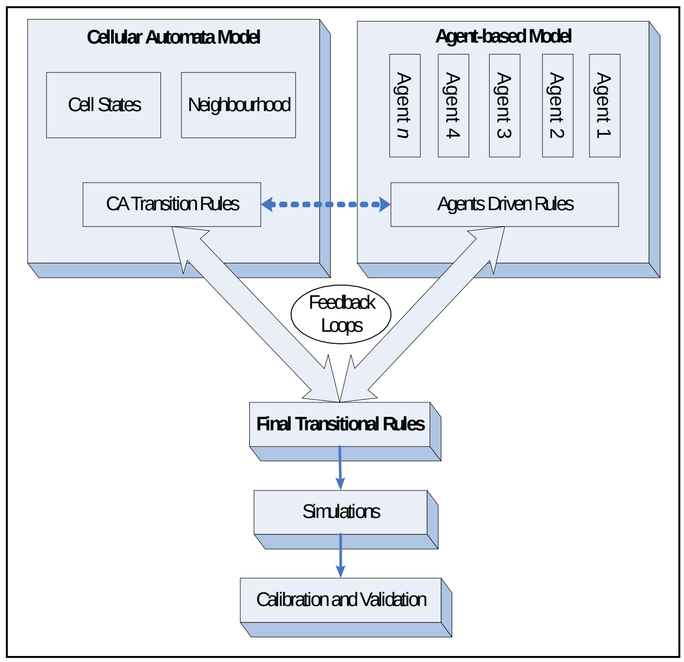
Framework for Integration of Agent-based and Cellular Automata Models for Dynamic Geospatial Simulations

H. S. Sudhira, T. V. Ramachandra, Andreas Wytzisk & C. Jeganathan



March 2005

Technical Report: 100

Centre for Ecological Sciences,
Indian Institute of Science,
Bangalore

Acknowledgements

We thank Indian Space Research Organization – Indian Institute of Science Space Technology Cell (ISTC/BES/TVR/148) for financial assistance to carry out field investigations, National Remote Sensing Agency (NRSA), Hyderabad and Global Land Cover Facility of the NASA and University of Maryland, USA, for providing the requisite satellite data for the study.

Authors address

T. V. Ramachandra	**Energy & Wetlands Research Group** **Centre for Ecological Sciences** **Indian Institute of Science** Bangalore 560 012 **E Mail: cestvr@ces.iisc.ernet.in**
H. S. Sudhira	**Energy & Wetlands Research Group** **Centre for Ecological Sciences** **Indian Institute of Science** Bangalore 560 012 **E Mail: sudhira@ces.iisc.ernet.in**
Andreas Wytzisk	**Intenational Institute for Geo-Information Science** **& Earth Observation, P O Box 6, 7500 AA** ENSCHEDE, THE NETHERLANDS E Mail: wytzisk@itc.nl
Jeganathan	**Geo-Informatics Division** **Indian Institute of Remote Sensing (NRSA)** Dehradun 2480 01, Uttaranchal, India jegan@iirs.gov.in

Abstract

Simulations using the CA technique in geo-spatial modelling have been attempted in the recent times. However it is seen that CA models do not address the external driving factors that are also responsible for the dynamics involved directly. In order to confront the limitation of the CA models not interacting with the externalities driving the process, the integration of agent-based models over a CA model is more appropriate. Hence, the current research is aimed at the development of framework for integration of CA and agent-based models for simulations. For enabling the efficient integration of the CA models and agent-based models at appropriate scales in space and time, the research outlines the framework for undertaking such simulations. This is suggested using the proposed Geo-Spatial Analyser for space variant simulations and incorporating the HLA framework for time variant simulations.

The urban sprawl dynamics is considered for demonstrating the prototype of an agent-based model exhibiting radial urban sprawl. This reveals the pattern of growth that takes place under different scenarios. For the application of agent-based and cellular automata models for a real situation, the case of Mangalore city, Karnataka, India is considered and applied. The Mangalore city is currently experiencing high rates of urbanisation as evinced from the study. For a scenario of an infrastructure initiative like the creation of a ring road, the implications of this are depicted using the combination of agent-based and CA models. The simulations reveal the nature of likely growth in the region due to infrastructure initiative. This work has contributed in the development of the framework for integration of agent-based models and the CA models for undertaking geo-spatial simulations. The case of urban sprawl dynamics was considered and demonstrated using the combination of agents-based models with the CA model for visualising the patterns of growth in conjunction with the drivers of sprawl.

