentos, reportes y el producto de SW a desarrollar) y conocimiento implícito o tácito (es decir las experiencias obtenidas a lo largo de uno o varios desarrollos de productos de Software).

Thomas Davenport en [2] define la Gestión de Conocimiento como *un método para simplificar el proceso de Compartir, Distribuir, Capturar y entender el conocimiento en una Compañía*. El conocimiento se encuentra implícito en los procesos, y por ende en el desarrollo de un producto de software, es por eso que el conocimiento que se genera en el proceso, se tiene que gestionar, para obtener el mayor aprovechamiento. La Gestión del Conocimiento dentro del proceso de desarrollo del software es importante, dado que ayuda a entender el dominio de la aplicación, proporciona a los desarrolladores información útil para futuros proyectos, permite a desarrolladores novatos entender hasta qué punto de la aplicación se está trabajando y proporciona apoyo a diferentes miembros del equipo de desarrollo a visualizar el conocimiento generado de otras actividades. La figura 1, describe un modelo para sistemas de Gestión del Conocimiento comúnmente empleado en las empresas y la figura 2 muestra las fases del ciclo de creación de conocimiento, de acuerdo con Davenport en [2].

1. **La fase de obtención del conocimiento.** Conocimiento que se genera a través de los motores de búsqueda o de cualquier otra herramienta de filtrado y selección.
2. **Uso/Explotación del Conocimiento.** El conocimiento se utiliza como una parte del proceso de trabajo.

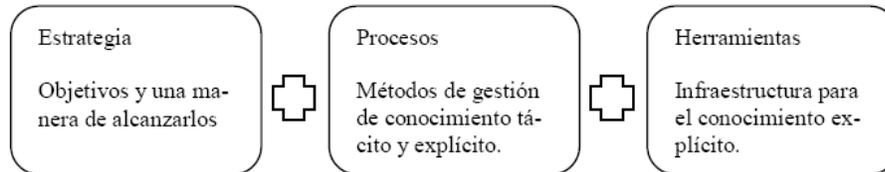


Figura. 1. Un modelo de componentes de un sistema de Gestión del Conocimiento.

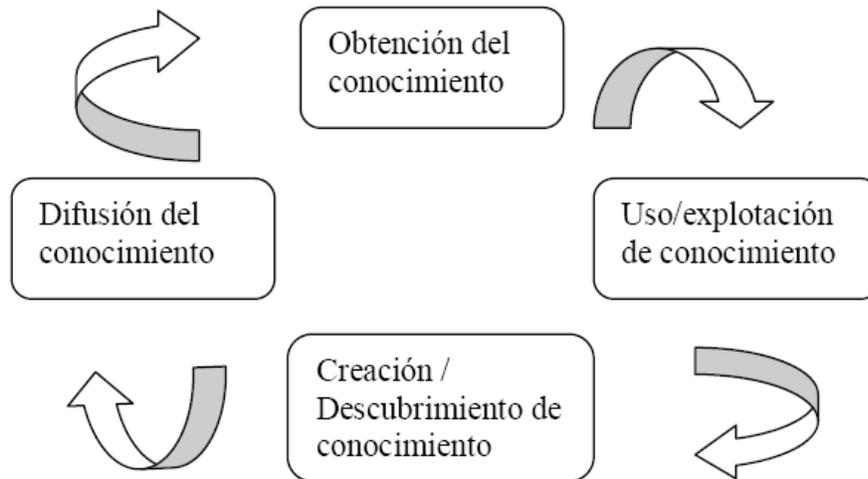


Figura. 2. Ciclo de creación de conocimiento.

3. **Creación/Descubrimiento de Conocimiento.** Permite la obtención de nuevo conocimiento, bien a partir de la creación directa del mismo, o bien a través de su descubrimiento.
4. **Difusión del Conocimiento.** El conocimiento se difunde entre los trabajadores para el aprovechamiento del mismo. Este conocimiento funge como punto de partida para la creación de nuevo, reiniciándose así el ciclo.

Como se mencionó previamente, en el ciclo de desarrollo de software se realizan diversas actividades en donde se tienen algunos problemas, uno de ellos es la falta de documentación que influye en la transferencia de conocimiento [3]. Otro problema es la rotación de personal: empleados nuevos no cuentan con un repositorio de conocimientos que les ayuden a resolver los problemas actuales en base a problemas afrontados en el pasado.

Lo anterior es una pequeña parte del porqué la administración del conocimiento ha sido utilizada cada vez más en las organizaciones de desarrollo de software [4], y su incorporación ha mostrado que tiene beneficios como la reducción de costos, mejoras en la calidad de los procesos y productos [5]. Como se menciona en [6], “existen diversos trabajos centrados en aplicar la administración del conocimiento en organizaciones de desarrollo de software, la mayoría de los cuales se enfocan en la reutilización de experiencias, tales como mejores prácticas o lecciones aprendidas, con el fin de mejorar la calidad de los nuevos productos, del proceso de desarrollo, o para facilitar la reutilización de artefactos de software. Comúnmente, este tipo de trabajos se enfoca en la captura de conocimiento relacionado con el desarrollo de nuevos productos, dejando de lado el mantenimiento del software actual”.

Un claro ejemplo que contrasta con lo anterior, es la metodología KoFI [6], esta metodología se basa en la “identificación de flujos de conocimiento por medio del estudio y modelado de los procesos de una organización, con enfoque en la identificación del conocimiento requerido y generado en las principales actividades del proceso, así como las fuentes donde éste es almacenado y obtenido” [7]. De esta manera, es posible para las personas relacionadas con los procesos, contestar a preguntas como ¿Qué conocimiento es importante para las personas que realizan los procesos? ¿Cuáles son las fuentes que proveen de conocimiento a un proceso? ¿Dónde se encuentra almacenado el conocimiento de un proceso?, entre otras. La metodología KoFI se compone de cuatro etapas: 1) Identificación de fuentes de conocimiento. 2) Identificación de tipos/temas de conocimiento. 3) Identificación de flujos de conocimiento. 4) Identificación de problemas que afecten el flujo de conocimiento. Esta metodología ha sido aplicada en diversos dominios [8], [9], [10] y [11].

Asimismo se han desarrollado herramientas para apoyar la gestión del conocimiento de software, una de ellas es la herramienta SKM [12] la cual está basada en la Web, usa ontologías y gestiona la actividad de diseño dentro del proceso de desarrollo de software. Esta herramienta resuelve una problemática en la actividad de diseño del proceso de desarrollo de software, registra las decisiones de diseño y permite visualizarlas. En HACOPDS se registran las decisiones en las cuatro actividades básicas del desarrollo de Software, la figura 3 visualiza la arquitectura lógica de la herramienta HACOPDS, esta implementa la gestión de usuarios, estos están relacionados con uno o más proyectos en los que gestionaran el conocimiento. Por otro lado la gestión de proyectos y la gestión de artefactos están estrechamente relacionados, ya que cada proyecto que se genere tendrá artefactos que se van a ir generando en cada fase del proceso de desarrollo de Software, las fases que se tomarán en cuenta en HACOPDS son *Requisitos*, *Análisis*, *Diseño*, *Implementación* y *Pruebas*, la herramienta crea un directorio de cada proyecto adjuntándosele los artefactos representativos y clasificados por las fases en las que pertenecen, en el mejor escenario la gestión de experiencias, se conforma de todos los problemas que se le presentan a los desarrolladores en cada fase del proceso de desarrollo.

3 Modelo de Conocimiento.

El modelo de gestión del conocimiento de Nonaka y Takeuchi [13], se fundamenta en el proceso de transformación del conocimiento, es decir, en las diferentes

fases por las que éste pasa en su transformación para ser utilizable por la organización. Desde este enfoque, si tenemos en cuenta la clase de conocimiento, el punto de vista de Nonaka trata la accesibilidad del conocimiento.

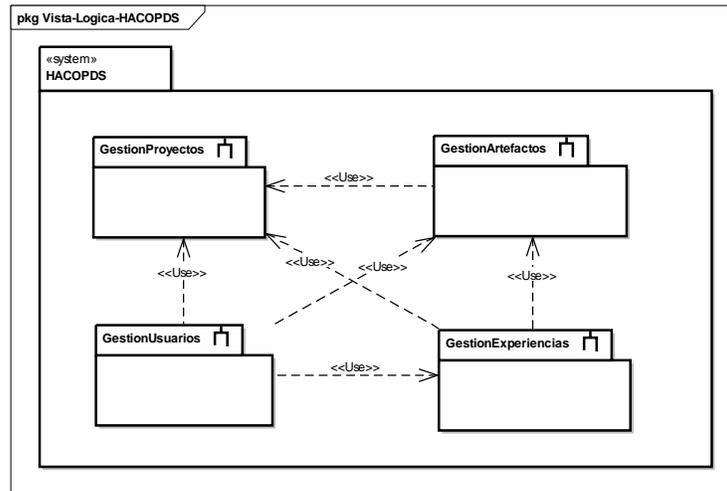


Figura 3. Arquitectura lógica de la herramienta HACOPDS.

Por lo tanto, éste puede ser de dos tipos: tácito y explícito. Es un proceso de interacción entre el conocimiento tácito y el explícito de naturaleza dinámica y continua. Se conforma mediante una espiral, que define la permanente transformación ontológica del conocimiento, desarrollada en cuatro fases. En la figura 4 se muestra dicho proceso.

Socialización (Conocimiento Acordado): es el proceso de adquirir el conocimiento tácito a través de la compartición de las experiencias mediante: exposiciones orales, reuniones, documentos, manuales y tradiciones, que añade el conocimiento nuevo o individual a la base colectiva que posee la organización.

