TOWARDS A MULTI-DIMENSIONAL MODELLING OF COMPLEX SOCIAL SYSTEMS USING DATA MINING AND TYPE-2 NEURO-FUZZY SYSTEM: RELIGIOUS AFFILIATION CASE OF STUDY

Manuel Castañón–Puga, Carelia Gaxiola–Pacheco and Juan Ramón Castro
Facultad de Ciencias Químicas e Ingeniería
Universidad Autónoma de Baja California
Tijuana, Baja California, México.
Email: {puga,cgaxiola,jrcastro}@uabc.edu.mx

Dora–Luz Flores
Facultad de Ingeniería, Arquitectura y Diseño
Universidad Autónoma de Baja California
Ensenada, Baja California, México.
Email: dflores@uabc.edu.mx

Ramiro Jaimes–Martínez
Instituto de Investigaciones Históricas
Universidad Autónoma de Baja California
Tijuana, Baja California, México.
Email: rjaimes@uabc.edu.mx

KEYWORDS

FUZZY AGENTS, DATA MINING, SOCIAL COMPLEXITY, RELIGION AFFILIATION

ABSTRACT

The purpose of this paper is, to describe a work-in process for application of distributed agency methodology to multi-dimensional preference model into a complex social system. This paper shows a study case focused on a modeling system for decision-making on cognitive structure religious affiliation preference. A type-2 neuro-fuzzy approach is used to configure cognitive rules into an agent to build a multi-agent model for social simulation.

INTRODUCTION

The social systems are complex entities that represent a whole that cannot be understood by looking at its parts independently Yolles (2006). Another characteristic is the interdependence of the parts conforming the whole: a change to one of the components in the system may potentially affect all others. Boulding (1956).

The main goal of this part of our research is to develop a computational model of change in religious affiliation that incorporates available mathematical and computational theories that have not been appropriately considered in models of complex social phenomena. Even though applications of Multi-Agent Systems (MAS) have been developed for the social sciences, MAS have been widely considered in some areas such as Artificial Intelligence (AI) Gilbert (2007). The state-of-the-art in computational capabilities has been incorporated in multiple areas Russell and Norvig (2004), particularly as it refers to distributed systems and distributed agencies López et al. (2002).

Previous and Related Work

There are some related work on oriented simulation methodology for modelling complex problems. In Suarez et al. (2008), is proposed "fuzzy agents" to represent agencies that can not be modelled with specific actors. Then, in Suarez and Castanon-Puga (2010) is proposed "distributed agency" as a simulation language for describing social phenomena.

In Suarez et al. (2010) shows an approach to decision-making system based on type-2 fuzzy inference system in order to implement fuzziness on cognitive agents. This model introduces subjectivity and uncertainty in agent perceptions and preferences levels, were decision-making system are influenced by different agencies. Preferences for choosing one option competing with each other, but also were influenced by endogenous and exogenous variables that could affect the levels of uncertainty and, therefore, affect the way the agents interpret the messages.

In Márquez et al. (2011b), a methodology is proposed based on neuro-fuzzy technique, where a data mining procedure based on a neural network are applied to a real data to configure a type-2 fuzzy inference system into an agent.

Some other important work related to these approaches is presented in Mendel and Wu (2010) where we find an example of the application of type-2 fuzzy logic to model a subjective decision-making system or perception.

Data Mining and Neuro–Fuzzy System

An Interval Type-2 Fuzzy Neural Network (IT2FNN) are used for automatically generate the necessary rules. The phase of data mining using Interval Type-2 Fuzzy Logic Systems (IT2FLS) Castillo et al. (2010); Castro et al. (2010) becomes complicated, as there are enough rules to determine which variables one should take into account. The search method of back-propagation and hybrid learning (BP+RLS) is more efficient in other methods, such as

Proceedings 26th European Conference on Modelling and Simulation ©ECMS Klaus G. Troitzsch, Michael Möhring, Ulf Lotzmann (Editors)
ISBN: 978-0-9564944-4-3 / ISBN: 978-0-9564944-5-0 (CD)

genetic algorithms Rantala and Koivisto (2002); Castro et al. (2008).

