Irregular Cellular Spaces: Supporting Realistic Spatial Dynamic Modeling over Geographical Databases

Tiago G. S. Carneiro[1], Gilberto Câmara[2], Raian V. Maretto[1]

[1] Computer Science Department – Federal University of Ouro Preto, MG, Brazil

[2] Image Processing Division – National Institute for Space Research, SP, Brazil

{tiago, raian}@iceb.ufop.br, gilberto@dpi.inpe.br

Abstract. *This paper presents a new computational model for representing the Geographic Space concept, called Irregular Cellular Space, which main goal is supporting the development of multiscale spatially explicit dynamic models integrated to geographical databases. This model has been implemented in a modeling software platform named TerraME which has been used for the development of some interesting environmental dynamic models and form simulation of dynamic spatial patterns of change.*
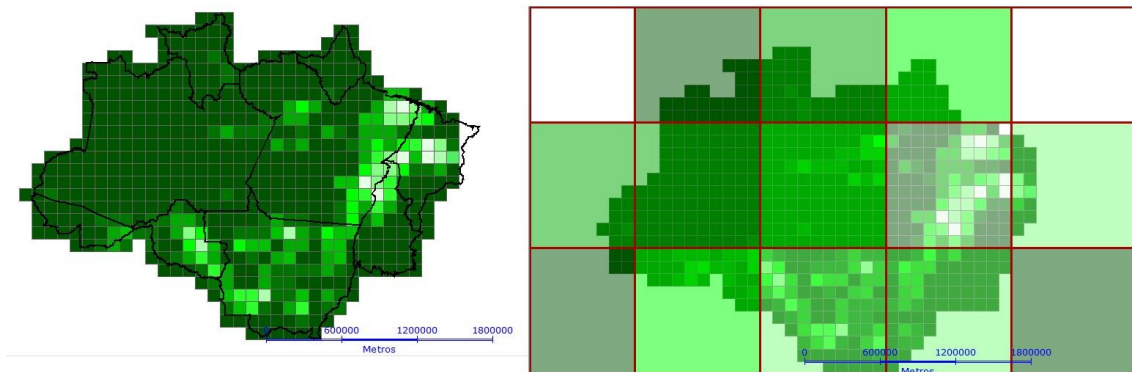
Resumo. *Este artigo apresenta um novo modelo computação para representar o conceito de Espaço Geográfico, chamado Espaço Celulares Irregulares, que tem como principal objetivo servir como suporte para o desenvolvimento de modelos dinâmicos espacialmente explícitos integrados a bancos de dados geográficos em múltiplas escalas. Este modelo foi implementado em um ambiente de modelagem chamado TerraME que tem sido utilizado para o desenvolvimento de modelos ambientais dinâmicos e para a simulação de padrões espaciais de mudança.*

1   Introduction

The modern Geographic Information Systems (GIS) lack on representing dynamical aspects from geographic space. The most provides only a static computational representation for geo-objects, geo-fields and fluxes. This fact had led to several proposals of integration between dynamical modeling and GIS platforms [Box 2002] [Villa and Costanza 2000] [North et al. 2006]. In general, the space have been represented as a **regular cellular space** (RCS), i. e., a regular two-dimensional grid of multi-valued cells grouped into neighborhoods, where the dynamic model rules operate and possibly change cells attribute values. Due to the regular structure of this spatial model, the resulting modeling platforms inherit all disadvantages from raster representations for the space concept: border effects and cell attributes aggregation

