

# Spatial Evolving Agents

## *An heuristic approach in Agent-Based and cellular automata modelling for finding optimal solutions in a geographic context*

Juan M. Escamilla-Molgora<sup>1\*</sup>

### Abstract

The aim of this review is to describe briefly two related heuristic methods for modelling and simulating spatial phenomena using geographic information systems. These methods are: *Cellular Automata* (CA) and *Agent-Based models* (ABM). An introduction explaining the basic concepts, the *rules* and how they work is explained in the first section. A brief introduction to genetic algorithms is presented in the second part. This exposition would serve as a background for understanding how these models can be used to solve complex problems. Specially problems that defined in an analytical or numerical manner could have a non polynomial time solution. Finally two iconic examples are given that support the proposed approach. The Axelrod model for cultural aggregation and an ecological application that models the change of vegetation coverage using cellular automata.

### Keywords

Agent-Based Models, Genetic Algorithms, Spatially Explicit Models, GIS Algorithms, Swarm Intelligence, Optimization Heuristics

<sup>1</sup> Department of Physical Geography and Ecosystem Science, Lund University, Sweden

\*Corresponding author: molgor@gmail.com

## 1. Introduction

In the GIS world there are many different definitions of models. In this review, model would be defined as a simplified representation of reality. A model can be constructed as a computer program that uses a simplified representation of one or more aspects of reality. This attributes are transformed in order to create a new representation of the real-world. Models can be classified as *static*: if the input and output correspond to the same point in time. Or *dynamic*: if the output represents a later point in time than the input (Longley et al., 2005). Modelling can serve a number of purposes. *Static* models provide indexes or indicators that can provide some predictors of impacts, sensitivities, or vulnerabilities. *Dynamic* models go further by attempting to project quantifiable impacts into the future and are used to assess different management or development scenarios (Castle and Crooks, 2006).

An automaton is a processing mechanism with features that change over time based on some internal rules. Automata process information from their surrounding as input. and their characteristics are altered according to rules that govern their reaction to those inputs (Castle and Crooks, 2006). In the scope of this review, *dynamic models* would be exemplified as *Cellular Automata* or *Agent-Based Models*.

### 1.1 Cellular Automata

Cellular automata are dynamical recursive models that are discrete in space, time and state. A simple cellular automata  $A$  is defined by a lattice  $L$ , a state space  $Q$ , a neighbourhood  $N$  and a local transition function  $f$  (decision rule) (Adamatzky, 1994).

$$A = (L, Q, N, F)$$

Each element (cell) in the lattice  $L$  is set in a state  $s \in Q$ . The cells can be linked in several ways, the most common is a geometric link. Meaning that two cells are connected if they share a common boundary. The cells are indexed by  $i \in I$  where  $I$  is the index set. Cells change its state in discrete time. The most common case is synchronously, meaning that all cells change their states simultaneously. The change in state of cell  $c$  depends on the cells that are *inside* the neighbourhood  $N$  of the  $c$  ( $N(c)$ ) and the corresponding transition function  $f(N(c))$ . For spatial (geographic) models, two-dimensional C.A. are used. For many applications, a simple neighbourhood of shape 3x3 is used for each cell (Balzter et al., 1998). The neighbourhood that consists of 8 cells surrounding the central cell is called *Moore Neighbourhood*. The neighbourhood that have only 4 cells *North*, *South*, *East*, *West* is called *Von Neuman Neighborhood* (Adamatzky, 1994).

C.A. can be *deterministic* if the change of state in any cell is always the same given the neighbourhood. Contrary, a C.A is *stochastic* if each cell has a probability of change

from one state to another (Balzter et al., 1998). In general, a deterministic C.A. is a type of convolution. The transition function  $f(N(c))$  also called *rules* has the form:

$$a_{t+1}^c = f(a_{t-r}^{c-r}, \dots, a_t^c, \dots, a_{t+r}^{c+r})$$

Where:  $a_t^c$  represents the state of cell  $c$  at time  $t$ ,  $r$  is the neighbourhood  $c$  and  $f$  is the local transition function that represents the transition rules. For each time step  $t$  the set  $\{a_t^c | \forall s \in I\}$  is frequently called the configuration of the C.A. at time  $t$  (Balzter et al., 1998). The iconic example of a two dimensional deterministic, nonetheless, chaotic (susceptible to initial conditions and non periodic patterns (Devaney, 2003)) is Conway's Game of Life in which the transition rule depends only whether there exists a constant number of *active* cells in the neighbourhood see (Gardner, 1971)

