INTEGRATION OF CELLULAR AUTOMATA AND GIS FOR SIMULATING LAND USE CHANGES

Nagaratna P Hegde, * Dr I V MuraliKrishna, **Dr K V ChalapatiRao
Associate Professor, CSE Dept, VCE, HYDERABAD-500031,E-Mail:nagaratnaph@yahoo.com
*Director, R & D, JNTU, HYDERABAD-500075, E-Mail : ivm@ieee.org
** Rtd Professor, Osmania University, HYDERABAD E-Mail: kvchalapatirao@yahoo.com

KEY WORDS: Geographic Information Systems; Simulation, CellularAutomata, Neural Network

ABSTRACT:

Cellular automata have been used as a simulation technique in the study of an impressively wide range of urban phenomena, including regional growth, urban sprawl, gentrification, residential growth, population dynamics, economic activity and employment, historical urbanization, land use evolution, and polycentricity to name but a few. A spatial model consists of a collection of processes performed on spatial data that will produce information, usually in the form of a map. These models can often be represented as process flow diagrams, like showing how the output from one process can be the input to a subsequent process. C A are ideal for simulating static entities in spatial models and processes that operate by diffusion. They are ideal for encoding spatial structures into simulation models.
The application of CA in land use/land cover/ urban modeling can give insights into a wide variety of urban phenomena. Urban CA models have better performance in simulating urban growth than conventional urban models because they are much simpler than complex mathematical equations, but produce results that are more meaningful and useful. Temporal and spatial complexities of urban systems can be well modeled by properly defining transition rules in CA models. CA simulation provides important information for understanding urban theories, such as the evolution of forms and structures. GIS is a technology that is used to view and analyze data from a geographic perspective. The spatial representation of an object and its related non-spatial attribute are merged into a unified data file. In practice the area under study is covered by a fine mesh or matrix of grid cells and particular ground surface attribute value of interest occurring at the center of each cell point is recorded as the value for that cell. It should be noted that while some raster models support the assignment of values to multiple attribute per discrete cell, others strictly to a single attribute per cell structure

1. INTRODUCTION

1.1 Simulation

Differential equations, partial differential equations and in some instances, empirical equations have been the underlying mathematical tools behind spatial simulation models. Approaches based on cellular automata models are proposed herein to replace the conventional tools. Issues such as the definition of transition rules, computer implementation with raster geographical information systems and model verification are discussed.

1.1.1 Types of Spatial problems

Types of spatial problems that can be approached using cellular structure and 'rules' are
O Spatially complex systems (e.g., landscape processes)
O Discrete entity modeling in space and time (e.g., ecological systems, population dynamics)
O Emergent phenomena (e.g., evolution, earthquakes)
Transition probabilities for the typical CA model depend on the state of a cell, the state of its surrounding cells, the physical characteristics of the cell (e.g., terrain, soil quality, vegetation, hydrology, and demographic characteristics), and the weights associated with the neighborhood context of the cell (e.g., proximity to other villages and the time since settlement). These weights and neighborhood conditions are determined from empirical analyses of LUCC based on social survey data, the GIS database that represents resource endowments of a site, and the spatial linkages between villages, land parcels, and other critical landscape features

1.2 Cellular Automata

Cellular automata (CA) models consist of a simulation environment represented by a grid of space (raster), in which a set of transition rules determine the attribute of each given cell taking into account the attributes of cells in its vicinities. These models have been very successful in view of their operationality, simplicity and ability to embody both logics- and mathematics-based transition rules. It is thus evident that even in the simplest CA, complex global patterns can emerge directly from the application of local rules, and it is precisely this property of emergent complexity that makes CA so fascinating and their usage so appealing.
Cellular Automata (CA) models were originally conceived by Ulam and Von Neumann in the 1940s to provide a formal framework for investigating the behavior of complex, extended systems. CA are dynamic, discrete space and time systems. A cellular automaton system consists of a regular grid of cells, each of which can be in one of a finite number of k possible states, updated synchronously in discrete time steps according to a local, identical interaction rule. The state of a cell is determined by the previous states of a surrounding neighborhood of cells.