problems dependent on the chosen grid resolution [Costanza and Maxwell 1994] [Kok and Veldkamp 2001], and the lacking of abstractions for representing moving objects, as flights and trucks following their routes, or geographical networks, as roads or communication lines. Figure 1 illustrates some of these disadvantages, border effects and aggregation of cell attribute values.
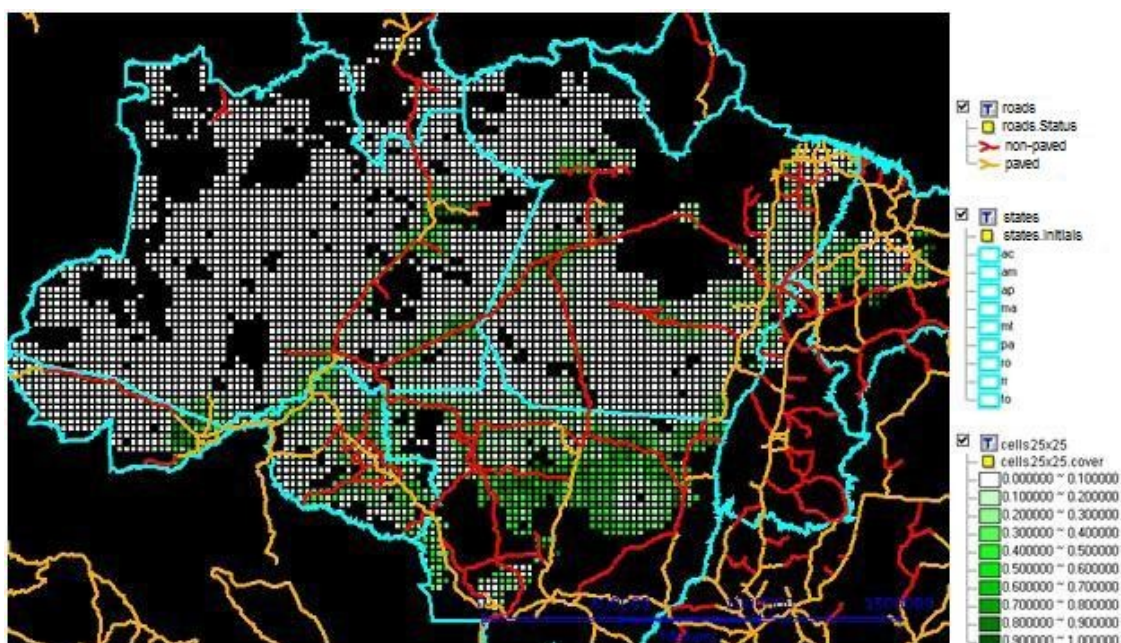
There are several reasons for highlighting cellular spaces as a promising space representation for computer based dynamic modeling and simulation. The existence of a simple and formal cellular space based model of computation, which concept and structure may be efficiently specialized for dynamic space representation on discrete environments, the Cellular Automata (CA) [von Neumann 1966] model. In a Euclidian two-dimensional grid, one may use basic analytic geometry knowledge to describe change circular or elliptical paths. It is easy to develop algorithms for representing process trajectories: to go to East just increment the X coordinate, to go to the South decrement the Y coordinate. However, besides its simplicity, the CA model has enough complexity to simulate spatial diffusive processes and emergent phenomena [Batty 1999] [Wolfram 1984].

Figure 1. Problems due to the choice of a raster structure for space representation: aggregation of cell attribute values and border effects. Maps color: light green means "deforested" and dark green means "forest".

This work presents a formal model for the geographical space concept, called **Irregular Cellular Space** (ICS), which extends the spatial structure from the RCS to support the development of GIS integrated spatial dynamic models which uses many space representations for supporting multiple scale modeling. For model evaluation, the ICS conceptual model has been implemented in environmental modeling software platform called TerraME. Its properties have been fully stressed in some interesting use cases

[Aguiar et al. 2005] [Almeida et al. 2008]. This paper exercises these properties in a simple and pedagogic deforestation model for the Brazilian Amazon region. Figure 2 shows the three ICS used in the deforestation model described at the end of this paper.