Exteriorización (Conocimiento Conceptual): es el proceso de convertir conocimiento tácito en conceptos explícitos, requiere hacer tangible dicho conocimiento, integrándolo en la cultura de la organización (por ejemplo, en bases de datos de conocimiento).

Combinación (Conocimiento Sistémico): es el proceso de generar conocimiento explícito al reunir conocimiento explícito proveniente de diversas fuentes, y que se puede categorizar, confrontar y clasificar para añadirse a las mencionadas bases de datos de conocimiento.

Interiorización (Conocimiento Operacional): es el proceso de incorporación del conocimiento explícito al conocimiento tácito individual (o grupal).

El módulo de descubrimiento de conocimiento y su visualización a la herramienta HACOPDS, considera la implementación de un modelo se diseñó una ontología, la cual es una representación del conocimiento de un dominio, para el diseño se usó la metodología Methontology [14] para formalizar dicho conocimiento, la ontología es una taxonomía que contiene los conceptos del dominio, además se definen las re-

laciones y atributos. En la figura 5 se muestra de manera parcial la ontología, ésta ha sido implementada en Protégé [15].

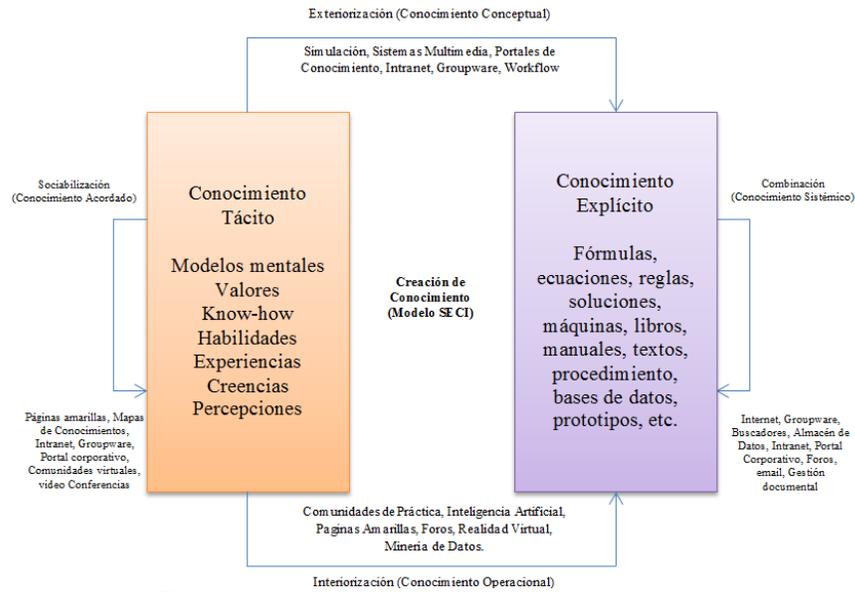


Figura 4. Creación de conocimiento [13].

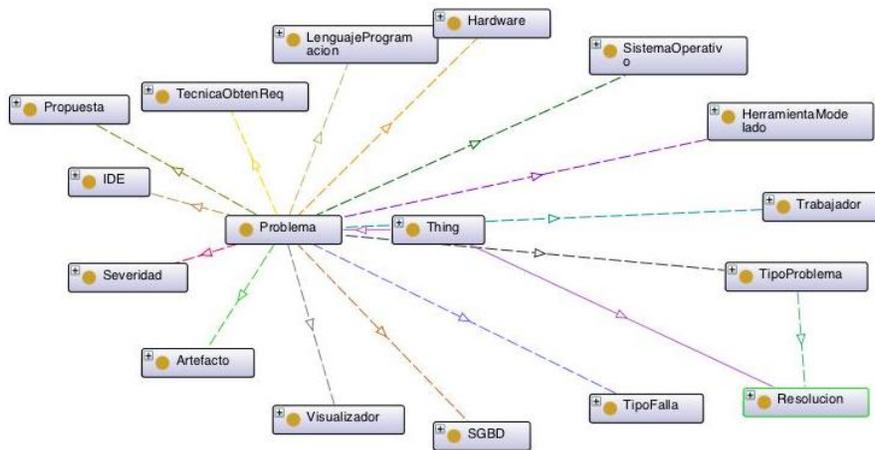


Figura 5. Vista parcial de la ontología.

Como se puede apreciar en la figura 4, el dominio que se ha considerado es acerca de los problemas que se presentan en las diversas etapas del desarrollo de software. De tal manera que se puede apreciar la relación del concepto *Problema* con diversos conceptos que participan en el desarrollo de algún producto de software, por ejemplo el sistema operativo, el lenguaje de programación, tipo de falla, entre

algunas. Estos conceptos contienen instancias que se usarán para determinar *el tipo de problema* que se presenta.

4 Aplicación del Conocimiento.

La aplicación del conocimiento se refiere al proceso de utilizar el conocimiento en la práctica para alcanzar un desempeño competente. El objetivo final de la gestión de conocimiento es el de utilizar el conocimiento para beneficio de la organización. Solamente el uso productivo del conocimiento traducirá los activos intangibles en resultado tangibles. Sin una efectiva utilización, todos los esfuerzos en el desarrollo, almacenamiento y transferencia del conocimiento son en vano.

El uso de la tecnología en la aplicación del conocimiento tiene una influencia positiva al facilitar la captura y accesibilidad del conocimiento organizacional. La tecnología también puede incrementar la cantidad de memoria organizacional disponible, así como permitir la aplicación del conocimiento a través del tiempo y espacio y agilizar la aplicación del conocimiento a través de la automatización.

En nuestro trabajo la aplicación del conocimiento está asociada con el descubrimiento y la visualización del conocimiento acerca de los problemas que se presentan en las fases de desarrollo de software. De tal manera que nuestra primera pregunta es ¿cómo obtenemos conocimiento?, la respuesta a esta pregunta en particular para nuestro caso de estudio es a través del uso de la inferencia a partir de nuestro modelo de conocimiento y con el uso de mecanismos de razonamiento. En la figura 6, se visualiza el diagrama de bloques que muestra el procedimiento general para la obtención del conocimiento, la herramienta HADCOPS cuenta con una base de datos relacional y adicionalmente se tiene una ontología, es por ello que se realizó un mapeo de la base de datos (donde se obtienen el conocimiento tácito) a la ontología como instancias y poder realizar procesos de inferencia como también consultas.

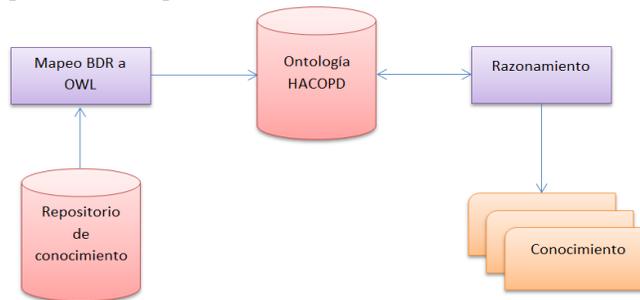


Figura. 6. Diagrama de bloques.

A través de la herramienta se introducen las características del problema (ver figura 7) que se ha presentado y se realiza un proceso de razonamiento que indicará si es un nuevo problema o uno ya existente. Este proceso inferencial se apoya en el uso de una base de conocimiento de reglas, estas reglas modelan los diversos problemas de los que se tiene un conocimiento previo. La figura 8 muestra parcialmente esta base de reglas.



Figura. 7. Interfaz de usuario para la consulta de experiencias.

```
@prefix prob: http://www.semanticweb.org/ontologies/2012/4/Ontology1338231475747.owl#

[problema1: (?A prob:tieneFalla prob:Error_de_Compilacion) , (?A prob:tieneLenguaje prob:Java) ,
(?A prob:tieneIDE prob:NetBeans) , (?A prob:tieneSistema prob:Windows) ,
(?A prob:tieneSGBD prob:PostgreSQL) , (?A prob:tieneHardware prob:PC) ,
(?A prob:tieneVisualizador prob:No_Aplica) , (?A prob:tieneSeveridad prob:Normal)

-> (prob:Problema1 rdf:type prob:TipoProblema) , (?A prob:tieneTipoProblema prob:Problema1) ,
print('El tipo del problema es Problema1')]

[solucion_problema1: (?A prob:tieneTipoProblema prob:Problema1) , (prob:Problema1 prob:tieneResolucion ?B)
-> (?A prob:tieneResolucion ?B),print('La solucion del nuevo problema es: ',?B)]
```

Figura. 8. Reglas.

Dado que la ontología esta representada en el lenguaje OWL [16], se puede aprovechar que el subsistema de inferencia de Jena 2 [17] está diseñado para permitir un rango de máquinas de inferencia o razonadores que pueden ser conectados a Jena. Tales máquinas se emplean para derivar afirmaciones adicionales RDF [18], las cuales se obtienen desde alguna base RDF junto con cualquier información adicional de la ontología, así como los axiomas y reglas asociadas con el razonador. EL uso primario de este mecanismo es soportar el uso de lenguajes como RDFS y OWL, los cuales permiten hechos adicionales a ser inferidos desde las instancias de datos y descripciones de clases. Aquí se usa el término inferencia para referirse al proceso abstracto de derivar información adicional y el término razonador para referirse a un objeto de código específico que ejecuta esta tarea. En nuestro caso de estudio se usa la plataforma Jena para llevar a cabo el proceso de inferencia para descubrir nuevo conocimiento. La figura 9 muestra parcialmente un programa para llevar a cabo el razonamiento con el conjunto de reglas que describen los problemas y el aprendizaje de estos nuevos problemas.