1.2.1 Working principle

The CA model in general works by

- simulating the present by extrapolating from the past using the image time-series,
- validating the simulations via the remotely sensed time-series of past conditions and through the available collection of field observations,

- allowing the model to iterate to the year of choice in future and
- comparing model outputs to an autoregressive time-series approach for annual conditions

1.2.2 The algorithm for CA and image

```
For each iteration
{
 For every cell
 {
If   cell is the same state as its group made by several adjacent neighbor cells  Keep the state of the cell   unchanged
      Else   choose the majority cells' value
  }
}
```

2. Urban Modelling

The formalism described in the first section adapted to meet the needs of urban researchers in several ways:

Cell space: Of course, the idea of an infinite spatial plain is unrealistic in an urban context. Cellular automata are therfore constrained in their cell space to finite dimensions. The regularity of this space is also questionable in urban contexts. Some cities are quite regular in their structure (at least from the perspective of block configuration, say, in a city like Manhattan), but most are markedly irregular, e.g., cities such as Dublin, Athens, Venice, etc. Recent research has adapted the structure of cellular automata spaces and rendered them irregular (see David O' Sullivan's research in this field).

Cell states: In the traditional cellular automaton, cell states are discrete (and quite often binary): alive or dead, one or zero. There is little in the city, however, that is discrete. Most conditions--land use, land value, land coverage, demogrpahic mix, density, etc.--are continuous, and of course urban spaces are multi-faceted. Therefore, cellular models of urban systems commonly contain several cell states simultaneously, and these states range in type from absolute to discrete and to continuous. An innovative adaptation to the traditional idea of the cell state is the introduction of fixed (states that cannot be altered by transition rules) and and unfixed cell states, corresponding respectively, for example, to water sites or land values.

Neighborhoods: The idea of the neighborhood in the formal cellular automaton is rather restrictive. Urban neighborhoods come in many shapes, configurations, and sizes. Complaining that neighborhoods such as the Moore and von Neumann restrict the level of spatial variation that cellular automata models can generate, many researchers have tinkered with neighborhoods to include the notion of 'action-at-a-distance'. Cell neighborhoods consisting of 113 cells, have in some cases been used in simulations.

Transition rules: Perhaps the greatest tinkering with cellular automata models comes in the formulation of the transition rules. It is here that cellular models of urban systems are generated with adherance to what we know in theory about cities. Recently, urban studies using cellular models have intriduced an innovative range of parameters into transition rules in a bid to enhance their realism. These parameters have included probabilistic functions, utility-maximization, accessibility calculations, exogenous links and constraints (linking cellular models to other models), weights, hierarchies, inertia, and stochasticity.

Figure 1 shows the comparison of the tradition CA(left side) and CA used by the land use.