Keywords: Agent-based Models, Cellular Automata, Geo-spatial Simulations, High-Level Architecture and Urban Sprawl

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Models in Space and Time

The advances in geo-informatics coupled with the availability of spatial and temporal information from remotely sensed data have aided to investigate and model the environmental systems. These models are executed for different scenarios or alternatives depending upon the specific objectives. Modelling the spatial and temporal dimensions has been an intense subject of discussion and study for philosophy, mathematics, geography and cognitive science (Claramunt and Jiang, 2001).

The Geographic Information System (GIS) has until now offered storage, retrieval and analysis of spatial and temporal databases. This has helped in modelling of the different environmental processes in the spatial and temporal dimensions. Different models for representing the spatial and temporal phenomenon have evolved from the traditional Location-based/Snapshot models to Entity-based models (Object-oriented approach) to Event-based models to Process-based models and so on. Inspite of the continuous developments in the representations of the geo-spatial data along with their temporal attributes, these systems with the goal of gaining insights about cause-effect relationships are not yet ready to answer the questions of patterns of change through time (Peuquet, 1999).

A simulation is the imitation of the operation of a real-world process or system over time (Banks and Carson, 1984). Simulation is the key technology for describing, assessing, analysing, forecasting, etc. the dynamics of real, planned or virtual systems (Schulze et. al., 2002). An important growing concern in the geo-informatics community is the simulation of the environmental processes using the different models in the spatial and temporal dimensions within a GIS. In the spatial domain, modelling and subsequent simulations have been extensively attempted for the various processes involved in environment, like the meteorology, hydrology, air pollution dispersion, urban growth, land-use / land-cover dynamics etc. The types of mathematical models are categorized by Cellier (1991) into three types based on the different interpretations of time. The first type is the set of continuous-time models, which are characterized by the fact that, within a finite time span, the state variables change their values infinitely often. These continuous-time models are typically represented by the differential equations. The second type is the set of discrete-time models, wherein the time axis is discretized. These models are commonly represented as sets of difference equations when the discretization is equally spaced. The last type of models is the set of discrete-event type models, wherein the time axis of such models are continuous (real, instead of integer). The discrete-event model differs from the continuous-time model by the fact that within a finite time-span only a finite number of state changes can occur.

Describing the approaches of dynamic land-use models, Liu and Anderson (2004) categorize them as process-based and transition-based modelling. Process-based models are known to describe the cau-

sality between the different components of the system explicitly. Transition-based models use probability or similar terms to summarize the changes happened over a time interval.
Mostly, simulations are run based on the models or equations by numerical methods rather than by analytical methods (Banks and Carson, 1984). Analytical methods employ the deductive reasoning of mathematics to solve the equation representing the model. For example, differential calculus can be used to determine the shortest route in a network model. Numerical methods employ computational procedures to 'solve' mathematical models. In the case of simulations, which employ numerical methods, models are 'run' than solved; that is an artificial history of the system is generated based on the model assumptions, and observations are collected to be analysed and to estimate the true system performance measures (Banks and Carson, 1984). The validation of a model is the determination of the model is an accurate representation of the system. Validation is usually achieved through the calibration of the model, an iterative process of comparing the model to actual system behaviour and using the discrepancies of the two, and the insights gained to improve the model.

Modelling and simulations in the spatial domain is now extensively done using the techniques of Cellular Automata (CA), although there are different approaches to model processes in space and time including the geo-statistical approaches, differential equations, etc. Simulations undertaken by CA models and geo-statistical techniques are of numerical type. CA is a cell-based approach to model processes in a two–dimensional space. Benenson and Torrens (2003) define an automaton as a discrete entity, which has some form of input and internal states. These states can change over time according to a set of rules to determine a new state in a subsequent time step. These rules control the transformation of a cell state to another cell state over the specific period of time depending on the neighbourhood of the cells. A detailed account on CA is discussed in the next chapter.