En la herramienta HACOPDS se realizarán diferentes tipos de consulta a la ontología utilizando el lenguaje de consulta SPARQL, este realiza consulta sobre documentos RDF, la estructura de estos documentos esta formada por un conjunto de tripletas (las tripletas están compuestas por un sujeto, predicado y objeto), para la creación del documento que contiene la ontología de la herramienta HACOPDS se

utilizo el lenguaje OWL, este por dentro está estructurado como un documento RDF, es decir contiene un conjunto de tripletas utilizadas para realizar consultas sobre esta. En la figura 10, se muestra un ejemplo de consulta de SPARQL utilizando el marco de trabajo jena. Existen lenguajes de consulta para documentos OWL como OWL-QL, para este caso, no se utilizó de acuerdo a [19], es débil al manejo de operaciones aritméticas y de comparación.

```

public class UsoReglasHACOPDS {
    public static void main(String args[] ) {
        Model datos = ModelLoader.loadModel("C:/HACOPDS.owl");
        List reglas = Rule.rulesFromURL("C:/reglasProblemas.txt");
        GenericRuleReasoner razonador = new GenericRuleReasoner(reglas);
        razonador.setMode(GenericRuleReasoner.HYBRID);
        razonador.setOWLTranslation(true);
        razonador.setTransitiveClosureCaching(true);
        InfModel inf = ModelFactory.createInfModel(razonador, datos);
        inf.prepare();

        try{
            String outputFileName = "C:/HACOPDS.owl";
            FileOutputStream salida= new FileOutputStream(outputFileName);
            inf.write(salida, "RDF/XML");
        }catch(Exception e){
            System.out.println("Error "+e);
        }
    }
}
    
```

Figura. 9. Usando Jena para el razonamiento.

```

public static boolean verifica(OntModel modelo_mapeo,String titulo, String descripcion,
    String tipoFalla, String sistema,String hardware, String ide, String gestor,
    String visualizador, String severidad){

    String queryString =
    "PREFIX prob: <http://www.semanticweb.org/ontologies/2012/4/Ontology1338231475747.owl#>" +
    "prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>" +
    "SELECT ?z " +
    "WHERE {" +
    "?z prob:tieneFalla prob:"+tipoFalla+"." +
    "?z prob:tieneSistema prob:"+sistema+"." +
    "?z prob:tieneHardware prob:"+hardware+"." +
    "?z prob:tieneIDE prob:"+ide+"." +
    "?z prob:tieneSGBD prob:"+gestor+"." +
    "?z prob:tieneVisualizador prob:"+visualizador+"." +
    "?z prob:tieneSeveridad prob:"+severidad+"." +
    "};

    com.hp.hpl.jena.query.Query query = QueryFactory.create(queryString);
    QueryExecution qe = QueryExecutionFactory.create(query, modelo_mapeo);
    ResultSet results = qe.execSelect();
    ResultSetFormatter.out(System.out, results, query);
    qe.close();
    if(results.getRowNumber()==0){
        return false;
    }else{
        return true;
    }
}
    
```

Figura. 10. Usando el lenguaje de consulta SPARQL dentro de la herramienta HACOPDS.

5 Conclusiones.

Una herramienta de gestión del conocimiento debe de emplear las 4 fases que menciona el ciclo de creación del conocimiento, la herramienta HACOPDS describe 3fases del proceso de creación del conocimiento, como consecuencia se creó el módulo encargado de descubrir el conocimiento para cubrir las 4 fases.

En el proceso de desarrollo de Software, los trabajadores o integrantes del equipo conforme van trabajando con el desarrollo de la herramienta y documentación se les presentan problemas que anteriormente ya se les ha dado solución, este artículo muestra una forma de darle al trabajador una posible solución, indicando los criterios y argumentos que fundamentaron la resolución del problema en el pasado, como consecuencia se reducirá el tiempo de desarrollo.

Para la obtención del conocimiento se utilizarán mecanismos de inferencia, estos mecanismos son una forma de encontrar y darle solución a estos problemas comunes. Cada problema tiene asociado una regla que permitirá darle solución a un problema con las mismas características del problema, el uso de reglas es una de las mejores alternativas para darle solución a problemas comunes ya que se aproximan más que a diferentes técnicas de similitud. La desventaja del uso de reglas es que para cada nuevo problema con características diferentes a los problemas ya almacenados que tienen asociada una regla, se debe realizar una nueva regla. Existen otras alternativas de solución que ayudaran a encontrar problemas que tengan o no asociada una regla y es el uso de métricas de similitud, como su nombre lo indica muestran similitud pero podrían ayudar a dar solución a un nuevo problema que no tenga asociada ninguna regla, esto se muestra como trabajo a futuro dentro del módulo de descubrimiento de conocimiento.

Referencias

1. Real Academia Española, <http://www.rae.es/rae.html>, (visitado el 01 de Marzo de 2011).
2. Davenport, T.; Prusak, L.: *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press (1998)
3. Lethbridge, T. C., Singer, J., y Forward, A.: *How software engineers use documentation: The state of the practice*. IEEE Software, 20(6): 35- 39 p. (2003)
4. Aurum, A., Jeffery, R., Wohlin, C., y Handzic, M. (Eds.): *Managing software engineering knowledge*. Springer. Berlin, Germany. 406 p.(2003)
5. Rus, I, y Lindvall, M: *Knowledge management in software engineering*. IEEE Software, 19(3): 26-38 p. (2002)
6. Rodríguez, O.: *Metodología para el diseño de sistemas de administración del conocimiento: su aplicación en mantenimiento de software*, Centro de Investigación Científica y de Educación Superior de Ensenada, Tesis de Doctorado, (2007)
7. Rodríguez Elías, O., Martínez García, A., Favela Jara, J., Viscaíno Barceló, A., Piattini Velthuis, M.: *Facilitando el acceso a las fuentes de información de un grupo de mantenimiento de software por medio de un mapa de conocimiento*. (2005)
8. Rodríguez-Elias, Oscar M.; Morán, Alberto L.; Lavandera, Jaqueline I.; Vizcaíno, Aurora; Soto, Juan Pablo.: *Applying the KoFI methodology to Improve Knowledge Flow in a Manufacturing Process*. In: Proc. of the 3rd International conference on Software

- and Data Technologies (ICSOFIT'2008), Porto (Portugal), 5-8 de Julio, pp. 309-314, ISBN: 978-989-8111-51-7. (2008)
9. Rodríguez-Elias, Oscar M.; Morán, Alberto L.; Lavandera, Jaqueline I.; Vizcaíno, Aurora.: Analyzing Manufacturing Process Knowledge Flow with KoFI", In: Proc. of the 20th International Conference on Software Engineering and Knowledge Engineering (SEKE'2008), San Francisco Bay, USA, 1-3 de Julio. Pp. 428-433, ISBN: 1-891706-22-5. (2008)
 10. Rodríguez Elias, Oscar M.; Vizcaíno Barceló, Aurora; Martínez García, Ana Isabel, Favela Vara, Jesús, PiattiniVelthuis, Mario G.: Aplicando KoFI: Un Caso de Estudio", In: Calidad de Producto y Proceso Software, Calero, Coral; Moraga, Ma. Ángeles; Piattini, Mario (Eds.), Ra- Ma, p. 483-507, (ISBN: 978-84-7897-961-5).(2009)
 11. Zavala Díaz A., Rose GómezC.E.,Rodríguez Elías O. M., Serna Encinas M.T.: Modelo de Conocimiento para un Sistema de Gestión de Conocimiento de una División de Estudios de Posgrado e Investigación Tecnológica, Avances de Investigación en Ingeniería en el Estado de Sonora, Volumen 1, ISBN 978-0-578-09609-4, (2011)
 12. Quiroga J.A.:Gestión de conocimiento en grupos de desarrollo de software, Universidad de los Andes, Bogotá, Colombia. Obtenido Marzo 05, 2011 de <http://paradigma.uniandes.edu.co/media/articulos/jor-quir-1.pdf>. (2011)
 13. Nonaka I, Takeuchi H.: La Organización creadora del Conocimiento. Oxford (1999)
 14. Bernal J.A.: Modelo ontológico para la administración del currículo del personal académico de la Universidad de Sonora, Maestría en Ciencias de la Computación, Instituto Tecnológico de Hermosillo. (2009)
 15. Protégé, <http://protege.stanford.edu/>, (visitado el 8 de Noviembre de 2011).
 16. OWL, <http://www.w3.org/TR/owl-features/>, (visitado el 8 de Noviembre de 2011).
 17. Jena, <http://jena.sourceforge.net/>, (visitado el 8 de Noviembre de 2011).
 18. RDF, http://www.w3.org/standards/techs/rdf#w3c_all, (visitado el 8 de Noviembre de 2011).
 19. Estudio comparativo de lenguajes para la búsqueda y recuperación de información semántica, <http://www.sociedadelainformacion.com/15/lenguajes.pdf>, (visitado el 01 de septiembre 2012).

Una aproximación basada en redes neuronales para el problema de implicación textual translingüe

Darnes Vilariño, David Pinto,
Mireya Tovar, y Beatriz Beltrán

Facultad de Ciencias de la Computación,
Benemérita Universidad Autónoma de Puebla, México
{darnes, dpinto, mtovar, bbeltran}@cs.buap.mx
Paper received on 20/07/12, Accepted on 02/08/12.