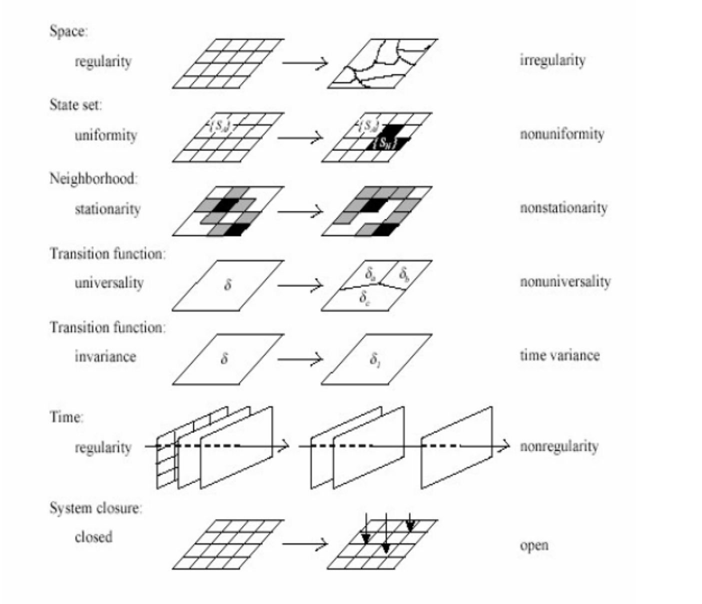

Figure 1 comparison of traditional CA and land use

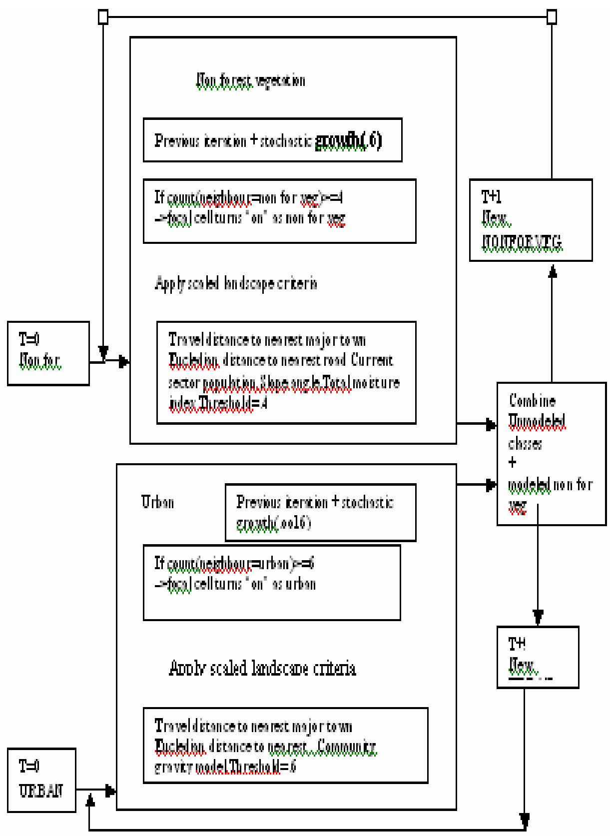

FIGURE 2 FLOW-CHART OF THE CA MODEL FOR URBAN GROWTH

Figure 2 is a generalized flow-chart of the CA model that can be used for spatial simulations of urbanisation. Only two land use/cover classes were modeled for this specific activity - forest to non-forest vegetation and forest/non-forest vegetation to Urban.

2.1.1 Rules to model forest to non-forest vegetation included:

· travel distance to the nearest three "major" communities as a measure of geographic accessibility -- lower values indicate a greater probability of change; distance was computed as Euclidean distance to the nearest road and then simple distance along the road network to the community -- lower values indicate a greater probability of change;

· sector population (i.e., a cluster of individual farm and their aggregate population) was computed as an indicator of rural labor -- higher values indicate a greater change probability;

· slope angle and soil moisture potential as indicators of resource endowments -- greater the slope angle, less is the probability of change, lower the soil moisture index, the greater the probability of change;

· the model parameters included a 0.06 stochastic value of percent random change, a kernel threshold of 4-cells for neighborhood change (within a 3 x 3 moving window), and a masking threshold of 0.4 (a user specified threshold in which cells above that value can change, whereas cells below that threshold can not change; this threshold value is not used as probability values to derive the change itself, only as an indicator of whether change can occur).

2.1.2 Rules to model forest/non-forest vegetation to urban :

·travel distance to the nearest three "major" communities as a measure of geographic accessibility - lower values indicate a greater probability of change; distance was computed as Euclidean distance to the nearest road and then simple distance along the road network to the community -- lower values indicate a greater probability of change;

·community gravity model that estimated yearly population of the "major" and "minor" communities within and adjacent to the ISA, divided by the log distance to the nearest community; actual population counts for 1990, 1999, 2000 were used for most communities; establishment date of communities was also used;

·the model parameters included a 0.0001 stochastic value of percent random change, a kernel threshold of 6-cells for a neighborhood change (within a 3 x 3 moving window), and a masking threshold of 0.6.

The CA process starts with a 1986 LULC classification using Landsat TM data. An annual time steps and a 30-m cell size is used to model LULC change for the period 1987-2010. Stochastic parameters are cells that are randomly selected to be "turned on" as new cells of the class being modeled. Threshold values are set to determine whether the focal cell of the 3 x 3 kernel will change based upon the other cells and a suitability scoring of the input layers. Input layers are processed using GIS functions; greater values indicate a greater likelihood of a cell changing from its initial state to possible outcome state(s). Suitability scores indicate areas of greater or lesser likelihood of change based on multiple criteria. A masking parameter is applied to this score to regulate change/no-change cells.

3. CA AND GIS INTEGRATION

Most current GIS techniques have limitations in modeling changes in the landscape overtime, but the integration of CA and GIS has demonstrated considerable potential.(Itami,1988 and Deadman et al., 1993). The limitations of contemporary GIS include, its poor ability to handle dynamic spatial models, poor handling of the temporal dimensions(Park and Wagner 1997). In coupling GIS with CA, CA can serve as an analytical engine to provide flexible framework for the programming and running of dynamic spatial models. Masanao and Couclelis(1997) address a generalized modeling formalism of CA, which is extended with Geo-algebra capable of expressing a variety of dynamic spatial models within a common framework.

3.1 Encoding Of Spatial Features : The Raster Model

The spatial representation of an object and its related non-spatial attribute are merged into a unified data file. In practice the area under study is covered by a fine mesh or matrix of grid cells and particular ground surface attribute value of interest occurring at the center of each cell point is recorded as the value for that cell. It should be noted that while some raster models support the assignment of values to multiple attribute per discrete cell, others strictly to a single attribute per cell structure.