The geo-statistical modelling by Kriging is another numerical method for spatial interpolation or estimation. Kriging, to some extent has similarities to a CA, that there is a function driving the interpolation. By definition of Kriging, it is a method of interpolation, which predicts unknown values from data observed at known locations. This method uses variogram to express the spatial variation, and it minimizes the error of predicted values, which are estimated by spatial distribution of the predicted values. In essence, Kriging uses data points that are minimal or less or sparsely sampled. In this regard, CA approaches assume that entire data set although discrete in space has continuous set of data values in cells and not only a few samples. Further CA offers opportunity to address the temporal dimension at least in its discrete nature at regular time steps, which are not so capable of being handled in other techniques to model processes of space and time. Thus CA has emerged as popular method for modelling processes in space and time.

In CA based modelling the temporal variability of spatio-temporal processes is less addressed than its spatial counterpart. CA based models are used for studying temporal dynamics (Clarke and Gaydos, 1998; Liu and Anderson, 2004) wherein the temporal dimension is mostly considered as duration and discrete. Normally in a CA model, the transitions take place from time t to time t+1 and would ideally simulate the physical states between time t=0 and t+1 (Batty and Jiang, 1999). Although time can be discretized at higher scale (discrete-time stepped), the state changes of the CA for certain transition rules can be assumed to be within specific time only and not at all the discretized time units. In certain situations there can be more transition rules for various state changes at varied time units underlying the importance of the events that take place based on the transitions and not all transitions that may or may not happen at every time step. In case of CA, the state changes are only within the discrete set of time, thus they are a fit case of both discrete-time model and discrete-event model.

Though CA based models have been used for simulating various dynamics extensively, CA are considered as immobile geographical automata (Benenson and Torrens, 2003). They are immobile because individual automata are not free to move in the space in which they reside; all spatial movement takes place through the diffusion of information through neighbourhoods. In addition, Batty and Jiang (2000) argue that the development of spatial interaction in spaces wider than the neighbourhood itself and in enabling the model dynamics to take account of system-wide conservation constraints, are usually destroyed by the CA transitions. Most importantly transition rules account only for the states and neighbourhood and not for externalities driving the processes.

Modelling of land-use / land-cover dynamics in both space and time is bounded by various causal factors driving the changes that can have varied relations in space and time apart from the inherent physical state dynamics. A specific use-case of involving such a land-use / land-cover dynamics considered in this research is the dynamics of urban sprawl. The phenomenon of urban sprawl is characterized by the uncontrolled or uncoordinated growth of the built-up area in the outskirts of a city or along the highways. The inherent causal factors and dynamics involved in the rapid changes occurring in the land-use / land-cover due to urban sprawl is considered as a fit case to apply the CA models for simulating for future scenarios. Typically, the CA models need to account for the external drivers that also drive a land-use / land cover change which are not accounted in the transition rules of the automata. Certain externalities can be system wide or specific to certain locations, for which the CA models have to evolve to address such requirements. Further there are also different significant processes that take place in the region in question, apart from those represented in the CA model. The new wave of research is driving towards integrating the agent-based models (multi-agent systems) with the CA models, such as in the case of modelling land-use / land cover dynamics by incorporating different drivers / agents involved.

The aim of agent design is to create a program, which interacts `intelligently' with its environment. The term ‘agent’ is usually applied to describe self-contained programs, which can control their own actions based on their perceptions of their operating environment. The key hallmarks of agent-hood are autonomy, social ability, responsiveness, and pro-activeness (Jennings and Wooldridge, 1996). Agents, have their origins in software engineering and artificial intelligence where they are used in networking, communications, etc. An agent is an encapsulated computer system that is situated in some environment and that is capable of flexible, autonomous action in that environment in order to meet its design objectives.

The initial application of agent-based models within a GIS was made in studying the dynamics of pedestrian behaviour in streets (Schelhorn et al., 1999). As such agent-based models have been used to model the discrete dynamics of small-scale spatial events for mobility in carnivals and street parades (Batty, Desyllas and Duxbury, 2003). Thus, agent-based models are mobile geographical automata with transition rules.

The combination of CA and agent-based models has enabled to address the interactions in space and time (Batty and Jiang, 1999) by the representation of different behaviours (drivers) as objects. Semboloni et al., (2004) present an interactive multi-agent simulation model on the web, CityDev, which is an interactive, dynamic model that can interact with the user on the web.