Resumen. En este artículo de investigación se presenta una aproximación basada en Redes Neuronales para resolver el problema de Implicación Textual entre diferentes idiomas. Se parte de la hipótesis de que la similitud entre dos textos y su correspondiente longitud puede ayudar a resolver el problema de implicación textual. Se ofrecen dos formas para calcular la similitud textual (léxica y por variación léxica-semántica). Dado que la pareja de textos a evaluar se encuentran en idiomas diferentes, se lleva a cabo un proceso de traducción para posteriormente calcular los grados de similitud entre las oraciones en un mismo idioma. Se usan los resultados reportados en la literatura para comparar aquellos obtenidos por la aproximación propuesta en este trabajo. Los experimentos reportados muestran que sí existe relación entre el tamaño de las oraciones, el grado de similitud y su correspondiente juicio de implicación textual (bidireccional, hacia adelante, hacia atrás o sin implicación).

Palabras Clave: Implicación textual translingüe, Redes Neuronales, Aprendizaje Automático.

1 Introducción

La implicación textual translingüe (Cross-Lingual Textual Entailment o CLTE por sus siglas en inglés) ha sido propuesta recientemente por [4, 5] como una extensión de la tarea de implicación textual [1], la cual se puede definir como sigue: Dado un texto (T) y una hipótesis (H) escritos en diferentes idiomas, la tarea de CLTE consiste en determinar si el significado de H puede ser inferido a partir del significado de T . En este artículo se reportan los resultados obtenidos al evaluar dos características básicas de los textos en cuestión: longitud y grado de similitud. Se determina el rendimiento de ambas características usando un enfoque basado en conocimiento y un enfoque supervisado. Los textos son también analizados cuando se aplica un procedimiento de expansión semántica para verificar si este proceso ayuda o no a mejorar el rendimiento de los algoritmos para la tarea planteada. Parte

de estos experimentos fueron reportados en la Tarea 8 del foro de evaluación semántica SemEval 2012 y que llevó por nombre “*Cross-lingual Textual Entailment for Content Synchronization*” [6].

Para los experimentos llevados a cabo en este artículo, se define la tarea de implicación textual translingüe de manera formal como sigue:

Dado un par de fragmentos de texto tópicamente relacionados (T_1 y T_2) escritos en diferentes idiomas, la tarea consiste en anotar este par automáticamente con alguno de los siguientes juicios de implicación textual:

- Bidireccional (*Bidirectional*) ($T_1 \rightarrow T_2 \ \& \ T_1 \leftarrow T_2$): los dos fragmentos se implican mutuamente (equivalencia semántica).
- Hacia adelante (*Forward*) ($T_1 \rightarrow T_2 \ \& \ T_1 \not\leftarrow T_2$): implicación unidireccional de T_1 a T_2 .
- Hacia atrás (*Backward*) ($T_1 \not\rightarrow T_2 \ \& \ T_1 \leftarrow T_2$): implicación unidireccional de T_2 a T_1 .
- Sin implicación (*No Entailment*) ($T_1 \not\rightarrow T_2 \ \& \ T_1 \not\leftarrow T_2$): no hay implicación textual entre T_1 y T_2 .

En esta tarea, se asume que tanto T_1 como T_2 son declaraciones verdaderas (*TRUE*); de aquí que no existen pares contradictorios. Los conjuntos de datos de entrenamiento y prueba se encuentran disponibles en las siguientes combinaciones de idiomas:

- Español/Inglés (SPA-ENG)
- Alemán/Inglés (DEU-ENG)
- Italiano/Inglés (ITA-ENG)
- Francés/Inglés (FRA-ENG)

A manera de ejemplo, considere los siguientes pares de oraciones (Español/Inglés) que cumplen con uno de los juicios de implicación textual:

Juicio de implicación: Forward

T_1 : La unificación italiana fue el movimiento político y social que aglomeró diferentes estados de la península itálica en el único estado de Italia en el siglo XIX.

T_2 : Italian unification was the political and social movement which agglomerated different states of the Italian peninsula into the single state of Italy.

Juicio de implicación: Backward

T_1 : News Corporation, que pertenece a Rupert Murdoch, ha decidido rechazar un contrato para tomar el control total de la cadena de transmisión BSkyB.

T_2 : News Corporation, held by Rupert Murdoch, has decided to give up the plan of acquiring full control of BSkyB, a broadcasting company headquartered in London.

Juicio de implicación: Bidirectional

T_1 : Ratko Mladic, también conocido como "El carnicero de Bosnia", fue arrestado después de haber sido buscado durante más de una década.

T_2 : Ratko Mladic, nicknamed "The Butcher of Bosnia," has been taken imprisoned after being sought for more than 10 years.

Juicio de implicación: No entailment

T_1 : Un tsunami que se generó en el Pacífico Sur por un poderoso terremoto submarino ha matado a 110 personas.

T_2 : A strong undersea earthquake started a tsunami in the Pacific, leading to the death of at least 110 people with the majority of fatalities in Samoa.

Las oraciones anteriormente mostradas fueron tomadas aleatoriamente del conjunto de datos de entrenamiento empleado en los experimentos presentados en este trabajo de investigación. Son usadas únicamente con el propósito de mejorar la comprensión del artículo, pero de ninguna manera reflejan un comportamiento generalizado sobre los pares de oraciones y los juicios de implicación textual.

El resto de este artículo se encuentra estructurado de la siguiente manera: la sección 2 describe las diferentes aproximaciones presentadas en este trabajo. Los resultados obtenidos son mostrados y discutidos en la sección 3. Finalmente, se discuten las conclusiones y trabajo a futuro en la sección 4.

2 Descripción de las propuestas

Para el experimento llevado a cabo en este trabajo, hemos considerado atacar el problema de CLTE usando principalmente dos características: tamaño de las oraciones (longitud textual) y la similitud textual entre oraciones del mismo idioma. Para el caso de similitud textual, hemos considerado un enfoque léxico y uno por variación léxica-semántica basado en expansión de términos (de manera concreta, sinónimos).

El juicio de implicación textual es llevado a cabo de dos maneras diferentes: 1) usando reglas de decisión planteadas de manera empírica, y 2) usando una red neuronal entrenada a priori.

Con la finalidad de encontrar el grado de similitud entre dos oraciones en el mismo idioma, fue necesario traducir las oraciones originales (escritas en dos idiomas distintos) al idioma contraparte, es decir, si el par de oraciones (T_1 , T_2) es Alemán-Inglés, entonces T_1 se traduce al inglés y T_2 se traduce al alemán. Para este propósito hemos usado Google Translate¹.

Antes de presentar los modelos de determinación del juicio de implicación textual usando reglas empíricas y redes neuronales, presentamos a continuación dos formas de calcular la similitud textual entre dos oraciones.

2.1 Cálculo de similitud textual

Sea T_1 la primera oración del par a analizar su implicación textual y escrita en el idioma origen (Español, Alemán, Italiano o Francés) y T_2 el fragmento de texto típicamente relacionado con T_1 (dado en idioma Inglés), entonces, se obtiene T_3 que es la traducción al Inglés de T_1 , y T_4 que es la traducción de T_2 al idioma origen (Español, Alemán, Italiano o Francés). El cálculo de similitud textual a nivel léxico y por variación léxico-semántica se describe a continuación. En ambos casos, se usa la oración original, es decir, no se realiza preprocesamiento alguno de las mismas.

¹ <http://translate.google.com.mx>

2.1.1 Similitud léxica

Primeramente se determina la similitud léxica entre dos textos escritos en el idioma origen (*SimS*), es decir, entre T_1 y T_4 . Adicionalmente, se calcula la similitud léxica entre las oraciones que se encuentran escritas en inglés (*SimT*), es decir, entre T_2 y T_3 .

Para este propósito se ha utilizado el coeficiente de similitud de Jaccard [7]. La ecuación (1) muestra la similitud léxica para dos textos escritos en el idioma origen. De la manera similar se define *SimT* como el grado de similitud para los dos fragmentos de textos escritos en el idioma inglés.

$$SimS = simJaccard(T_1, T_4) = \frac{|T_1 \cap T_4|}{|T_1 \cup T_4|} \quad (1)$$

2.1.2 Similitud léxico-semántica

Para calcular el grado de similitud léxico-semántica se considera una expansión de los términos originales en las oraciones usando los sinónimos de cada palabra (en el idioma origen y en el idioma destino). Se emplearon cinco diccionarios que contienen los sinónimos de los idiomas considerados en el marco del SemEval 2012 (inglés, español, alemán, italiano y francés)² para este propósito.

En la Tabla 1 se muestra tanto el número de términos, como el número de sinónimos en promedio por término considerados para cada idioma.

Tabla 1. Diccionarios de sinónimos usados para la expansión de términos.

Idioma	Términos	Num. promedio de sinónimos por término
Inglés	2,764	60
Español	9,887	45
Alemán	21,958	115
Italiano	25,724	56
Francés	36,207	93

Sea $T_1 = w_{1,1}w_{1,2} \dots w_{1,|T_1|}$ y $T_2 = w_{2,1}w_{2,2} \dots w_{2,|T_2|}$ las oraciones origen y destino, respectivamente. Sea $T_3 = w_{3,1}w_{3,2} \dots w_{3,|T_3|}$ y $T_4 = w_{4,1}w_{4,2} \dots w_{4,|T_4|}$ las versiones traducidas de las oraciones en el idioma origen y destino, respectivamente. El conjunto de sinónimos para la palabra/término $w_{i,k}$ se expresa como $Synset(w_{i,k})$, y se obtiene a partir de los diccionarios anteriormente mencionados. Para lograr un mejor grado de similitud entre ambos textos, cada palabra se ha generalizado utilizando el truncador de Porter [8].

