Computational framework for modeling the dynamic evolution of large-scale multi-agent organizations

Alina Lazar, Robert G. Reynolds

Artificial Intelligence Laboratory, Wayne State University, Detroit, MI 48202

ABSTRACT

The process by which complex social entities such as the state emerged from lower level structures and other supporting economies has long been of prime interest to anthropologists and other disciplines as well. This is because the emergence of such a social structure can have a profound impact on the societies' physical and social environment. However, the task of developing realistic computational models that aid in the understanding and explanation of state emergence has been a difficult one. The goal of this project is to produce a large-scale knowledge-based computational model of the origins of the Zapotec State centered at Monte Alban, in the Valley of Oaxaca, Mexico.

Contact:
Alina Lazarr
Email: all@cs.wayne.edu

Robert Reynolds
Email: SAVATHOME@email.msn.com

Computational framework for modeling the dynamic evolution of large-scale multi-agent organizations

Alina Lazar, Robert G. Reynolds

The results of the data mining (Weiss 1998) process can be used in many different ways. Therefore, the form of the knowledge collected will have a major impact on the efficiency and effectiveness of its use in a given application. In this paper we examine the problem of extracting knowledge for use by agents in a large-scale multi-agent system (Russell 1995). Here, the knowledge is ontological knowledge that represents constraints that the physical and social environments placed upon the agents and their interactions. The ontological knowledge represents the semantic building blocks around which the world models are formed. For an agent in a particular model, only the things in his ontology can exist and it cannot perceive things that are not represented in the ontology. An *ontology* (Fox et al 1998) is a basic level of knowledge representation scheme, a formal definition of entities and their properties, interactions, behaviors and constraints. Each agents' decisions need to be checked against these constraints prior to their execution. In a complex multi-agent system, hundreds of thousands of agents may need to check these constraints regularly which means that a successful data mining activity will need to produce a relatively small set of syntactically simple rules for the process to be efficient. Fox et al (1998) have used data mining techniques to produce corporate ontogenies.

There are several factors that can influence the nature of the ontological constraints that are produced. First, the nature of the data collection and measurement process and the uncertainty induced into the data set by the presence of noise. Second, the nature of the representation used to express the extracted patterns; e.q. whether it allows for uncertainty or not. Third, the data mining technique employed and the assumptions that it makes about the collected data. Fourth, how these constraints will be stored, accessed, and used by the agents involved.

For a given data set one can compare the different data mining techniques in terms of the syntactic and semantics of the induced constraints. In this application we are interested in simulating the emergence of the archaic state in the Valley of Oaxaca, Mexico. A state is among the most sophisticated and powerful structures that has emerged from the social evolution process. In the modern world these are termed "nation states' with a government composed of a hierarchical decision-making structure where the decision-makers are either elected or appointed. States are supported by various economies and are able to interact with each other via warfare, trade, etc . Most states in the ancient world-often called archaic states-were ruled by hereditary royal families. These archaic states exhibited much internal diversity with populations numbering from tens of thousands to millions. They had a bureaucracy, organized religion, a military presence, large urban centers, public buildings, public works, and services provided by various professional specialists. The state itself could enter into warfare and trade-based relationships with other states and less complex neighbors.

The process by which complex social entities such as the state emerged from lower level structures and other supporting economies has long been of prime interest to anthropologists and other disciplines as well. This is because the emergence of such a social structure can have a profound impact on the societies' physical and social environment. However, the task of developing realistic computational models that aid in the understanding and explanation of state emergence has been a difficult one. This is the result of two basic factors:

1. The process of state formation inherently takes place on a variety of temporal and spatial scales. The emergence of hierarchical decision-making can be viewed as an adaptation that allows decision-makers to specialize their decisions to particular spatial and temporal scales.
2. The formation of the state is a complex process that is fundamentally directed by the social variables but requiring dynamic interaction between the emergent system and its environment. Identifying the nature of these interactions is one of the reasons why the process of state formation is of such interest.

The goal of this project is to produce a large-scale knowledge-based computational model of the origins of the Zapotec State (Flannery 1996), centered at Monte Alban, in the Valley of Oaxaca, Mexico. State formation took place between 1400 B.C. and 300 B.C. While archaic states have emerged in various parts of the world, the relative isolation of the valley allowed the processes of social evolution to be more visible there. Extensive surveys (Blanton 1989, Blanton, Kowalewski 1982, Kowalewski 1989) of the 2100 square kilometer valley, were undertaken by the Oaxaca Settlement Pattern Project in the 1970's and 1980's. The location and features of over 2,700 sites dating from the archaic period (8000 B.C.) to Late Monte Alban V (just prior to the arrival of the Spaniards) were documented.