Object-oriented approaches to cells, where cells represent land parcels, administrative areas, or even individual buildings, almost invariably require non-regular lattices, but may ease the problem of defining transition rules (O'Sullivan and Torrens, 2000). The agent-based models would have all the characteristics of cellular automata, but, unlike the cellular automata, agents in an agent-based model would be programmed for spatial mobility within the regions they inhabit. The focus here is on to develop the framework for integrating the CA models and agent-based models validating the same for the use case of urban sprawl dynamics.

To depict the fusion of cellular automata and multi-agent systems in a spatial context for simulating discrete, dynamic and action-oriented spatial systems, Benenson and Torrens (2003) propose the term Geographic Automata Systems (GAS). The GAS framework defines the automata collectively considering, type of automata, states, transition rules, location, movement rules, neighbours, and neighbourhood transition rules. Thus, the GAS framework addresses the fusion of CA and ABM in a spatial context, by tight coupling of GAS with GIS. With this, there can be many automata with different scales for space and time, defined based on the process/phenomenon in question. However, when there are two or more automata waiting to be executed at the same space and time, such cases are handled only by synchronous or asynchronous updating.

Furthermore, the management of time in the GAS framework is less advanced to handle multiple simulations simultaneously, for which their framework uses 'synchronous' and 'asynchronous' updating. The GAS framework uses the Object-based Environment for Urban Systems (OBEUS) for validation. In the 'synchronous' updating – "all automata are assumed to change simultaneously and conflicts can arise when agents compete over limited resources. Resolution of these conflicts depends on the model's context, a decision OBEUS leaves to the modeller". Also in case of 'asynchronous' updating – "automata change in sequence, with each observing a geographic reality left by the previous automata. Conflicts between automata are thereby resolved; but the order of updating is critical as it may influence results. OBEUS demands the modeller sets up an order of automata updating according to a template: randomly, sequence in order of some characteristic" (Benenson et al., 2004).

1.2. Scales and Representations of Models in Space and Time

The scales and representations of models in space and time can be addressed depending on the number of dimensions in which they are looked at. The current research perceives the models in space in two-dimensional surface and that of time are considered to be linear. According to Goodchild (2001) the scale implies the level of spatial detail or spatial resolution, often defined as the shortest distance over which the change is recorded and thus having some units. In the geo-spatial representations of models in the raster forms in a lattice structure, the cell side corresponds to spatial resolution. However, the raster representations can be of regular lattice as in the case of remote sensing images or irregular lattice structures like that of an triangulated irregular network model. The universally followed scale for planar models in space is the metric system with the units of meters. However in the geographic space, the earth is considered as an ellipsoid and so specific systems of representation have evolved in two-dimensional space as coordinate systems and datums as references for the measurement of height of the surface on the earth. Among the coordinate systems is the Universal Transverse Mercator (UTM) projection in meters and the geographic coordinates as latitudes and longitudes. As there are different possible geometries, projections, coordinate systems and datums defined for the representation of geographic feature in space, the Open Geo-spatial Consortium (OGC) has adopted a conceptual schema under the Abstract Specification Topic 1 – Feature Geometry, which is identical with the ISO 19107 – Geographic Information – Spatial Schema (OGC, 2003).

On the temporal aspects of scales, the time is linearly discretized in equally spaced intervals as (nano-, micro-, milli-) seconds that make up the minutes, hours, days, weeks, months, years, etc. Obviously most models have a temporal resolution that falls under any one of the classes above mentioned. While modelling the geographic processes, the scales of these features modelled can vary in space and time. And so, for an effective and meaningful simulation of these models, the synchronization of the different models has to be addressed to obtain tangible results in both space and time. The spatial synchronization would refer to the process of coupling the models with different geographic spatial scales to an appropriate scale, for example resampling an image for a particular spatial resolution. In the similar context the temporal synchronizations would address the coupling of different time-variant models. Nevertheless, most processes in the environmental systems take place at their own respective scales in both space and time. Even though if certain processes were individually modelled using CA and/or agent-based models, these models wouldn't have a common spatial and temporal resolution. Although, the geo-spatial community has for long until now attempted to model and simulate these different processes on the desktop aided by the advances in the computational capabilities, a significant challenge faced by the geo-spatial community is the need of dynamic systems for modelling and simulation capable of integrating or synchronizing the different models of environmental processes at their respective scales.