Para calcular la similitud léxico-semántica entre dos palabras escritas en el idioma origen ($w_{1,i}$ y $w_{4,j}$), se usa una ecuación similar a la ecuación (1). De la misma

² <http://extensions.services.openoffice.org/dictionaries>

manera se puede obtener $Sim(w_{2,i}, w_{3,j})$, que es la similitud por variación léxico-semántica entre las palabras escritas en el idioma inglés.

$$Sim(w_{1,i}, w_{4,j}) = \begin{cases} 1, & \text{si } (w_{1,i} = w_{4,j}) \text{ OR } w_{1,i} \in SynSet(w_{4,j}) \\ & \text{OR } w_{4,j} \in SynSet(w_{1,i}) \\ 0, & \text{en otro caso} \end{cases} \quad (2)$$

Ambas oraciones consideran la existencia de similitud cuando dos palabras son idénticas, o cuando al menos alguna de las dos palabras aparece en el conjunto de sinónimos de la otra palabra.

La similitud entre los fragmentos de texto T_1 y T_4 ($SimS$) se calcula por medio de la ecuación (3), y de la misma manera se obtiene $SimT(T_2, T_3)$ que es la similitud léxico-semántica completa entre los fragmentos de texto T_2 y T_3 .

$$SimS(T_1, T_4) = \frac{\sum_{i=1}^{|T_1|} \sum_{j=1}^{|T_4|} Sim(w_{1,i}, w_{4,j})}{|T_1 \cup T_4|} \quad (3)$$

2.2 Determinación del juicio de implicación textual

En esta sección se presentan dos aproximaciones para determinar el correspondiente juicio de implicación textual. Para ambos casos se utilizan dos formas de calcular la similitud textual (léxica y léxico-semántica), obteniendo cuatro experimentos que son evaluados y comparados en la Sección 3.

2.2.1. Aproximación empírica

En este caso, se ha considerado el tamaño de la oración como una característica en la determinación del juicio de implicación textual. De manera particular, se compara la longitud de la oración T_2 con respecto a la oración T_3 (ambas escritas en el idioma inglés). Un análisis empírico nos ha llevado a considerar que si la longitud de T_2 es menor que la longitud de T_3 y existe además un grado de similitud textual (léxica o semántica) superior a 0.5 para los pares de oraciones escritos en los mismos idiomas, entonces, se establece un juicio tipo *Forward*. Por otro lado, si la longitud de T_2 es mayor que la longitud de T_3 y existe además un grado de similitud textual (léxica o léxico-semántica) superior a 0.5 para los pares de oraciones escritos en los mismos idiomas, entonces, se establece un juicio tipo *Backward*. En el caso en el que, la longitud de T_2 sea igual que la longitud de T_3 , la longitud de T_1 sea igual que la longitud de T_4 , y existe además un grado de similitud textual (léxica o léxico-semántica) superior a 0.5 para los pares de oraciones escritos en los mismos idiomas, entonces, se establece un juicio tipo *Bidirectional*. Finalmente, si ninguno de los casos anteriores se cumple, se establece un juicio de *No entailment*. Estas reglas de decisión pueden verse más claramente en el Algoritmo 1.

Algoritmo 1.

```
If  $|T_2| < |T_3|$  then
  If  $(simT > 0.5 \text{ and } simS > 0.5)$ 
    then Forward
```

```

ElseIf  $|T_2| > |T_3|$  then
  If ( $simT > 0.5$  and  $simS > 0.5$ )
    then Backward
ElseIf ( $|T_1| = |T_4|$  and  $|T_2| = |T_3|$ ) then
  If ( $simT > 0.5$  and  $simS > 0.5$ )
    then Bidirectional
Else No entailment

```

Dado que se usan dos tipos de similitud textual, se han obtenido dos aproximaciones basadas en la propuesta empírica anterior. Hemos llamado *EMP_Lex*, a la propuesta que usa la similitud textual léxica, mientras que aquella que usa la similitud por variación léxico-semántica se ha denominado *EMP_Sem*.

2.2.2. Aproximación supervisada

Para esta aproximación se utilizó una red neuronal tipo perceptrón multicapa de propagación hacia atrás (implementado en Weka [2]), con 20 neuronas en la capa de entrada, 7 neuronas en la capa intermedia y 4 de salida. Las neuronas de la capa de entrada combinan las seis características consideradas para la entrada de la red que son: cuatro longitudes de oración (para T_1 , T_2 , T_3 y T_4), y los dos valores de similitud textual ($SimS$ y $SimT$). Los parámetros de la red neuronal fueron ajustados logrando un 95.7% de instancias correctamente clasificadas en el proceso de entrenamiento.

Dado que se usan dos tipos de similitud textual, se han obtenido dos aproximaciones basadas en la propuesta supervisada anterior. Hemos llamado *NN_Lex*, a la propuesta que usa la similitud textual léxica, mientras que aquella que usa la similitud por variación léxico-semántica se ha denominado *NN_Sem*.

3. Experimentos

Los corpora usados en los experimentos provienen del conjunto de datos para implicación textual translingüe presentados en [3] y proporcionados por los organizadores de la Tarea 8 de SemEval 2012 [6]. En el caso de las reglas de decisión empíricas, se emplearon los conjuntos de entrenamiento únicamente para ajustar los parámetros, mientras que en el caso de la red neuronal, dichos conjuntos se usaron para entrenar los pesos de la red. La cantidad de oraciones contempladas tanto en los conjuntos de entrenamiento como de prueba fue de 500 para cada idioma.

En la tabla 2 se muestran los resultados globales obtenidos por las cuatro aproximaciones propuestas en este artículo. Adicionalmente, hemos incluido los puntajes máximo (*Mejor*), mínimo (*Peor*), promedio (*Promedio*) y mediana (*Mediana*), reportados en la competencia de la Tarea 8 de SemEval 2012; esto con fines de comparación.

Tabla 2. Resultados globales obtenidos por las propuestas y su comparación con aquellos obtenidos en la Tarea 8 de SemEval 2012

	SPA-ENG	ITA-ENG	FRA-ENG	DEU-ENG
Mejor	0.632	0.566	0.57	0.558
Promedio	0.407	0.362	0.366	0.357
Mediana	0.346	0.336	0.336	0.336
Peor	0.266	0.278	0.278	0.262
EMP_Lex	0.350	0.336	0.334	0.330
EMP_Sem	0.366	0.344	0.342	0.268
NN_Lex	0.476	0.456	0.486	0.458
NN_Sem	0.398	0.398	0.392	0.386

Los puntajes muestran que la aproximación supervisada se encuentra por arriba de la mediana de los resultados reportados en la literatura. Si bien, aun no se alcanza el mejor valor obtenido en la competencia, se ha observado que las dos características básicas que se han considerado son sumamente importantes, pues se acercan considerablemente al mejor puntaje. Cabe aclarar que no se han hecho análisis semánticos profundos, lo cual es claramente una ventaja de las propuestas planteadas. Sin embargo, se considera que es un buen punto de partida y que al agregar un análisis más profundo de las oraciones en cuestión, seguramente se obtendrá un mejor rendimiento.

Tabla 3. Resultados obtenidos por la red neuronal (*NN_Lex*) para cada uno de los juicios de implicación textual (sin expansión de términos).

Idiomas	Backward			Forward			No-Entailment			Bidirectional			Promedio general
	P	R	F ₁	P	R	F ₁	P	R	F ₁	P	R	F ₁	
SPA-ENG	0.57	0.50	0.53	0.57	0.52	0.55	0.39	0.37	0.38	0.40	0.49	0.44	0.476
3	4	6	4	8	0	2	6	4	0	6	3		
ITA-ENG	0.55	0.52	0.53	0.54	0.46	0.50	0.38	0.43	0.40	0.37	0.40	0.39	0.456
1	0	5	2	4	0	8	2	9	5	8	1		
FRA-ENG	0.56	0.55	0.55	0.56	0.57	0.57	0.39	0.38	0.38	0.42	0.43	0.42	0.486
1	2	6	7	6	1	3	4	9	2	2	7		
DEU-ENG	0.54	0.54	0.54	0.51	0.58	0.54	0.43	0.29	0.35	0.34	0.40	0.37	0.458
4	4	4	0	4	5	0	6	1	9	8	6		

Tabla 4. Resultados obtenidos por la red neuronal (*NN_Sem*) para cada uno de los juicios de implicación textual (con expansión de términos).