### 1.1.1 more about transition functions

Consider a Neighbourhood  $N$  of size 3 and a state space  $Q = \{0, 1\}$ , for a one-dimensional C.A. The transition function is going to take as input a boolean vector of size 3. Therefore, the domain will be the  $2^3 = 8$  in size. Now, for the output, the function will return 0 or 1, meaning that the total number of rules is size  $2^8 = 256$ . It can be seen that the number of rules or transition functions increases exponentially depending on the size of the neighbourhood. Finding a optimal transition rule for a particular problem often yields to a NP-Complete problem. Several heuristics methods involving fitness measures exists and genetic algorithms is an approach often used (Mitchell, 1998). Perhaps the more detail description of cellular automata, specially the 256 rules of the elementary 1-D automaton is the book *A new Kind of Science* Wolfram (2002).

### 1.1.2 Remarks on automata

The C.A. can represent only the changes in state of the phenomena observed (objects). C.A. can be a good approximation to reality when the abstraction entity is something that does not interact with the object itself. Sometimes it is necessary that the objects transform the environment, and dialectically, the environment affects the response of the objects. For model this type of problems Agent-Based Models are proposed.

## 1.2 Agent-Based Models

Although there is no general consensus on a formal definition of the term *agent*. There are several characteristics that an Agent-Based Model should have (taken from Castle and Crooks (2006)):

**Pro-active** Agents have goals to achieve based on their behaviours'. For example, agents that try to find a certain path in the space.

**Reactive** Agents can have knowledge of their surroundings. Agents can have also prior knowledge or memory maps. Extending the example, the agents can have prior knowledge of the space, in which direction is the path.

**Interactive** Agents have the possibility to share information between other agents. Agents can query other agents and the environment in an arbitrary size neighbourhood.

**Mobile** Agents can have motility in the space.

**Adaptable** Agents can be designed to modify their inner state based on the environment and the relations with other agents, permitting agents to adapt with a form of memory or learning. Agents can adapt at the individual level (e.g. learning alters the probability distribution of rules that compete for attention), or the population level (e.g. learning alters the frequency distribution of agents competing for reproduction).

Agents can also be designed to be adaptive, which can produce Complex Adaptive Systems (CAS; Holland, 1995). In this sense, the agents look for the best suitable conditions, based on their proper characteristics. Using the property of heterogeneity, mentioned before, it is possible to generate many agents with slight variations and make an *artificial selection* of the sub-population that fits best certain wanted characteristics.

## 1.3 Genetic algorithms

**Why genetic?** *Artificial Selection* is the process in which humans selects members of a biologic population with certain features and make hybrids between the members this sub-population. Generation after generation better individuals emerge from this population. A good example of artificial selection are the dog breeds. Each breed have been raised accordingly to achieve a goal. Better muscle mass, good for hunting, vermin picker or even just for petting. A similar process have been abstracted into the machine learning and optimization methods. The process consist of instantiate similar but not identical agents. Define a fitness measure that is a function that assess how the agent's responses to the environment. During the process, a subset of the population with the best fitness values is filtered. Variations, called *mutations* are induced into this population. The sub-population reproduce among itself making a new population of siblings. The process is repeated again until the remaining population's fitness gets a high value.

### 1.3.1 The chromosome

Every agent have intrinsic methods that sense the environment in a certain way and responds according to these factors. This response is called *behaviour*. Each agent can have different

**Autonomy** Agents are independent units capable of process and share information between other agents without the influence of a centralized control.

**Heterogeneity** Agents can be defined with distinct characteristics.

**Active** Agents are active because they interact with the environment and share the unique characteristics (heterogeneity) in the simulation. Active features involves:

*behaviours* that correspond to different configurations of the sensed space. The set of all *behaviours* for a given sensed space is called chromosome.