Figure 2. Three Irregular Cellular Spaces: (1) 25x25km2 sparse squared cells which main attribute is landCover *(white = "100% forest" and green = "0% forest")*; (2) each polygon representing one Brazilian State is a cell which main attribute are *name = {"MG" | "SP" | " RJ" | ...| "AM"}* the demanded area to be deforested; and (3) each roads is a cell which main attributes are status *(red = "paved" and orange = "non-paved")* and *brazilianSate = {"MG" | "SP" | " RJ" | ...| "AM"}*.

2 Basic Definitions

Following Castells (1999), this work views geographical space as a combination of "spaces of fixed locations and spaces of fluxes", where the concept of 'spaces of fixed locations' represents the geographical space as arrangements of objects on the **absolute space** (location), and the concept of 'spaces of fluxes' indicates spatial arrangements based on **relative space** (situation). Couclelis (1997) has proposed the notion of **proximal space** which aims to formalize this point of view.

In the regular cellular space model, a proximal space model, cells are indexed by their two-dimensional coordinates (x, y) meaning their relative position in the lattice of cells, and the cells neighborhood relationships are stationary following the Moore or Von

Newman pattern [Couclelis 1997]. In the ICS model there is no rigid structure for the space representation. The cellular space is any irregular arrange of cells which geometrical representation may vary from a regular grid of same size squared cells to a irregular set of points, lines, polygons, nodes and arcs, pixels, or even voxels. The n-dimensional space elements are indexed by a family of modeler defined functions, named **spatial iterators**, which map cell index values onto cells references. Topological relationships are expressed in terms of **Generalized Proximity Matrixes** (GPMs) allowing the representation of non-homogenous spaces where the spatial proximity relations are non-stationary and non-isotropic. A GPM is a weighted matrix where the spatial relations are computed taking into account both absolute space relations such as Euclidean distance or adjacency and relative space relations such as topological connection on a network [Aguiar and Câmara 2003].

3   The Irregular Cellular Space Model

(definition 1) The ICS is a set of cells defined by the structure *(S, A, G, I, T)*, where:

- $S \subseteq R^n$ is an n-dimensional Euclidian space which serves as support to the cellular space. The set S is partitioned into subsets, named cells, $S = \{S_1, S_2,..., S_m \mid S_i \cap S_j = \varnothing, \forall i \neq j, \cup S_i = S\}$.

- $A = \{(A_1, \preceq),(A_2, \preceq),...,(A_n, \preceq)\}$ is the set of partially ordered domains of cell attributes, and where a_i is a possible value of the attribute $(A_i, \preceq)$, i.e., $a_i \in (A_i, \preceq)$.

- $G = \{G_1, G_2,...,G_n\}$ is a set of GPMs – Generalized Proximity Matrix (Aguiar, Câmara et al. 2003) used to model different non-stationary and non-isotropic neighborhood relationships, allowing their use of conventional relationships, such as topological adjacency and Euclidian distance, but also relative space proximity relations, based, for instance, on network connection relations.

- $I = \{(I_1, \preceq), (I_2, \preceq),..., (I_n, \preceq)\}$ is a set of domains of indexes where each $(I_i, \preceq)$ is a partially ordered set of values used to index cellular space cells.

- $T = \{T_1, T_2,..., T_n\}$ is a set of spatial iterators defined as functions of form $T_j:(I_i, \preceq) \rightarrow S$ which assigns a cell from the geometrical support S to each index from $(I_i, \preceq)$. Spatial iterators are useful to reproduce the spatial patterns of change since they permit easy definition of trajectories that can be used by the model entities to traverse the space applying their rules. For instance, the distance to urban center cell attribute can be sorted in an ascendant order to form

an index set $(I_i, \leq)$ that, when traversed, allows an urban growth model to expand the urban area from the city frontier.

3.1 Spatial Iterators: modeling spatial trajectories of change

(definition 2) A spatial iterator $T_i \in T$ is an function defined as $T_i{:}(I_i, \leq) \rightarrow S$ that maps modeler built partially ordered sets of index $(I_i, \leq) \in I$ into cells $s_i \in S$.