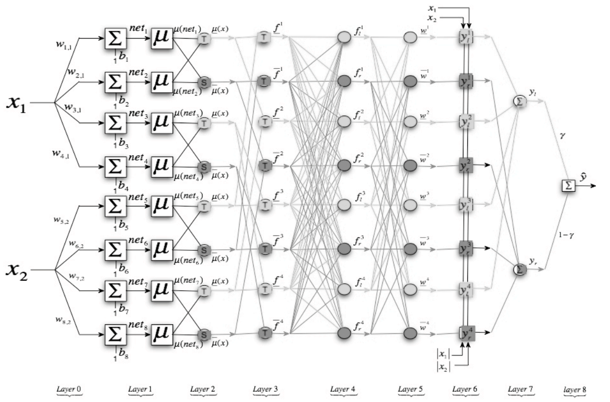

Figure 1: Generation of the necessary rules using an Interval Type–2 Fuzzy Neural Network (IT2FNN).

Since the IT2FNN method seems to produce more accurate models with fewer rules is widely used as a numerical method to minimize an objective function in a multidimensional space, find the approximate global optimal solution to a problem with N variables, which minimize the function varies smoothly Stefanescu (2007).

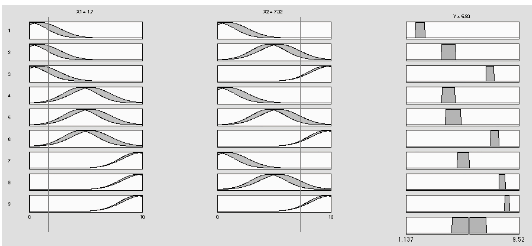

Figure 2: Rules on a Type-2 Fuzzy Inference System.

With the application of this grouping algorithm we obtain the rules, the agent receives input data from its environment and choose an action in an autonomous and flexible way to fulfill its function Peng et al. (2008).

Religious Affiliation

When literature talks about of religious change, usually refers to the attachment or religion affiliation Ortiz (2006). Although some authors have argued that the concept can not be limited to this dimension, membership is one of the most important variables to study the religious phenomena Fortuny (1999).

The religious field is conformed by several dynamics systems. For example, we can identify some organizational entities:

- institutional
- socio-demographic groups
- individual

Within these multiple dimensions interrelated complex processes are occurring, such changes of allegiance, change in commitment and participation, socialization and subjectivity of standards (through doctrines, values, practices), reformulation and affirming traditions. These multiple dimensions shape the religious field, and generically is known as religious change.

Religious Affiliation in México

In México, religious affiliation has undergone major changes since the 1950's until today. Based on population censuses, the growth rates of the evangelical population has been higher than the total Catholic population [1] Jaimes-Martínez (2007).

Baja California has one of the percentages of highest evangelical population of Northern states [2].

CASE OF STUDY

Tijuana is a city located in north-western of México. Belongs to the state of Baja California, and is one of the fastest growing city in the country due to high migration rates. The population is mainly composed by migrants from southern of the country. They came to the border to further job opportunities, or looking to migrate to the United States, staying in the city long time.

Tijuana's multi-cultural and religious complexity

Tijuana is an example of social and religious change. Its boundary condition has been one factor that has become a city in full development and expansion, not only by the strength of the Southern California economy, but by the early efforts to boost manufacturing by the federal government.

These factors, combined with growing internal and international migration, have transformed a town of Tijuana from a town with 12,181 inhabitants in 1930 to one with 1.2 million in 2000[3] Alegría and Ordónez (2002). It was so from NAFTA, Tijuana was consolidated as a major call centres maquiladora industry, with an evident increase in employment and production, but not productivity or living standards and welfare Arias (2008).

According to some authors, the economic balance, social and cultural development of these global processes, regional and local has had complex effects on Tijuana's society, where stands the reconfiguration of identities and new forms of social and cultural reproduction Jiménez (2006).