For example suppose that the agent is a car and the problem to model is to avoid collisions with other cars. The agent sense the environment within a radius of 5 meters. Supposing that the model is implemented in a raster of 1 meter resolution, the number of cells in a radius of 5 meters is 100. Now, suppose that each cell could be either occupied or not. All the sensed input is evaluated using the decision function that outputs either 0 or 1 if cell would be occupied or not the next time step. The chromosome would be the set of all the  $2^{100}$  *behaviours*. As in C.A. the sensed space could be indexed accordingly to fit the binary representation of an integer. Therefore, a complete chromosome could be seen as a binary string of size  $2^{100}$ .

The simplest form of genetic algorithm involves three types of operators: *selection*, *crossover* (*single point*), and *mutation* (after (Mitchell, 1998)).

- **Selection** This operator selects chromosomes in the population for reproduction. The fitter (using the fitness measure) the chromosome, the more times it is likely to be selected to reproduce.
- **Crossover** This operator randomly chooses a locus and exchanges the subsequences before and after that locus between two chromosomes to create two offspring. For example, the strings 11000100 and 11111111 could be crossed over after the third locus in each to produce the two offspring 11011111 and 11100100. The crossover operator roughly mimics biological recombination between two single chromosome (haploid) organisms.
- **Mutation** This operator randomly flips some of the bits in a chromosome. For example, the string 00000100 might be mutated in its second position to yield 01000100. Mutation can occur at each bit position in a string with some probability, usually very small (e.g., 0.001).

The process executes iterated until a certain fitness threshold is reached.

## 2. Reviewed examples

### 2.1 The effect of habitat destruction pattern on species persistence: A cellular model

Dytham (1995) proposed a cellular automaton to model the persistence of two plant species in environments with different patterns habitats. The model assumes that two competing species can only coexist if one is the competitive dominant (*species c*) and the other possesses superior dispersal abilities (*species d*). There are four possible patch states in his model: patch permanently destroyed; patch empty but available for colonization; patch occupied by species *c*; patch occupied by species *d*. Colonisation has a range of only one patch. The probability of a patch being colonised is higher if

more of its direct neighbours contain the species. Colonizations and extinctions occur during a discrete time-step. For a new colonised patch is not possible to produce offspring. The impact of four habitat destruction patterns in the population was simulated on a grid of 50x50 cells. The model is a *stochastic cellular automata* defined with a Moore neighbourhood. Dytham's results indicate that species' coexistence heavily depends not only on the amount of available habitat, but also on the destruction pattern. For more details see (Balzter et al., 1998) and (Dytham, 1995).

### 2.2 Segregations of culture

Robert Axelrod in 1997 proposed a simplified agent-based model that described how social entities (modelled as agents) can form clusters with similar features. This model has been an iconic example in agent-based modelling because of its simplicity and the realistic patterns that emerge from this simple rules. Each object, named *social entity*, represent a person within a certain urban context. The idea is to discretize the social features of a social entity. For example: political parties affinities, education, yearly income, age, etc. A distance function is proposed to compare similarities between the objects. The *behaviour* is that objects move in proportion to the distance between them. Performing the same steps in an iterated fashion would yield to a stabilized pattern of clustered agents, where the more similar objects are closer. For a complete description and analysis of this model see (Axelrod, 1997).

## References

- A. Adamatzky. *Identification Of Cellular Automata*. Taylor & Francis Group, 1994. ISBN 9780748401727.
- R.M. Axelrod. *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*. Princeton studies in complexity. Princeton University Press, 1997. ISBN 9780691015675.
- Heiko Balzter, Paul W. Braun, and Wolfgang Köhler. Cellular automata models for vegetation dynamics. *Ecological Modelling*, 107(2–3):113 – 125, 1998.
- Christian Castle and Andrew Crooks. Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. September 2006.
- R.L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley studies in nonlinearity. Westview Press, 2003. ISBN 9780813340852.
- C. Dytham. The effect of habitat destruction pattern on species persistence: A cellular model. *Oikos*, 74(2), 1995.
- M. Gardner. On cellular automata, self-reproduction, the garden of eden and the game life. *Sci. Am.*, 224:112–117, 1971.

Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind. *Geographic Information Systems and Science*. John Wiley & Sons, April 2005. ISBN 0470870001.

M. Mitchell. *An Introduction to Genetic Algorithms*. A Bradford book. MIT Press, 1998. ISBN 9780262631853.

Stephen Wolfram. *A New Kind of Science*. Wolfram Media, January 2002.