The following functions should be defined by the modeler in order to construct the set

of indexes $(I_i, \leq)$ and later uses it to build a spatial iterator.

- (definition 2.1) filter:$Sx(A_i,\leq) \rightarrow$ Boolean is a function used to filter the ICS, selecting the cells that will form the spatial iterator domain. It receives a cell $s_i \in$ S and the cell attributes $a_i \in (A_i, \leq)$ as parameters and returns "true" if the cell s_i will be inserted in $(I_i, \leq)$ and "false" if not.

- (definition 2.2) $\leq{:}(Sx(A_i,\leq))x(Sx(A_i,\leq)) \rightarrow$ Boolean is the function used to partially order the subset $(I_i, \leq)$ of cells. It receives two cell values as parameters and returns "true" if the first one is greater than the second, and otherwise it returns "false".

- (definition 2.3) SpatialIterator:$SxAxRxO \rightarrow T$ is a constructor function that creates a spatial iterator value $T_i \in T$ from instances of functions of the families R and O, where R are the filter functions as in definition 2.1 and O are the $\leq$ function as in definition 2.2. The SpatialIterator function is defined as: SpatialIterator(filter, $\leq$) = $\{(a_i, s_i) \mid$ filter(s_i, a_i) = true $\forall a_i \in (A_i, \leq)$ and $\forall s_i \in S; a_i \leq a_j$ $\forall i \leq j; s_i$ = spatialIterator(*filter*, $\leq$) $\forall s_i \in S$ and $a_j \in (A_i, \leq)$ where i = j$\}$.

Figure 3 shows a source code piece, written in the TerraME modeling language, where a spatial iterator is created in order to simulate the deforestation process in Land Use and Cover Change (LUCC) models. The first parameter is the ICS for which the spatial iterator "it" is being created. The filter function, second parameter, selects only cells form "cs" whose land cover is "forest". The function $\leq$, third parameter, orders the cells according to their distance to the nearest road, making cells closer to roads more suitable to change. To construct a spatial iterator that traverses a two-dimensional ICS according to its Euclidian coordinates, from North to the South and from West to the East, one may define the function $\leq$ as: $\leq(c1, c2)$ = { "true" if $c1.x < c2.x$; or "false" if $c1.x > c2.x$; or $(c1.y < c2.y)$ otherwise}.

113

```
it = SpatialIterator{
        cs,
        function( cell ) return cell.cover == "forest"; end,
        function( c1, c2 ) return c1.distRoad > c2.distRoad; end
}
```

Figure 3. A TerraME representation of a spatial iterator used for modeling a deforestation process which spreads along the roads
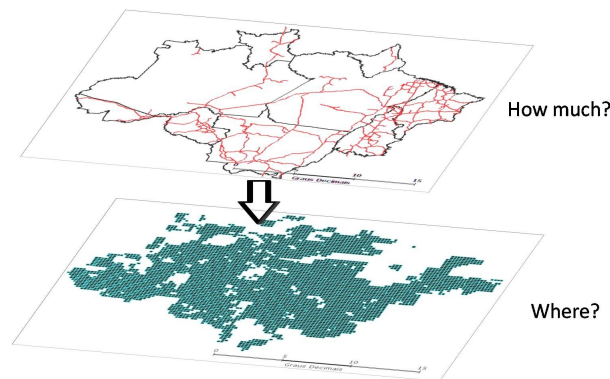
3.2 Dynamic Operations on ICS

Different operations have been defined for traversing the ICS space or the cells neighborhoods applying rules which may cause changes: **ForEachCell**, **ForEachNeighbourhood**, and **ForEachNeighbor**.