Several hundred variables were recorded for each site. In addition, they surveyed the 6.5 square kilometer urban center of Monte Alban, a site that contained over 2,000 residential terraces. This site was the focus for early state formation in the valley.

Both surveys provided the knowledge needed to create our multi-agent simulation model. We then produced a spatial temporal database that contained the results of both surveys and used data mining techniques from Artificial Intelligence (Russell 1995) to produce knowledge about site location, warfare, trade, and economic decisions to be used for the construction of the multi-agent model. However, in order to do this we needed to add more data about the spatial and temporal context to both the regional and urban center surveys. Specifically, we had to add variables that allowed us to locate each site spatially and temporally to a level of precision consistent with the scale of our simulation. For example, temporal periods are characterized by the presence of pottery of different styles. That data was available only in text form. All of this pottery data, over 130 variables for each residential terrace, was scanned into the computer, corrected for errors, and added to the Monte Alban data set. This data allowed us to identify the periods that each terrace was occupied. Pottery data was also integrated into the regional data set.

In addition, the survey had produced hundreds of pages of hand drawn maps for both the Monte Alban and regional surveys that contained the spatial context for the location of each site. Since our goal was to ask specific questions about the spatial and temporal context we needed to tie each site into its mapped location. We then proceeded to digitize each of the maps and to associate each site object with its corresponding data record. This allowed us to produce a geographical information system (GIS) that serves as our “virtual valley of Oaxaca”. This acts as a vehicle for our data mining activities and as a knowledge base for the multi-agent simulation and allows the results of the simulation to be displayed and compared with the actual data in a spatial context. It is envisioned that the resultant GIS system will be a useful tool for researchers and students from various fields to study the emergence of complexity in the future.

In order to perform the data mining activities, we extended traditional data mining techniques and developed new ones in order to deal with the complexities inherent in the Oaxaca database. At the regional level we used Utgoff’s incremental decision tree algorithm (IDTI) (Utgoff 1989) to generate the Decision Trees for each region and phase of the valley. The approach was used to generate decision trees that discriminated between sites that were targets for warfare and those that were not for a given period (Reynolds 1999). However, given the many disperate steps under which the data was collected and organized it was felt that perhaps some improvements might be made by using a technique that took into account the presence of uncertainty in the data -especially in regions and periods when the social and settlement patterns were complex and prone to data collection error.

Since the majority of the data was discrete rather than continuous in nature we selected rough sets as a vehicle for representing uncertainty here. We employed an evolutionary technique, Genetic Algorithms, to control the search in this case because Genetic Algorithms had been successfully used with Rough Sets previously. The decision systems or rule sets produced by both approaches were then compared in terms of their ability to decide about the location of sites that are targets for warfare in this period. We then compared the two approaches over all relevant phases of social evolution in the valley in terms of their syntactic structure and complexity..

Next the semantic impact of the two different rule sets on the evolution of complex social structures in the valley was tested with a multi-agent simulation based upon the model of Marcas and Flannery (1997). A multi-agent system model of the origins of an archaic state was developed. In the model agent interaction is mediated by a collection of rules where the rules can pertain to various types of interaction such as trade, warfare, and ritual.

Our goal is to run the simulation with the two different set of warfare rules to determine the semantic differences that each places on the emergence of complexity in the valley. Social networks are produced as the result of each run and the extent to which a valley wide control network emerges, a signature of state formation, is determined. Different levels of warfare can be generated using each set of rules. The results suggest that in terms of the model of Marcus and Flannery, the observed rate of emergence of social complexity in the real world can only be achieved by allowing for a sufficient amount of warfare among agents. Also, the patterns produced by the rough set approach were more likely to exhibit the valley wide networks than the crisp set approach using decision trees.

REFERENCES

Ågotnes T (1999, February). Filtering Large Propositional Rule Sets While Retaining Classifier Performance, Master's thesis, Norwegian University of Science and Technology.

Ågotnes T, Komorowski J and Øhrn A (1999, September). Finding High Performance Subsets of Induced Rule Sets: Extended Summary. *Proceedings Seventh European Congress on Inteligent Techniques and Soft Computing (EUFIT'99), Aachen, Germany* (H. J. Zimmermann and K. Lieven, eds.).

Blanton RE (1989). *Monte Albán Settlement Patterns at the Ancient Zapotec Capital*. Academic Press.

Blanton R E, Kowalewski S, Feinman G & Appel J (1982). *Monte Albán's Hinterland, Part I, the Prehispanic Settlement Patterns of the Central and Southern Parts of the Valley of Oaxaca, Mexico*. The Regents of the Univerity of Michigan, The Museum of Anthropology.