Idiomas	Backward			Forward			No-Entailment			Bidirectional			Promedio general
	P	R	F ₁	P	R	F ₁	P	R	F ₁	P	R	F ₁	
SPA-ENG	0.51	0.57	0.54	0.53	0.53	0.53	0.22	0.19	0.20	0.28	0.28	0.28	0.398
1	6	1	6	6	6	6	2	8	1	8	5		
ITA-ENG	0.54	0.58	0.56	0.50	0.37	0.42	0.27	0.24	0.25	0.30	0.38	0.34	0.398
1	4	2	0	6	9	0	8	8	8	4	2		
FRA-ENG	0.46	0.46	0.46	0.61	0.56	0.59	0.22	0.29	0.25	0.31	0.24	0.27	0.392
0	4	2	7	8	2	4	6	5	9	0	4		
DEU-ENG	0.51	0.40	0.45	0.49	0.51	0.50	0.29	0.28	0.28	0.28	0.35	0.31	0.386
5	0	0	2	2	2	4	0	7	6	2	5		

Con la finalidad de analizar a profundidad los resultados obtenidos con las aproximaciones basadas en redes neuronales (*NN_Lex* y *NN_Sem*), en las Tabla 3 y 4 se muestran los valores de precisión (*P*), *recall* (*R*) y *F₁* obtenido para el juicio de implicación textual. Ahí se puede observar que al menos para la aproximación plan-

teada, el uso de sinónimos generaliza demasiado, incrementando el valor de similitud textual y confundiendo los juicios de implicación finales. Analizando los valores obtenidos, se puede ver que las características usadas son significativas para el caso de juicios *Forward* y *Backward*, sin embargo, no son suficientes para determinar con buen grado de precisión los juicios *Bidirectional* y *No entailment*. Como trabajo a futuro se plantea revisar cuáles son las características más relevantes para estos dos últimos juicios de implicación textual y así mejorar los resultados obtenidos.

4 Conclusión

En este trabajo se presentan cuatro aproximaciones para resolver el problema de implicación textual translingüe. Dos aproximaciones usan un enfoque empírico empleando similitud textual léxica y por variación léxica-semántica. Las otras dos aproximaciones usan un enfoque supervisado contemplando también ambos tipos de similitud textual.

Se ha observado que la determinación del juicio *No entailment* es complicada, pues existen oraciones que no se implican textualmente, pero comparten términos comunes. Para el caso del juicio *Bidirectional*, se ha obligado a que las longitudes de las oraciones sean exactamente iguales, lo cual es una restricción muy estricta. Por lo que, se debe buscar una estrategia para relajar esta restricción.

A pesar de las limitaciones anteriores, se considera que el uso de las longitudes de las oraciones en la determinación de un juicio de implicación textual translingüe es una característica interesante, pues es relativamente simple de calcular y que no requiere un análisis semántico profundo.

Se está en proceso de evaluar el efecto de la traducción en los porcentajes obtenidos cuando se determina el juicio de implicación. Como trabajo a futuro, se piensa usar corpus paralelos para obtener diccionarios estadísticos que puedan mapear de mejor manera el significado de las oraciones.

Referencias

1. Dagan, Ido and Oren Glickman: Probabilistic Textual Entailment: Generic Applied Modeling of Language Variability. In PASCAL Workshop on Learning Methods for Text Understanding and Mining, Grenoble, January 2004.
2. Ian H. Witten and Eibe Frank. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2005).
3. Manning, Christopher D., Schütze, Hinrich: Foundations of Statistical Natural Language Processing. MIT Press, 6a Edición (2003)
4. Mehdad, Yashar, Matteo Negri, and Marcello Federico: Towards Cross-Lingual Textual Entailment. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles, California, 321–324 (2010).
5. Mehdad, Yashar, Matteo Negri, and Marcello Federico: Using bilingual parallel corpora for cross-lingual textual entailment. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, HLT '11, Stroudsburg, PA, USA, Volume 1, 1336–1345 (2011).

6. Negri, M., A. Marchetti, Y. Mehdad, L. Bentivogli, and D. Giampiccolo: Semeval-2012 Task 8: Cross-lingual Textual Entailment for Content Synchronization. First Joint Conference on Lexical and Computational Semantics (*SEM), pages 399–407, Montreal, Canada, June 7-8, (2012).
7. Negri, Matteo, Luisa Bentivogli, Yashar Mehdad, Danilo Giampiccolo, and Alessandro Marchetti: Divide and conquer: crowdsourcing the creation of crosslingual textual entailment corpora. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '11, Stroudsburg, PA, USA, 670–679 (2011).
8. Porter, M: An algorithm for suffix stripping. Program, 14(3):130–137 (1980).

Multiagent-based decision making in strategy games

Diana Fernández-Tirado¹, Adolfo Pacheco-Nieto² and Gilberto G. Flores-Vidaña³

Instituto Tecnológico de Estudios Superiores de Monterrey, Monterrey, N.L., México.

¹A00746263@itesm.mx, ²A00812892@itesm.mx, ³A00735176@itesm.mx

Paper received on 25/07/12, Accepted on 06/09/12.

Abstract. The video game industry strategy aims to generate more intelligent games, where characters have the ability to make decisions and reach agreements in order to increase the challenge and keep gamers captives. This article proposes a negotiation and decision-making model for agents in a strategy game based on Contract Net, and shows how an approach based on Multi-agent System and Contract-Net can have better results than an independent agent. The system analysis and design was modeled using GAIA methodology in which we can describe the interactions and behavior of the agents.

Keywords: Multi-agents Systems, Contract-Net, GAIA, strategy game.

1 Introduction

In the area of manufacturing we have process planning and scheduling that are two, usually considered, different activities that are handled by techniques of Distributed Artificial Intelligence [6] such as Multi-Agent Systems. There is extensive research around these problems as we can see in the survey presented by Shen [7] whereas in Lu [8] proposes a novel agent bidding based approach that searches for optimal solutions in manufacturing planning and scheduling. Decision-making and implementation of agreements in games of strategy is a challenge that can be addressed with Multi-agent techniques like Contract-Net.

The problem addressed in this article is to increase the intelligence of the Knights and Workers (as agents in the game) in the following scenario: the scene shown in the screen is a map, a civilization is located in an arbitrary point and starts with few Workers and the MainBuilding. In order to create more Buildings and Workers or train Knights, the civilization needs resources (gold and wood), Workers can obtain these resources from mines and trees. The Workers also can create buildings that are able to defend the civilization from enemies' attacks. All around the map there are enemies, that can be walking around or staying in a certain point, these enemies can assemble and decide (or not) to attack the civilization that is being built in the map. If the enemies destroy the MainBuilding of the civilization, the game is over. The Knights can improve their strength, when a Knight attacks an enemy it gains experience and eventually it is able to kill stronger enemies. The goal

is to build a strong civilization and kill the MightyDragon that resides located in the map.

This paper shows the process to implement a Multi-agent System using the GAIA methodology to describe interactions between agents and Contract Net protocol for negotiation between agents, the project presents a scheme based on utilities in the strategy of Knight-Teamwork (in pursuit of the goal of killing the Mighty Dragon).

The remaining of this paper is organized as follows: section 2 presents the use of GAIA methodology in the system analysis and design of the game, in section 3 the prototype (developed in Net Logo) is described and a formula to calculate the utility used for decision-making in the Contract Net protocol is proposed, in section 4 the experiments made for this project along with the results obtained are detailed, and finally in section 5 we draw our conclusions.

2 Analysis and Design

Wooldridge, Jennings and Kinny [4] argue that the traditional models for software analysis and design are not suitable for agent-based systems because they are unable to represent unique aspects of agents. They present the GAIA methodology, which has been tailored specifically for these kinds of systems.

The GAIA methodology is a process that leads the developer to go from the requirements to a design with such a detail that can be implemented straight [4]. Gaia has two main concept categories: abstract and concrete, the result of this process is: (1) during the analysis stage the roles and interaction models are defined and (2) during the design stage the agent, services and acquaintance models are described.

2.1 System Organization

In the game named “Knights and Dragons” the environment is composed by the map (in which the game will be developed) with the resources located all around the map (mines and trees), and the goal is to build a civilization strong enough to kill all enemies in the map.

Raising and protecting a civilization is a complex process that needs to be handled by several kinds of agents. In order to achieve the goal of the civilization, agents must agree how to interact, and thus will allow them to get the highest benefit to the civilization.

The agents are related to their environment because the map is where the civilization will develop, and the agents (Workers and Knights) will be scouting around the map in order to find more resources that will allow the civilization to keep growing, get stronger, defend against enemies and kill them. The general structure of the agents is shown in Figure 1.

The game starts with the MainBuilding, which will create Workers in order to start gathering resources needed to create buildings and increase the army. To create more Workers we need to check if there is enough food, otherwise more Huts are needed, in this way the civilization does not get limited by food.

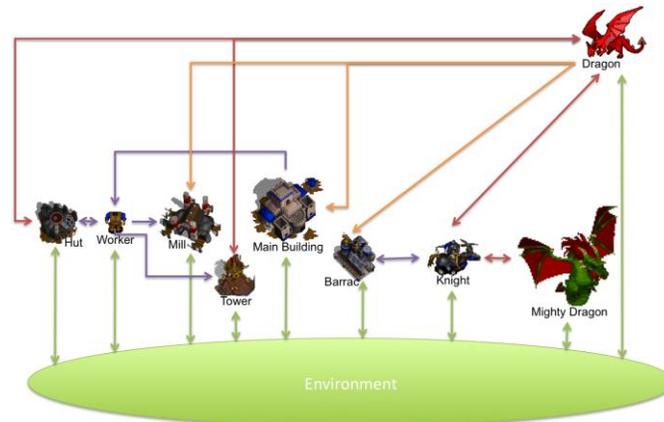


Figure 1: General Structure of Agents.

The game starts with the MainBuilding, which will create Workers in order to start gathering resources needed to create buildings and increase the army. To create more Workers we need to check if there is enough food, otherwise more Huts are needed, in this way the civilization does not get limited by food.

The Workers will start gathering resources or creating buildings, and if they require to create a building they need to ask for resources to the MainBuilding and this one will let them know if there are enough resources for the new building.

As soon as the Barrack gets built, it can start creating Knights, it has to be aware of how many Knights had been created, so the civilization does not have Knights they do not really need (yet).