- **(definition 3)** ForEachCell:TxF$\rightarrow$A denotes the function that uses the spatial iterator $T_i \in T$ to traverse an ICS applying a modeler defined function $f_m \in F$, where F is the family of functions from the form f_m:SxNxA$\rightarrow$A that calculates the new values for the attributes $a_j^t \in A_j$ from the cell $s_j \in S$ received as parameter. These functions also receives two others parameters: $n \in N$ a natural number corresponding to the relative cell position in the partially ordered set $(I_i, \leq) \in I$ used to define the spatial iterator T_i, and $a_j^{t-1} \in A$ the old values of the attributes a_j^t.

- **(definition 4)** ForEachNeighbourhood:SxGxF$\rightarrow$A is a function which traverses the set of neighborhoods, G, from the cell received as parameter and applies a modeler defined function $f_v \in F$ to each cell neighborhood $g_i \in G$, where F is the family of functions from the form f_v:G$\rightarrow$Bool. The function f_v receives a neighborhood g_i as parameter and returns a Boolean value: true if the ForEachNeighbourhood function should keep traversing the cell neighborhoods, or false if it should stop.

- **(definition 5)** ForEachNeighbor:SxGxF$\rightarrow$A is a function which receives three parameters: a cell $s_i \in S$, a reference to one of neighborhood $g_i \in G$ defined for this cell, and a function $f_n \in F$, where F is the family of functions from the form f_m:(SxA)x(SxA)xR$\rightarrow$Bool. The ForEachNeighbor function traverses the

neighborhood g_j and for each defined neighborhood relationship it applies the function f_m with the parameters $f_m(s_j, s_j, w_{ij})$, where $s_j \in S$ is the s_i neighbor cell and w_{ij} is a real number representing the relationship weight.

3.3 Application on Land Use and Cover Change Modeling

LUCC models distinguish between the projections for the quantity of change and for the location where these changes will take place [Veldkamp and Lambin 2001]. First, a sub-model which has rules that govern the amount of change (the "how much?" question) runs. It is called "demand model". Then, another called "allocation model" determines where the projected change will take place (the "where?" question). This structure is shown in Figure 4. At the next step, the LUCC models are back to the first stage until the simulation finishes.

Figure 4. The general structure of LUCC model, with two sub-models: the "demand model" and the "allocation model".

A pedagogic LUCC model which simulates the deforestation process in Brazilian Amazon region has been developed. It stresses ICS properties for supporting multiple scales modeling through the use of multiple computational representations for the space concept. Three ICS models have been used to represent the deforestation process at two different scales, Figure 1. The demand model uses the cellular space composed by Brazilian States (polygons) and the other formed by Brazilian roads (lines). Each State has two attributes {deforestDemand, forestArea}. Let S_{BR} be a Brazilian State. At each simulation step (year), the amount of area demanded for deforestation in the State SBR is calculated as "SBR.deforestDemand = realRate * SBR.foresArea", where "realRate" is the annual deforestation rate, which is calculated as "realRate = absRate * paved",