[1] The evangelical population has experienced rates of 8.90, between 1970 and 1980, while the total population was 3.16. Although at present growth rate 2.46 points, it is still higher than that of the population is Catholic and total population.

[2] Baja California has 7.90% and evangelical population, surpassed only by one of the first entities to which the Protestant missionaries arrived in the nineteenth century, Tamaulipas, to 8.65%. Nationally, the percentage of evangelicals is 5.20%.

[3] Tito Alegría and Gerardo Ordóñez consider the growth process of Tijuana covers from 1930 to 2000, thanks mainly to the economic expansion of Southern California.

In this sense, the religious sphere in Tijuana has a great religious diversification as a result of different waves of migration that have shaped their society. Therefore, religious affiliation is also an indicator to study these processes of reconfiguration and realignment[4] Jaimes-Martínez (2007).

Preference for religious affiliation in Tijuana

The city has a great diversity of faiths and religious traditions. Although more numerous the Christian (Catholic, Protestant, evangelical non-biblical), there are Buddhists, Muslims, Jews and a variety of groups and beliefs generically known as New Age[5].

Considering this, we can say that every group or social stratum in Tijuana has a wide range of choice, or affinity, in the religious field in the city. Each of them is not only an expression of traditions, customs and religious practices of different groups have brought to Tijuana from their places of origin, but the dynamic formulation of these beliefs in the new environment.

Importance of modelling Tijuana habitants religious affiliation

Qualitative tools are used mainly by social scientists to analyze religious change. This preference is justified by explanatory depth they are able to work on the meaning of the action of the subject. However, its explanatory power is limited to the subjects observed, without looking at a larger scale or level, for which we assume a similar behavior.

For larger scales of analysis is necessary to use quantitative tools, which have a greater ability to generalize, but is constrained to observe the direction and motivation for action. Therefore, tools such as computational modeling are an alternative to unite both fields.

METHODOLOGY FOR THE MODELLING OF COMPLEX SOCIAL SYSTEM USING NEURO-FUZZY AND DISTRIBUTED AGENCIES

To build the model of change of religious affiliation will follow the distributed agency methodological steps Márquez et al. (2011a):

1. Determining the levels of agency and their implicit relationships
2. Data mining
3. Generating a rule-set
4. Multi-Agent Modelling (Implementation on a agent based simulation tool)
5. Validating the model
6. A simulation and optimization experiment
7. Analysing the outputs

Although the methodology covers the entire life-cycle of a research process, on this paper we are describing the data mining and generating rule set steps. We are focused on the neuro-fuzzy approach in order to set up a rule set into agents.

MODELLING TIJUANA CITY

The principal difference between MAS and our proposed approach is that in our methodology the space includes transformations performed by a higher level of agency. This upper-level agent is composed of lower-level subcomponents the may enjoy agency in their own right. It is the responsibility of this intermediate agent to present its subcomponents with individual phase-spaces that are tailored to induce the desired behaviour from the lower-level agents which inhabit it, when it chooses according to its own objective function.

Therefore, for our proposed work-in-progress case study, if we consider a municipality an agent, this upper-level agent is composed by subcomponents, which in our case study of the city of Tijuana, Mexico, will be represented by the AGEBS that compose this city. AGEBS is the terminology used to describe the different areas of the city that are in turn are composed of neighbourhoods.

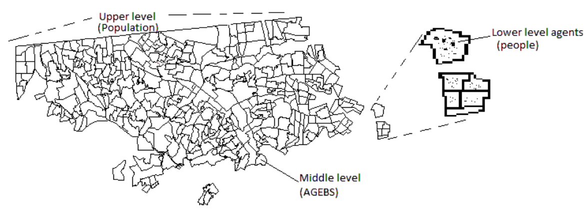

Figure 3: Levels of agents represented on the social system.

Levels of agency

In our study we use three levels of agency: the upper-level agent is represented by the religion groups or organizations, the intermediate agents are represented by the AGEBS and the lower level agents are the individual inhabitants of the city.

Using a recent census of the reunion sites distribution of the different religious organizations operating in the city, we know the exact places where they carry out activities of proselytizing. This information gives us hints of the influence of the presence of organizations in its environment and its impact on socio-demographic variables.