Crowston K (1994) Evolving Novel Organizational Forms. *Computational Organization Theory*. Carley KM, Prietula MJ, eds. Lawrence Erlbaum Associates Publisher, Hillsdale, New Jersey.

Marcus J, Flannery KV, 1996, *ZAPOTEC CIVILIZATION – How Urban Societies Evolved in Mexico's Oaxaca Valley* Thames and Hudson Ltd, London.

Fogel DB (1995). *Evolutionary Computation - Toward a New Philosophy of Machine Learning*. IEEE Press.

Fox M, Barbuceanu M, Gruninger M & Lin J (1998). An Organizational Ontology for Enterprise Modeling. *Simulation Organizations Computational – Models of Institutions and Groups.* Prietula, MJ, Carley KM, Grasser, L eds. AAAI Press/ MIT Press Menlo Park, California, Cambridge Massachusetts.

Goldberg, D. E. (1989, January). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company, Inc.

Holland JH (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Komorowski J & Øhrn A (1999) Modelling Prognostic Power of Cardiac Tests Using Rough Sets. *Artificial Intelligence in Medicine*, Vol. 15, No. 2, pp. 167-191.

Kowalewski SA, Feinman, GM, Finsten L, Blanton RE & Nicholas LM (1989) *Monte Albán's Hinterland, Part II, Prehispanic Settlement Patterns in Tlacolula, Etla, and Ocotlan, the Valley of Oaxaca, Mexico*. Vol. 1. The Regents of the University of Michigan, The Museum of Anthropology.

Lazar A & Sethi IK (1999, November). Decision Rule Extraction from Trained Neural Networks Using Rough Sets. *Intelligent Engineering Systems Through Artificial Neural Networks* (CH Dagli, AL Buczak, & J Ghosh, eds.), Vol. 9, (New York, NY), pp. 493-498, ASME Press.

Øhrn A (1999, December). *Discernibility and Rough Sets in Medicine: Tools and Applications*. PhD thesis, Norwegian University of Science and Technology, Department of Computer and Information Science.

Øhrn A (2000). The Rosetta C++ Library: Overview of Files and Classes. Tech. Rep., Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

Øhrn A (2000). Rosetta Technical Reference Manual. Tech. Rep., Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

Øhrn A & Komorowski J (1997, March). Rosetta - a Rough Set Toolkit for Analysis of Data. *Proceedings of Third International Joint Conference on Information Sciences, Durham, NC*, Vol. 3, pp. 403-407.

Øhrn A, Komorowski, J., Skowron, A. & Synak, P. (1998). The Design and Implementation of a Knowledge Discovery Toolkit Based on Rough Sets: The Rosetta System. *Rough Sets in Knowledge Discovery* (L. Polkovski & A. Skowron, eds.), Heidelberg, Germany: Physica Verlag.

Pawlak Z (1991). *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academinc Publishers.

Reynolds RG (1984). A Computational Model of Hierarchical Decision Systems," *Journal of Anthropological Archaeology*, No. 3, pp. 159-189.

Reynolds RG (1994). An Introduction to Cultural Algorithms *Proceedings of the Third Annual Conference on Evolutionary Programming, River Edge, NJ* (A. V. Sebald and L. J. Fogel, eds.), pp. 131-139, World Scientific Publishing.

Reynolds RG (1999). The Impact of Raiding on Settlement Patterns in the Northern Valley of Oaxaca: An Approach Using Decision Trees. *Dynamics in Human and Primate Societies* (T. Kohler and G. Gumerman, eds.), Oxford University Press.

Reynolds RG and Al-Shehri H, "Data Mining of Large-Scale Spatio-Temporal Databases Using Cultural Algorithms and Decision Trees", in Proceedings of 1998 IEEE World Congress on Computational Intelligence, Anchorage, Alaska, May 4-9, 1998.

Russell SJ & Norvig P (1995). *Artificial Intelligence a Modern Approach*. Prentice Hall, Upper Saddle River, New Jersy.

Scacchi W (1998) Modeling, Simulating, and Enacting Complex Organizational Process – A Life Cycle Approach. *Simulation Organizations Computational – Models of Institutions and Groups.* Prietula, MJ, Carley KM, Grasser, L eds. AAAI Press/ MIT Press Menlo Park, California, Cambridge Massachusetts.

Utgoff P E. (1989) Incremental Induction of Decision Trees In Machine Learning, P. Langley, ed. Pp. 161-186 Boston Kluwer.

Vinterbo S (1999, December). *Predictive Models in Medicine: Some Methods for Construction and Adaptation*. PhD thesis, Norwegian University of Science and Technology, Department of Computer and Information Science.

Weiss SM, Indurkhya N (1998). *Predictive Data Mining A Practical Guide.* Morgan Kaufmann Publishers, Inc., San Francisco, California.