3.2 Use of commercial GIS software.

Within this approach, there are two choices of developing spatial simulation models – loose-coupled integration or tight-coupled integration with GIS. If commercial GIS software meets the needs of building a simulation model, a tight-coupled integration is the ideal solution. Under this situation, spatial simulation model will be represented in GIS macro language, such as Arc/Info AML, or Arcview Avenue. Takeyama and Coulclelis (1997) developed the basic concepts of Geo-Algebra. Geo-Algebra is a mathematical generalization of map algebra, and capable of expressing a variety of dynamic spatial models and spatial data manipulations within a common framework. But Clarke (1998) stated that Geo-Algebra was not sufficient to represent a land use CA simulation model and suggested the more flexible approach – loose-coupled integration. When commercial GIS software could not handle the complexity of the spatial simulation model, and the model also requires some basic spatial data management, display and analysis, a loose-coupled approach usually is suggested. Loose-coupled integration develops the simulation model with C, C++,JAVA or other programming languages, and connects it with commercial GIS software. GIS saves the efforts to develop a spatial data view/analysis system.

4. ERRORS

A series of inherent model errors can be identified for CA models. They are related to
the following aspects:

· Discrete entities in space and time;
· Neighborhood definitions (types and sizes);
· Model structures and transition rules;
· Parameter values;
· Stochastic variables

4.1 neural network based cellular automata

It is very tedious to calculate the parameter values in conventional CA model. So 3 layer NNW model can be used .

Input layer takes site attributes as the input and output layer represents the land use type. Hidden layer is responsible for the nonlinearility of the network. The NNW is trained with the history data usingsupervised trainging algorithm(backpropagation). The network weights captures relation with the input and output after training the network. During simulation for a given value of site attributes how the land use changes can be shown.

5. CONCLUSION

NNw is convenient to use but these models are blackbox in nature. The meaning of the parameter values are difficult to explain because the relationship among neurons are quite complex.
CA model allows experiment to conducted on simulated systems rather than the real thing .So it is cheaper . it allows the alternative scenario to be evaluated . Different policy options can be considered and their impacts on the future is considered

6. REFERENCES

1 Theoretical and Prac-tical Issues on Cellular Automata, Proceedings of the Fourth International Conference on Cellular Automata for Research and Industry (ACRI 2000), Cellular Models of Urban Systems David O'Sullivan Paul M. Torrens June 2000

2 HOW CELLULAR MODELS OF URBAN SYSTEMS WORK. (1.THEORY) Paul M. Torrens ,paper series of Centre for Advanced Spatial Analysis(CASA)

3 An optimized cellular automata approach for sustainable urban development in rapidly urbanizing regions: D.P. Ward ,Department of Geographical Sciences and Planning, University of Queensland, Brisbane, Queensland 4072 Australia, Geocomputation 99

4 Cellular Automata and GIS - GEOG516 Presentation Feb 1999

5 http://www.stephenwolfram.com/publications/articles/ca/85-twenty/

6 Anuj Kumar Singh, 2003. Modelling land use land cover changes using cellular automata in a geo-spatial environment. M Sc Thesis, ITC Netherlands. Available Online: http://www.itc.nl/library/Academic_output/2003/MSc_theses_2003.asp

7 Batty, M.; Couclelis, H.; and Eichen M. 1997. "Editorial: Urban systems as cellular automata". *Environment and Planning B*. 24:159-164.

8 Geoffrey G. Roy and Folke snickers, CityLife : A study of cellular automata in urban dynamics , GISDATA-4 SERIES on SPATIAL ANALYTICAL PERSPECTIVES ON GIS , CHAPTER 16,213-228

9 Batty, M., Xie, Y. and Sun, Z. L. (1999) Modeling urban dynamics through GISbased cellular automata, *Computers, Environment and Urban Systems*, Vol. 23, 205-233.

10 David O'Sullivan Paul M. Torrens, "Cellular Models of Urban Systems" CASA (Centre for Advanced Spatial Analysis University College London) WORKING PAPER SERIES paper no 22 June 2000

11 Antonio S Camara, Francisco Ferreira and Paulo Castro . "Spatial Simulation Modelling " , GISDATA-4 SERIES on SPATIAL ANALYTICAL PERSPECTIVES ON GIS , CHAPTER 15,201-212

12 XIA LI AND ANTHONY GAR-ON YEH "NEURAL-NETWORK-BASED CELLULAR AUTOMATA FOR SIMULATING MULTIPLE LAND USE CHANGES USING GIS", INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE. VOLUME 16, NUMBER 4 / JUNE 01, 2002

13 Jeannette Candau "Calibrating a cellular automaton model ofurban growth in a timely manner"4th International Conference on Integrating GIS and Environmental Modeling(GIS/EM4):Problems, Prospects and Research Needs. Banff, Alberta, Canada, September 2 - 8, 2000