We are looking for relationships between demographic and economic factors (subtracted from AGEBS) and distribution of meeting places of religious organizations. We believe that factors such as poverty, marginalization

[4] Between 1990 and 2000 Tijuana just recorded a growth rate of 8.94 evangelical population, while at the national level was 2.46.

[5] Syncretic movements oriental religions such as Buddhism, introducing ideas of self-motivation, personal growth, alternative medicine, psychology, etc.

and other characteristics related to socio-demographic issues influencing the decision-making system of individuals in a complex and distributed way. Similarly, religious organizations act as agents who are influenced by other agencies distributed.

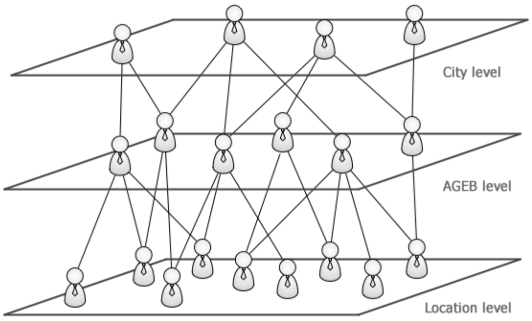

Figure 4: Multiple layers of agents represented on the social system.

Data sets

In the particular case of the city of Tijuana, the data set used came from the Instituto Nacional de Estadistica y Geografia (INEGI), the Mexican governmental organization in charge of gathering data at a federal level including aspects that are geographical, socio-demographic and economical. The data set of the city of Tijuana is divided into 363 areas, known as AGEBS. "The urban AGEB encompass a part or the totality of a community with a population of 2500 inhabitants or more... in sets that generally are distributed in 25 to 50 blocks"INEGI (2006).

The data sets for this case study were originally compiled in an information system that is intrinsically geographical. These systems helped in the generation, classification and formatting of the required data—a fact which facilitates the edition of the different thematic layers of information, in which one can quantify the spatial structure to visualize and interpret the areas and different spatial patterns in Tijuana.

For this paper, we going to use de following variables to exemplify the proposed approach using information from 2010 population census in México INEGI (2010).

- P15YMAS = Population over 15 years old on locations.
- P15YMSE = Population over 15 years old on locations without education.
- GRAPROES = Education.
- PEA = Working population.
- PEINAC = Non working population.
- PCATOLICA = Catholic population on locations.
- PNCATOLICA = Non catholic population on locations.

Neuro-fuzzy inference system

Using the neuro-fuzzy system for the automatic generation of rules, this phase of the data extraction from the data may become complicated, as the process needs to appropriately establish the number of sufficient norms and variables that the study needs to take into account.

Using this grouping algorithm, we obtain the appropriate rule-set assigned to each agent representing an AGEB or a inhabitant of it, the agent receives inputs from its geographical environment and in turn much choose an action in an autonomous and flexible fashion Gilbert (2007); Drennan (2005); Wooldridge and Jennings (1995).

The purpose of this structure without central control is to garner agents that are created with the least amount of exogenous rules and to observe the behavior of the global system through the interactions of its existing interactions, such that the system, by itself, generates an intelligent behavior that is not necessarily planned in advance or defined within the agents themselves; in other words, creating a system with truly emergent behavior Botti and Julian (2003); Russell and Norvig (2004).

From the 2010 census information, we create a Type-2 Fuzzy Inference System as how we could represent different agencies as a decision-making system into agents.

City level Type-2 Fuzzy Inference Systems

The figure 5 shows a type-2 fuzzy inference system for Tijuana city. It depicts a set of input-output variables and a rule set. Output variables are catholic and non-catholic as a response of the system. We could use the difference between both values to make decisions into an agent as a preference decision-making system.