where "paved" is the percentage of paved roads in the State SBR and "absRate" is the absolute deforestation rate, which is directly proportional to the density of roads in the State SBR. The equation used to calculate this rate is shown at Figure 5, where "SBR.kmRoads" is the sum of the perimeter of all the roads of the state SBR, "SBR.area" is the area of the state SBR, "totalKmRoads" is the sum of the perimeter of all the roads of all states, "totalArea" is the sum of the areas of all states and "deforestRate" is a model parameter provided by the modeler which represents the average deforestation rate for the whole Legal Amazon area. As the roads change their status dynamically (Figure 7), that is, at each four years ten percent of the perimeter of the roads from each State changes its "status" attribute value from "non-paved" to "paved", then the States deforestation rates are also dynamic. Initially, all cells were 100% forest, in other words, the value of each "landCover" cell attribute was equal to the cell area ($25x25km^2$). The total forest area of a State can be calculated as "$S_{BR}.forestArea = S_{BR}.forestArea - S_{BR}.deforestDemand$".
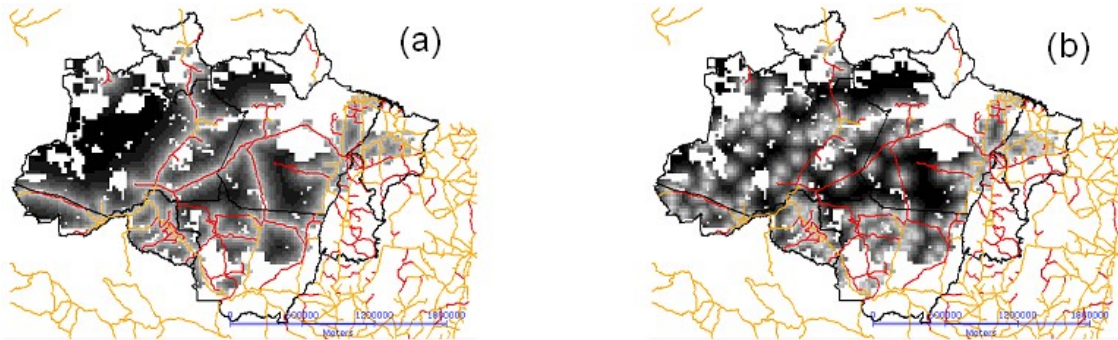
$$absRate = \frac{\dfrac{S_{BR}.kmRoads}{S_{BR}.area}}{\dfrac{totalKmRoads}{totalArea}} * deforestRate$$

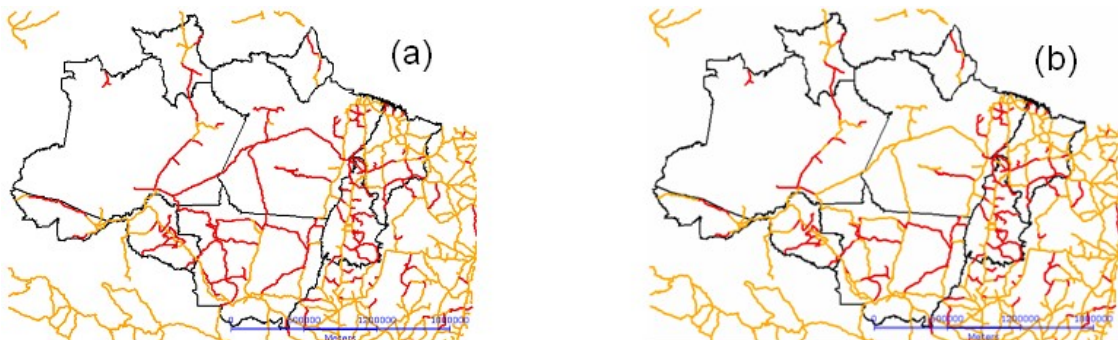**Figure 5.** Equation used to calculate the absolute deforestation rate.