The Knights will start scouting all around the map and if they find an enemy they will call for help from other Knights.

Buildings like the Hut and the Tower will help defend the town if enemies attack, when they notice enemies' attacks, the Workers will get into Huts in order to protect themselves and at the same time attack enemies.

The Dragons will be all around the map, they can be protecting a gold mine or just flying around the map, they can get to the town and attack, and even can agree with other Dragons to attack.

The Mighty Dragon is the strongest enemy in the map, the Knights have to find it and kill it, but as it is so strong they will need first to gain experience by killing other Dragons of different levels.

2.2 Roles Description

In the context of Gaia methodology a role is an abstract description of an entity's expected function and it is defined by four attributes: protocols, permissions, activities and responsibilities [3]. The role model states the roles an agent can perform in the environment described earlier. Role description, activities, protocols, responsibilities and permissions for the MainBuilding are presented as a significant example in Table 1.

Table 1: Schema for Role MainBuilding.

ROLE SCHEMA: MainBuilding	
<p>Description: Is the base of the town, it keeps the gold and the wood. This building can create Workers, and must decide if there are enough Workers or needs to create more. If this building is destroyed the civilization cannot survive.</p>	
<p>Activities:</p> <ul style="list-style-type: none"> • Check-GoldStock • Check-WoodStock • CheckFoodLimit • Die • CreateWorker 	<p>Protocols:</p> <ul style="list-style-type: none"> • IncreaseWoodStock • IncreaseGoldStock • RequestHutCreation • GetResource • GetHost • ReceiveAttack
<p>Permissions:</p> <p>Reads</p> <ul style="list-style-type: none"> • IsCreatingWorker // TRUE FALSE • IsInRequestForHutCreation // TRUE FALSE <p>Changes</p> <ul style="list-style-type: none"> • FoodLimit // Real Number • GoldStock // Real Number • WoodStock // Real Number 	<p>Responsibilities:</p> <p>Liveness: (([Upgrade] [RequestHutCreation]) (GetHost . GetResource . CreateWorker)) (CheckGoldStock . CheckWoodStock . CheckFoodLimit . IncreaseWoodStock . IncreaseGoldStock) ([ReceiveAttack] . [Die])) w</p> <p>Safety Properties:</p> <ul style="list-style-type: none"> • This building must survive • Create Workers when the town requires them • Gold Stock > 50 and Hut Space > 1 • Gold Stock > 0 • Wood Stock > 0 • Workers > 0 • Health points > 0

2.3 Interaction Model

The interaction model is described in GAIA as the interactions between the roles that lead them to reach their goal [3]. Examples of the protocols used to describe interactions between agents are:

IncreaseGoldStock		
Worker	MainBuilding	LevelOfGold
After the Worker collects the gold, it can carry it from a mine; it goes to the MainBuilding or Mill and deposits the resource.		GoldStock
GetResource		
AbstractBuilding, Barracks, Hut	MainBuilding	AmountOfGold, AmountOfWood, Host
A building asks for resources to the MainBuilding to create elements (Workers, Towers).		WoodStock, GoldStock, FoodLevel
Attack		
Knight, Dragon, Tower MightyDragon, Hut,	Knight, Dragon, Tower MightyDragon, Hut,	Enemies
When an enemy is in attack range of the unit, the unit		

attacks the enemy.		
AskForHelp		FindEnemies,
Knight	Knight	AttackFromEnemy
When a Knight find one or more enemies or it is attacked by enemies, the Knight calls for other Knights to help him.		

2.4 Activities Description

Role's activities are actions that an agent can perform, and can imply interactions with other agents or not [3]. The activities realized by the roles described in the document are:

CreateWorker: When the MainBuilding detects that more Workers are needed, if the civilization has the amount of gold needed to create a Worker and there are enough Huts, the MainBuilding may decide to create a Worker.

CheckForEnemies: This activity can be done for some buildings, and it allows to know the existence of enemies nearby, so defensive or offensive actions can be taken.

CheckFoodLimit: The MainBuilding validates the maximum amount of food allowed by total Huts.

CheckFoodLevel: The MainBuilding takes control of the food level that exist in all the available Huts.

CheckGold/CheckWoodStock: The MainBuilding or Mill verifies constantly the amount of gold/wood that the civilization has.

DefendMine: This activity is used by the Dragons to protect a gold mine from Knights. And this activity force the Dragons to stay in the mine waiting if a group of humans appear.

Die: This activity can be used by any role; the role that uses it disappears of the game.

Ignore: This activity is used by the MightyDragon to ignore the Workers nearby.

Scouting: The Knights explore the map in order to find wood, mines or Dragons.

Runaway: This activity takes place if a Dragon or a Knight thinks it is better to flee than to attack.

Walking: This activity is used by the Dragons to explore the map if they aren't protecting a mine.

2.5 Agent Model

During this step of the design stage we will identify “the agent types that will make up the system, the agent instances that will carry out these agent types at runtime, and the mapping between roles and agent types”[3].

The agent model (Figure 2) shows which roles are associated with each agent and how many agents of each type may exist in software. This model indicates that there is only one TownHallAgent because the game objective is to keep it safe. In

the simulation strategy game the creation of a set of DefensiveTowerAgents is needed to defend the town from the Dragons' attacks. To increase the FoodLimit of civilization is important to create several HutAgents. There is only one agent who creates KnightAgents, the BarracksAgent. The MillAgents are used to gather wood and gold.

During the game there will be several WorkerAgents that create the buildings, collect gold and wood, besides defending the people using the HutAgent. The KnightAgents are responsible for performing the exploration of the map and release the mines from DragonAgents; there may be several Knights in the game. The DragonAgents protect mines or walk around the map. Finally there must be only one MightyDragonAgent because its strength is higher than the other Dragons.

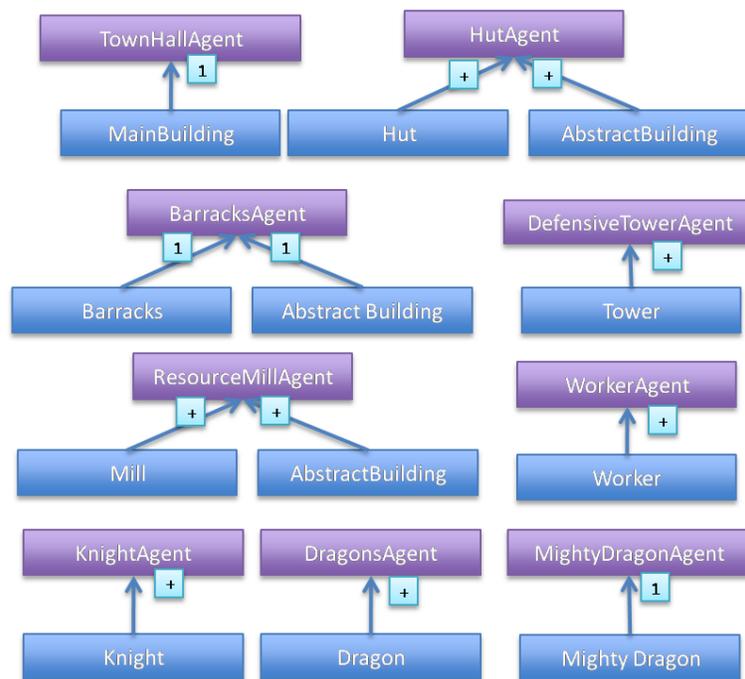


Figure 2: Agent Model.

2.6 Acquaintance Model

The goal of this step of the process is to document the communication lines that exist between the agents [3]. This model presents the existing interaction between agents (Figure 3), which are grouped in two main categories: buildings and characters. In the buildings category we can find MainBuildingAgent, the main function of this agent is to create Workers and keep gold/wood. In order to increment defenders for the town, the BarracksAgent creates and trains Knights. In the civilization the WorkerAgent can create TowerAgents, which can attack enemies and protect other buildings.

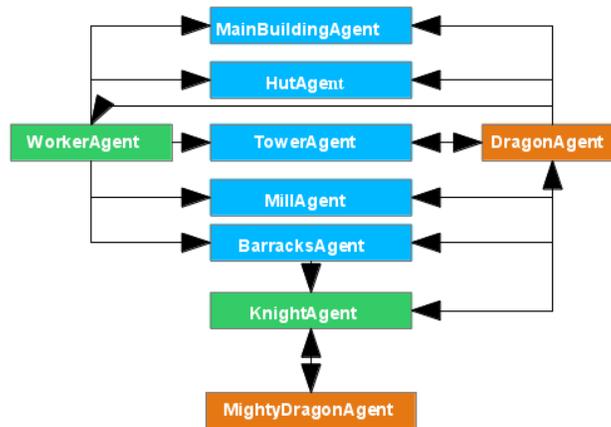


Figure 3: Acquaintance Model.

The characters category is compound of the WorkerAgents, KnightAgents, DragonAgents and MightyDragonAgent. The main interactions for WorkerAgent are the recollection of resources and creations of buildings. The KnightAgent attacks enemies, defends the town and goes scouting. MightyDragonAgent attacks the KnightAgents that are around it.

3 Prototype description



Figure 4: Screen of the software civilization simulator.

The Prototype for this article was implemented using Net-Logo. The prototype includes the following features:

Civilization: There are rules about the civilization, resources and building generation. Those are described below:

- The civilization must be established, and grow near a mine. In order to start mining a new mine, a Mill must be built near this new mine.
- Free mines will be guarded by enemies, in order to start getting gold from a new mine Knights need to kill them, so the Workers can start mining it.
- The Workers will be harvesting wood, and the trees will be cut off.
- The civilization will have a max food limit; this means how much Workers and Knights the civilization is able to feed. The Huts help improve the actual food limit (1 Hut = 10 food), but there is a top limit.