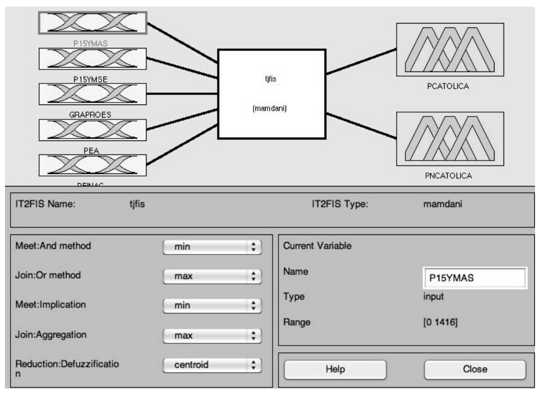

Figure 5: Fuzzy Inference System for Tijuana City.

The Figure 6 depicts the resolution example of the rules by the fuzzy inference system. Different quantitative input values could be introduced and the system resolve creating different responses. Depending of the combination of inputs, we can expect different responses of the system. An agent will use this inference system as a decision-making system to show different behaviours depending of the situation.

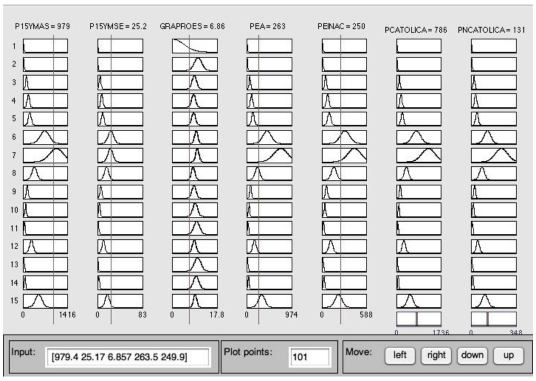

Figure 6: Fuzzy Inference System Rule Set Evaluation for Tijuana City.

The Figure 7 represents the response of the system to catholic preference, and the Figure 8 for non-catholic preference. We can see that there are response differences, so we can us it to make decisions.

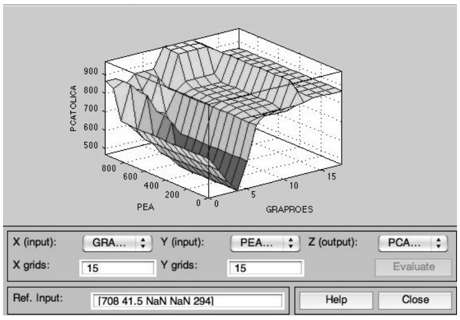

Figure 7: PEA vs. GRAPROES Type-Reduced Surface View for Tijuana City PCATOLICA output.

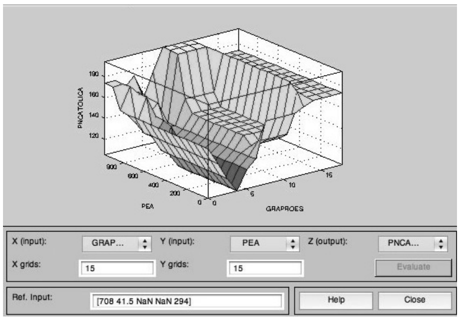

Figure 8: PEA vs. GRAPROES Type-Reduced Surface View for Tijuana City PNCATOLICA output.

Distributed agents do not necessarily define agents in lower-levels of description, but rather consider all levels of agency that are interconnected in a type of organism that spreads throughout the system Suarez et al. (2008); Suarez and Castanon-Puga (2010); Gilbert (2007).

AGEB level Type-2 Fuzzy Inference Systems

On AGEB layer, we can build fuzzy inference systems for agents that represents locations. Figure 9 and Figure 10 depicts the FIS response for different locations into the same AGEB. As we can see, there are differences between AGEB agents. At this level, we could be representing AGEB agents into a city context.

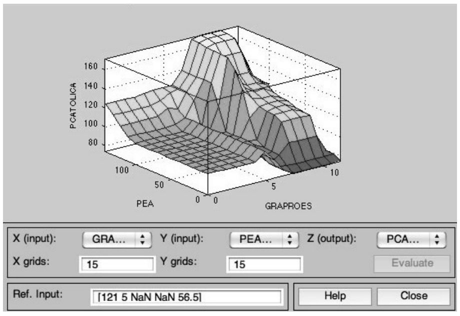

Figure 9: PEA vs. GRAPROES Type-Reduced Surface View for AGEB 32 PCATOLICA output.