The location model uses only the sparse cellular space of small squared cells to determine where the changes will take place. It is based on a common approach: to compute a change potential surface. At this approach, each small squared cell will have a numeric value indicating how prone it is to change (deforestation). Then the model traverses the cellular "surface" in an ascending order of potential, applying the changes [White et al. 1998]. Some LUCC models use multi-linear regression for change potential computation, such as the CLUE model [Veldkamp and Fresco 1996]. Other approaches include a stochastic combination of diffusive and spontaneous change, such as the DINAMICA model [Soares et al. 2002]. In other to reproduce the "fishbone" spatial pattern of land occupation, where the deforestation spreads along the roads, the change potential of each cell has been computed through the spatial iterator defined in Figure 3, which resulting change potential surface has been shown in Figure 6(a). To simulate the land occupation process which is based only on the spatial expansion of the
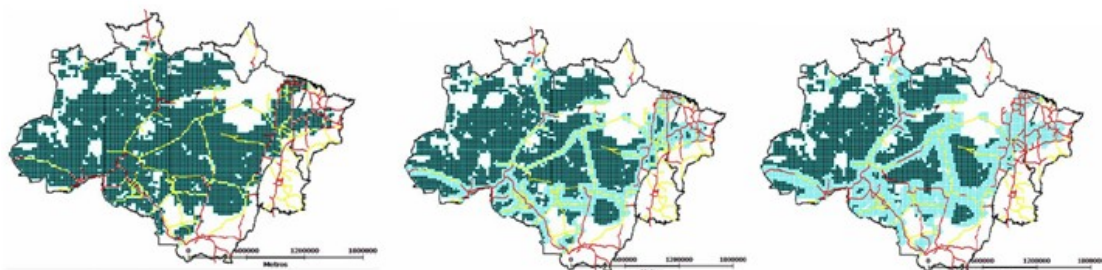
old human settlements, the change potential should be computed as in Figure 6(b). The surface on Figure 6(b) has been defined by the function $\leq (c_1, c_2) = \{c_1.distUrban < c_2.distUrban\}$, where "distUrban" is the cell distance to the nearest urban center. Figure 8 shows the model results for the first, fifth and tenth years of simulation.

Figure 6. Change potential surfaces defined through spatial iterators based on (a) distance to road or (b) distance to urban centers. The gray scale surface reflects the potential for change of each cell: dark gray means low potential for change and light gray means high change potential.

Figure 7. In central Amazon area: non-paved roads (red) in the past (b) become paved (orange) in the future (b).

Figure 8. Deforestation model results: dark green are the 100% forest cells and light blue are the 100% deforested cells.

4 Related Works

To understand the needs related to the space representation in GIS integrated modeling platforms, we examine the proposed extensions of the CA model on the LUCC modeling literature. Several theoretical papers have proposed CA extensions for a better representation of geographical phenomena (Couclelis 1997; Takeyama and Couclelis 1997; Batty 1999; O'Sullivan 2001). In the specific case of LUCC modeling, recent works extend the original CA model and make it more suitable for representing the complexity of human-environment interaction (White, Engelen et al. 1998; Soares, Cerqueira et al. 2002; Almeida 2003). However, most extensions are not useful for multiple scale spatial dynamic modeling.

As an alternative for single-scale modeling of environmental changes, some authors have proposed the layered CA model (Straatman, Hagen et al. 2001), where every cell in one layer has one parent cell in the upper layer and an arbitrary number of child cells in the lower layer. This arrangement allows the combination of models that operate in different spatial resolutions. However, the layered CA model requires a decision about the spatial stratification, where each cell is dependent on a parent cell and controls a number of child cells. The layered CA falls short of providing adequate support for multiscale modeling, since it handles only layers of fixed spatial resolutions. This approach constrains the generality of the system, since the different processes are constrained to fit the hierarchical spatial structure.

5 Results and Future Works

The contributions from this work can be divided in three parts:

(a) The formal model for representing the Geographic Space concept: the ICS allows realistic multiple scale dynamic modeling through the simultaneous use of different computational representations for the space concept;

(b) The TerraME spatial dynamic modeling platform: the ICS model has been implemented in a software platform, named TerraME, which supports GIS integrated environmental model development;

(c) ICS applications: the ICS model has been used in some important modeling studies in order to simulated the human-environment interaction in the Brazilian Amazon region [Aguiar et al. 2005] and in some Brazilian National Parks [Almeida et al. 2008].

Many problems in the multiple scales spatial dynamic modeling has been not addressed on this work. However, the ICS model may be a first step towards a computational model for representing dynamic spaces on GIS environments. Among the future works, many other operations can be defined for the ICS model, for instance, operation to create neighborhood relationships or operations to couple cellular spaces form different resolutions, i. e, to couple two different scales.