The game simulator is shown in Figure 4.

The user interaction interface has controls for:

- Max Food Limit.
- *Mines*: Maximum number of mines around the map and the maximum number of Workers that can be mining (a same mine).
- *Trees*: Maximum number of trees to be placed in the map.
- *Worker*: Gathering capacity, creation cost and initial number of Workers.
- *Knights*: Training cost, health points and strength.
- *Buildings*: Gold and lumber required to create different type of buildings.

3.1 Contract-Net implementations

Contract Net is a high-level protocol for communication among software agents (nodes) in a distributed problem solver that facilitates the task-sharing[2], that means It allows tasks to be distributed among a group of agents. For Contract Net there are two different types of agents, an Initiator and a Participant. At any time, any one agent can be an Initiator, a Participant or both [5]. Negotiation is implemented as two-way communication, in which an agent evaluates the offer of assigning a contract or receiving one from its own perspective [1], the protocol describes a number of basic messages that provide the necessary functionality for contract to be formed between agents.

Some of these messages are:

- **Task Announcement**: An agent does a call in which reports a task, with a complete specification of it, to be achieved.
- **Bid**: Agents, who listen had the call, decide for themselves if they are capable to perform or support the task and submit a tender.
- **Award**: Agent who did the call, choose between bids and selects who to award the contract.

The prototype implements four scenarios that apply Contract-Net:

Building construction team. In this scenario a Worker can request assistance from other(s) Worker, to be able to make buildings more quickly.

Knight's Teamwork. When a Knight finds a Dragon on the map, it asks for help to other Knights to attack the Dragon. Once the leader Knight estimates it is feasible to kill the Dragon, they go back and attack it.

Dragon Teamwork. When Dragons get to the Town, they ask other Dragons nearby to help them destroy the civilization.

Workers defend in the Huts. When Dragons attack the village, the Workers go inside the Hut and the defense of the village starts.

3.2 Knight decision-making

We calculate the utility of the Knights who fought against a Dragon using the following formula:

$$Utk = WHp \times Hp + Wstr \times Str - Wd \times d$$

In which:

- Utk: Utility of the group of Knights.
- WHp: Represents the Weight for Health Points
- HP: Health Points of the Knight.
- Wstr: Is the Weight to balance the strength of one Knight.
- Str: The current strength of a Knight
- Wd: The distance weight between the manager and potential contractor.
- d: The distance between the manager and potential contractor.

The use for this formula is shown in the Figure 5, with the following weights: $Whp = 0.5$, $Wstr = 0.7$, $Wd = 50$. The circles with letter K represent the KnightAgents, and the circles with letter D represents the DragonAgent. The gray circle represents the vision of the Knight. The results for this scenario were: Utility for $K_1 = 70$, Utility for $K_2 = -157$ and Utility for $K_3 = 34$. The order in which the Knight will choose the others is: K_1 , K_3 and K_2 .

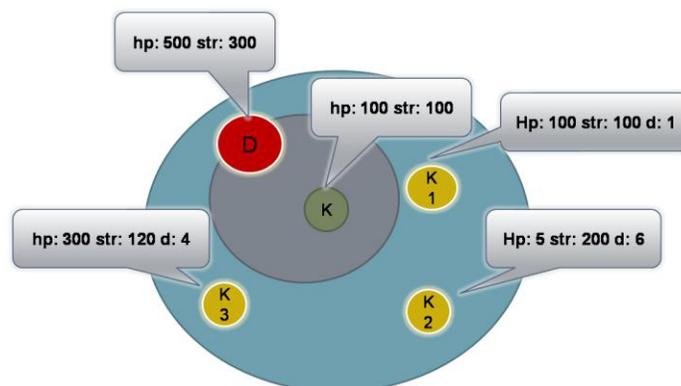


Figure 5: Selecting Knights.

However it is not sufficient to consider only the utility to attack the Dragon. To make better decisions is important to evaluate if the Knights who are in the contract are able to eliminate the Dragon. To resolve this problem, the following formulas were designed:

Number of attacks by Knight

$$Ak = \frac{Hp}{Hit_D}$$

Sum of Ak of all Knights.

$$S = \sum_{i=0}^n Ak_i$$

Total damage utility for Knights

$$DUK = \sum_{i=1}^n \{ [W_o Ak + W_p (S - Ak)] Hit_i \}$$

Damage Utility for Dragon.

$$DUD = \frac{Hp \times Hit_D}{\sum_{i=1}^n Hit_i}$$

Where:

- n is the number of total Knights that accept the contract.
- Hp: Help points
- Hit: Represents a percentage of the agent strength. If this percentage is higher, the damage generated by the agent will be higher.
- W_o : This factor can appreciate how important it is to calculate the chances of attack that has a knight.
- W_p : Indicates the probability that a knight can attack more times than allowed by their Health Point. If this value is closer to 1 the system assumes that the agent will have a greater opportunity to attack.

If $DUK > DUD$ then the Knights will attack, otherwise they run away.

4 Experiments and Results

The experiments compare the four types of Contract-Net strategy versus a random behavior of the agents. The first test does not include any Contract-Net, the behavior of the agents was random. As expected, the game consumes a lot of ticks and commonly the dragons win. The civilization can withstand using the TownHall and some of the knights that it can create. The results for this scenario are shown in the Table 2.

Table 2: Random Strategy

Result	Ticks	Spent Lumber	Lumber Harvested	Spent Gold	Gold Mined
Lose	40,371	2,880	3,052	12,950	11,980
Win	36,913	18,390	21,245	58,445	84,460
Lose	41,106	3,560	3,690	14,015	13,040
Lose	62,815	3,205	3,410	12,940	11,980
Lose	54,427	3,730	5,300	12,915	11,960

The second scenario uses a basic contract-net schema for dragons in which some dragons can create groups and attack the town. Using this strategy the dragons reduce the number of ticks in which they destroy the civilization. The results for this test are shown in the Table 3.

The last scenario adds three contract-net strategies for the civilization: Building construction team, Knight's Teamwork, and the defense of Workers in the Huts. This experiment shows the advantage for the civilization, the Knights can defeat the dragons every time and reduces the number of ticks. The knights can free some mines

and the workers can mine more gold. In one of the experiments the civilization kills the MightyDragon in only 3908, increasing its efficiency and winning all the games. The results for this scenario are shown in Table 4.

Table 3: Contract Net Dragon Team Work.

Result	Ticks	Spent Lumber	Lumber Harvested	Spent Gold	Gold Mined
Lose	16,399	2,710	2,784	12,960	11,980
Lose	25,384	3,675	4,182	13,285	12,300
Lose	23,588	2,205	2,210	12,940	11,950
Lose	27,472	3,380	3,680	14,470	13,480
Lose	29,310	3,630	4,992	12,975	12,020

Table 4: All Contract Net Strategies.

Result	Ticks	Spent Lumber	Lumber Harvested	Spent Gold	Gold Mined
Win	9638	3640	3661	11245	15790
Win	3908	2780	2831	7535	6920
Win	8849	3575	3662	10280	18750
Win	9958	3190	3381	9470	15820
Win	6779	3900	4280	11730	17030.

5 Conclusions

The application of methodologies for the development of Multi-Agent Systems requires an analysis focused on roles and protocols that seek to describe the activities and forms of interaction of agents that function as autonomous units with ability to interact and reach agreements. In contrast to the development schemes of traditional systems, this approach requires a radical shift in the process of analysis and design, which seeks to increase efficiency in the coordination of tasks.

For strategy games the achievement was to bring more intelligence to agents that perform the activity for which they were created, because of the use of techniques like Multi-Agent Systems, Contract-Net made possible that agents were able to increase efficiency, improve effectiveness and reduce significantly the time when they perform a specified task.

This work shows the basis of coordinated interaction in Multi-Agent Systems; however, it is necessary to increase the decision-making capabilities of the agents at the time of reaching an agreement with the goal of generating more intelligent strategy games and with a close connection with the reality.

References

- [1] Chad La Fournie (2003) Cooperation using Negotiations and the Contract-Net Protocol, Department of Computer Science, University of Calgary Canada.
- [2] Reid G. Smith (1980) Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver, IEEE Transaction on Computers Vol. c-29

- [3] Villarreal, P., Alesso, M., Rocco, S., Galli, M.R, and Chiotti, O. (2003). Approaches for the Analysis and Design of Multi-Agent Systems. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*. No.21, pp. 73-81.
- [4] Wooldridge, M., Jennings, N., Kinny, D. "The Gaia Methodology for Agent- Oriented Analysis and Design". *Journal of Autonomous Agents and Multi-Agent Systems*. Vol. 3, no 3, 2000.
- [5] Zafeer Alibhai (2003) What is Contract Net Interaction Protocol?, IRMS Laboratory, SFU
- [6] J. Madejski (2006) Survey of the Agent-Based Approach to Intelligent Manufacturing, Silesian University of Technology, Gliwice Poland.
- [7] Weiming Shen, Lihui Wang, Qi Hao, "Agent-Based Distributed Manufacturing Process Planning and Scheduling: A State-of-the-Art Survey". *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 36, No 4, July 2006.
- [8] Ruiqiang Lu, David Zhengwen Zhang (2010), "A Novel Agent Bidding Based Optimization Approach in Manufacturing Plannig and Scheduling", University of Exeter, Exeter United Kingdom.