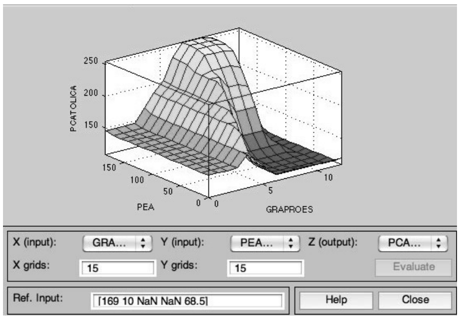

Figure 10: PEA vs. GRAPROES Type-Reduced Surface View for AGEB 51 PCATOLICA output.

Location level Type-2 Fuzzy Inference Systems

On location layer, we can build fuzzy inference systems for agents that represents locations. Figure 11 and Figure 12 depicts the FIS response for different locations into the same AGEB. As we can see, there are differences between locations agents. At this level, we could be representing locations agents into a AGEB context.

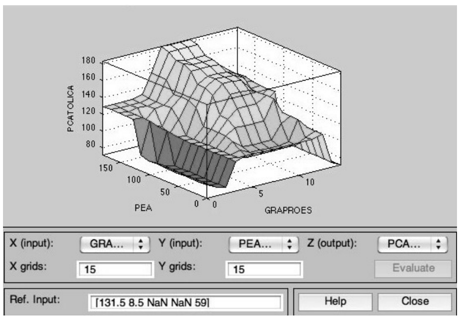

Figure 11: PEA vs. GRAPROES Type-Reduced Surface View for Location 187 PCATOLICA output.

Testing Religion Affiliation Type-2 Fuzzy Inference System in NetLogo

With the use of a NetLogo extension of a type-2 fuzzy inference machine, we can test the fuzzy inference sys-

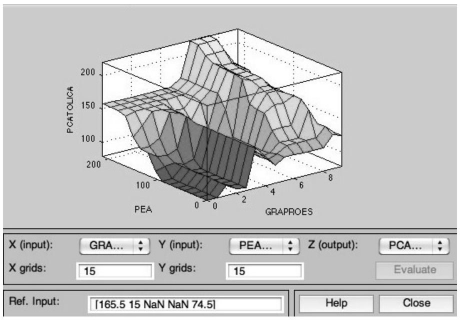

Figure 12: PEA vs. GRAPROES Type-Reduced Surface View for Location 283 PCATOLICA output.

tem for catholic population and the non-catholic population of Tijuana city. Figure 13 shows an screen shot of the test in NetLogo with the extension mentioned before. The test include a set of 100 agents with their initial attributes. The test consist in the evaluation of all agents inside of a type-2 fuzzy inference system, representative of each city. The result was a different response of each agent, according of the city in wich was evaluated.

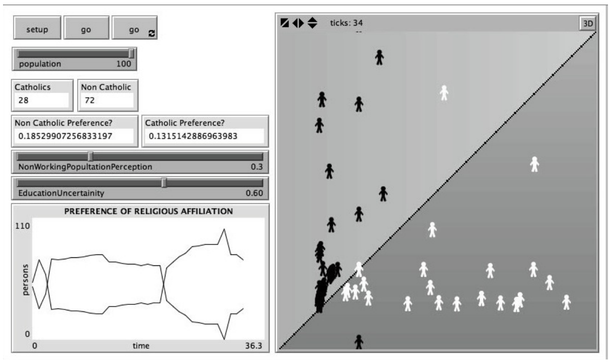

Figure 13: Screen Shot of Religion Affiliation Type-2 Fuzzy Inference System in NetLogo test prototipe.

The test evaluates variations in the limits of some input variables' membership functions, simulating a perception variation of the individuals in their respective context (their city). The result of this is a variation in the choice of religious afiliation in response of variations of the perception values.