6 References

Aguiar, A. P. D., Kok, K., Câmara, G., Escada, I. (2005) "Exploration of patterns of land-use change in the Brazilian Amazon using the CLUE framework." Proceedings of the Open Meeting of the Human Dimensions of Global Environmental Change Research Community, 6. Bonn, Germany.

Aguiar, A. P., Câmara, G., Monteiro, A. M., Souza, R. C. (2003) Modeling Spatial Relations by Generalized Proximity Matrices. Proceedings of Brazilian Symposium in Geoinformatics, 2003.

Almeida, R. M., Macau, E. E. N., França, H., Ramos, F.M. (2008) "Modelo de propagação de fogo em incêndios florestais e a teoria de percolação", XXXI National Conference on Applied and Computational Mathematics, Brazil.

Batty, M. (1999) "Modeling urban dynamics through GIS-based cellular automata". Computers, Environment and Urban Systems, v. 23, p.205-233.

Box, P. W. (2002) "Spatial units as agents: Making the landscape an equal player in agent-based simulations. In: Gimblett, H. R. (ed). Integration of agent-based modelling and geographic information systems. London UK: Oxford University Press, 2002.

Castells, M. (1999) "A Sociedade em Rede.", São Paulo: Paz e Terra.

Costanza, R. and T. Maxwell (1994)"Resolution and Predictability: an Approach to the Scaling Problem." Landscape Ecology v. 9, no. 1, pp 47-57.

Couclelis, H. (1997) "From cellular automata to urban models: New principles for model development and implementation". Environment and Planning B-Planning & Design, v. 24, n. 2, p. 165-174.

Kok, K.; Veldkamp, T. (2001) "A. Evaluating impact of spatial scales on land use pattern analysis in Central America." Agriculture Ecosystems & Environment, v. 85, n.1-3, p. 205-221.

North, M.J., Collier, N.T., Vos, J.R. (2006) "Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit," ACM Transactions on Modeling and Computer Simulation, Vol. 16, Issue 1, pp. 1-25, ACM, New York, New York, USA.

O'Sullivan, D. (2001) Graph-cellular automata: a generalised discrete urban and regional model. Environment and Planning B-Planning & Design, v. 28, n. 5, p. 687-705, 2001.

Soares, B. S., Assunção, R. M. (2001). "Modeling the spatial transition probabilities of landscape dynamics in an Amazonian colonization frontier". Bioscience, v. 51, n. 12, p. 1059-1067.`

Straatman, B., Hagen, A. (2001). "The Use of Cellular Automata for Spatial Modelling and Decision Support in Coastal Zones and Estuaria" M. M. T. R. I. f. K. a. Systems. Maastricht, The Netherlands: Maastricht University.

Takeyama, M.; Couclelis, H. (1997) "Map dynamics: Integrating cellular automata and GIS through Geo-Algebra" International Journal of Geographical Information Science, v. 11, n. 1, p.73-91.

Veldkamp, A.; Fresco, L. O. (1996) CLUE: a conceptual model to study the Conversion of Land Use and its Effects. Ecological Modelling, v. 85, p. 253-270.

Veldkamp, A.; Lambin, (2001) E. F. "Predicting land-use change". Agriculture Ecosystems & Environment, v. 85, n. 1-3, p. 1-6.

Villa, F., Costanza, R. (2000) .Design of multi-paradigm integrating modeling tools for ecological research Environmental Modeling and Software, Elsivier, Volume 15, Issue 2, pp.169-177

Von Neumann, J.. (1966) "Theory of self-reproducing automata". Illinois: A.W. Burks.

White, R.; Engelen, G. (1998). „Vulnerability assessment of low-lying coastal areas and small islands to climate change and sea level rise – Phase 2: Case study St. Lucia. Kingston, Jamaica: United Nations Environment Program" - Caribbean Regional Coordinating Unit.

Wolfram, S. (1994) "Cellular automata as models of complexity." Nature, v. 311, p. 419-424.