Finally, the test evaluates changes in uncertainty values of input variables' membership functions, simulating a variation on the security of individuals regarding the risk or doubt. The result is that if we make changes in the uncertainty values, the response is a change in the choice of religious affiliation.

CONCLUSIONS

The development of a general methodology for the description and analysis of complex systems remains an open research task for computer scientists and all other scholars interested in the subject matter. Naturally, such a vast language will imply the use of different techniques and theories, given that the perspectives of each scientific discipline-and ultimately of each researcher-may vary greatly.

The religious affiliation can be modelled with Fuzzy Inference Systems. Establishing different layers of interaction between agents and analysing their influence on decision-making system of agents in each level, we can represent the complexity of the phenomenon of individual preference to a religious affiliation.

We use the case study of the city of Tijuana, as it has an updated census of the distribution of meeting places of religious organizations in the city and their respective socio-demographic information.

We use a neural-fuzzy approach to develop a computational model of the decision-making system of agents in order to build a multi-agent system. We represent different levels of agency with different cognitive agents. Each agent in the system are a fuzzy cognitive agent that can choose religion options based on preferences.

ACKNOWLEDGEMENTS

We would like to thank the Mexican National Council for Science and Technology (Consejo Nacional de Ciencia y Tecnología, CONACYT) and Universidad Autónoma de Baja California for the economic support granted for this research.

REFERENCES

Alegría, T. and Ordónez, G. (2002). Regularización de la tenencia de la tierra y consolidación urbana en tijuana, b.c. Research report, El Colegio de la Frontera Norte, México.

Arias, A. L. (2008). Cambio regional del empleo y productividad manufacturera en méxico, el caso de la frontera y las grandes ciudades,1970-2004. *Frontera Norte*, 20(40):79–103.

Botti, V. and Julian, V. (2003). Estudio de métodos de desarrollo de sistemas multiagente. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 18.

Boulding, K. (1956). General systems theory the skeleton of science. *Management Science*, 6(3):127–139.

Castillo, O., Melin, P., and Castro, J. R. (2010). Computational intelligence software for interval type–2 fuzzy logic. *Journal Computer Applications in Engineering Education*.

Castro, J. R., Castillo, O., Melin, P., Mendoza, O., and Rodríguez-Díaz, A. (2010). An interval type–2 fuzzy neural network for chaotic time series prediction with cross–validation and akaike test. *Soft Computing for Intelligent Control and Mobile Robotics*, 318:269–285.

Castro, J. R., Castillo, O., Melin, P., and Rodríguez-Díaz, A. (2008). A hybrid learning algorithm for a class of interval type–2 fuzzy neural networks. *Journal of Information Sciences*, 179(13):2175–2193.

Drennan, M. (2005). The human science of simulation: a robust hermeneutics for artificial societies. *Journal of Artificial Societies and Social Simulation*, 8(1).

Fortuny, P. (1999). *Creyentes y creencias en Guadalajara*. CIESAS, México.

Gilbert, N. (2007). *Computational social science: Agent-based social simulation*, pages 115–134. Bardwell, Oxford.

INEGI (2006). Segundo conteo de población y vivienda 2005. instituto nacional de estadística geografía e informática.

INEGI (2010). Censo de población y vivienda 2010. instituto nacional de estadística geografía e informática.

Jaimes-Martínez, R. (2007). *La paradoja neopentecostal. Una expresión del cambio religioso fronterizo en Tijuana, Baja California*. PhD thesis, El Colegio de la Frontera Norte, México.

Jiménez, G. (2006). Cultura, identidad y metropolitanismo global. Online.

López, A., Hernández, C., Pajares, J., and Aguilera, A. (2002). Sistemas multiagente en ingeniería de organización. técnicas computacionales de simulación de sistemas complejos.

Márquez, B. Y., Castañón Puga, M., Castro, J. R., and Suarez, E. D. (2011a). Methodology for the Modeling of Complex Social System Using neuro-Fuzzy and Distributed Agencies. *Journal of Selected Areas in Software Engineering (JSSE)*, March:1–8.

Márquez, B. Y., Castanon-Puga, M., Castro, J. R., and Suarez, E. D. (2011b). A distributed agency methodology applied to complex social system, a multi-dimensional approach. In Zhang, R., Cordeiro, J., Li, X., Zhang, Z., and Zhang, J., editors, *Proceedings of the 13th International Conference on Enterprise Information Systems, Volume 1, Beijing, China, 8-11 June, 2011.*, pages 204–209. SciTe Press.

Mendel, J. M. and Wu, D. (2010). *Perceptual Computing, Aiding People in Making Subjective Judgments*. IEEPress.

Ortiz, O. O. (2006). Cambio religioso en la frontera norte. aportes al estudio de la migración y las relaciones transfronterizas como factores de cambio. *Frontera Norte*, 18(35):111–134.

Peng, Y., Kou, G., Shi, Y., and Chen, Z. (2008). A descriptive framework for the field of data mining and knowledge discovery. *International Journal of Information Technology and Decision Making*, 7:639–682.

Rantala, J. and Koivisto, H. (2002). Optimised subtractive clustering for neuro-fuzzy models.

Russell, S. and Norvig, P. (2004). *Inteligencia Artificial. Un Enfoque Moderno*. Pearson Prentice Hall, 2 edition.

Stefanescu, S. (2007). Applying nelder mead's optimization algorithm for multiple global minima. *Romanian Journal of Economic Forecasting.*, pages 97–103.

Suarez, E. D. and Castanon-Puga, M. (2010). Distributed agency, a simulation language for describing social phenomena. In *IV Edition of Epistemological Perspectives on Simulation*, Hamburg, Germany. The European Social Simulation Association.

Suarez, E. D., Castanon-Puga, M., Flores, D.-L., Rodriguez-Diaz, A., Castro, J. R., Gaxiola-Pacheco, C., and Gonzalez-Fuentes, M. (2010). A multi-layered agency analysis of voting models. In *The 3rd World Congress on Social Simulation WCSS2010*, Kassel, Germany.

Suarez, E. D., Rodriguez-Diaz, A., and Castanon-Puga, M. (2008). *Soft Computing for Hybrid Intelligent Systems*, volume 154, chapter Fuzzy Agents, pages 269–293. Springer.

Wooldridge, M. and Jennings, N. (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review*.

Yolles, M. (2006). Organizations as complex systems, an introduction to knowledge cybernetics.

AUTHOR BIOGRAPHIES

MANUEL CASTAÑÓN PUGA was born in Tijuana, Baja Califonria, México. Dr. Castañón-Puga is a full time Professor at Chemistry Sciences and Engineering Faculty of Baja California Autónomus University. Currently is leader of the "Complexity and computation" research group. His email is `puga@uabc.edu.mx` and his personal webpage at `http://cyc.tij.uabc.mx/puga`.

CARELIA GAXIOLA PACHECO was born in Sinaloa, México. Dr. Gaxiola-Pacheco is a full time Professor at Chemistry Sciences and Engineering Faculty of Baja California Autónomus University. Currently is member of the "Complexity and computation" research group. Her email is `cgaxiola@uabc.edu.mx` and her personal webpage at `http://cyc.tij.uabc.mx/gaxiola`.

DORA-LUZ FLORES was born in Tijuana, Baja California, México. Dr. Flores is a full time Professor at Engineering, Architecture and Design Faculty of Baja California Autónomus University. Currently is member of the "Complexity and computation" research group. Her email is `dflores@uabc.edu.mx` and her personal webpage at `http://cyc.tij.uabc.mx/flores`.

RAMIRO JAIMES MARTÍNEZ was born in México. Dr. Jaimes-Martínez is a full time Researcher at Historical Research Institute of Baja California Autónomus University. Currently is leader of the "History and society" research group. His email is `rjaimes@uabc.edu.mx`

JUAN RAMÓN CASTRO was born in Tijuana, Baja California, México. Dr. Castro is a full time Professor at Chemistry Sciences and Engineering Faculty of Baja California Autónomus University. Currently is member of the "Computational intelligence" research group. His email is`jrcastro@uabc.edu.mx`