In the case of land use / land cover dynamics involving an urban system, three levels of change processes can be identified according to their time scale: slow processes which take 3–5 years or longer to complete and affect the physical structure of a city, e.g. industrial, residential, and transport construction; medium processes such as economic, demographic, and technological changes which affect the usage of the physical structures; and fast processes completed in less than a year, such as mobility of labour, goods, and information (Wegener, 1994). All processes would have their own respective start and end time apart from the spatial extents they are going to affect. Each of these individual and/or group of processes can be thought of as distributed simulation systems as an analogy to the computer science and engineering domain. Thus a typical framework for simulation should be addressing the synchronization of different models modelled using CA and/or agent-based models at appropriate spatio-temporal scales. In this backdrop the upcoming framework relies on the specifications adopted by the OGC to promote the interoperability and reusability of the simulations.

In the modelling and simulation domains of the computer science and engineering numerous advances are made with respect to synchronization of such distributed simulations and parallel simulations. A key distinction is that the modelling and simulation community are able to achieve the synchronization of the time-variant simulation models using the standardized simulation framework, the High Level Architecture (HLA), an IEEE standard (1516), initially developed by the United States Department of Defence (Dahmann, Fujimoto and Weatherly, 1998; IEEE, 2000).

HLA is the architecture for reuse and interoperation of simulations. The intent of the HLA is to provide a structure that will support reuse of capabilities available in different simulations, ultimately reducing the cost and time required to create a synthetic environment for a new purpose and providing developers the option of distributed collaborative development of complex simulation applications. Thus the HLA offers a federation approach and addresses the interoperability and reuse of individual components. In order to facilitate interoperability, each member (federate) of a distributed simulation (federation) is equipped with appropriate interfaces to interact via a run-time infrastructure.

With the available architecture (HLA) for distributed simulations in the computer science and engineering domain, the current direction of research is to enable the spatial component based on this architecture so as to address spatial synchronizations apart from the temporal synchronization already available. Moreover, an ideal framework for simulation should be reusable and interoperable for different use cases and multidisciplinary research teams thereby minimizing the cost of effort and time for modellers to rebuild a simulation framework specific to their use case or research areas. An interoperable simulation framework should facilitate the modeller to utilize the framework more flexibly by allowing an open system or on a common platform. An open system is an information processing system that complies with the requirements of open systems interconnection (OSI) standards in communication with other such systems. An open systems interconnection (OSI) defines the accepted international standard by which open systems should communicate with each other. It takes the form of a seven-layer model of network architecture, with each layer performing a different function. OSI is used to develop interfaces and integrate two dissimilar systems, for example, PCs and UNIX or UNIX and mainframes. The OSI is sponsored by the International Standards Organization (ISO) (Patterson & Gittings, 1996). In the geo-spatial modelling domain, the proposed framework is to comply with the open system interconnection standards for facilitating the functioning of the framework in any system.

The focus of this research is on overcoming the limitations of the CA models by integrating CA with the agent-based models considering the issues of scales for synchronization of these models in space and time. Consequently and conceptually, this research framework is looking at coupling the agents and CA models, and incorporating a geo-spatial analyser (GSA) to handle the spatial synchronization and the HLA framework to handle the temporal synchronization, the discreteness in both spatial and temporal scales would be more flexible and robust to handle the issues arising out of multiple discrete-time stepped simulations.

1.3. Objective

The main objective is to develop the framework for the integration of agent-based and a cellular automaton model. This investigation will integrate agent-based models and CA model for synchronizing in their respective spatial and temporal scales. Specific questions addressed are:

1. What are the requisite conditions and methods for integrating agent-based models and CA at appropriate scales?
2. What are the requirements for framework to be interoperable and reusable for different processes?
3. What are the calibration and validation techniques (accuracy assessment methods) for evaluating and ensuring the usability of the simulations?

1.4. Method

The prerequisites for the integration of agents and CA were identified through literature review. Analytical solutions for the integration of the agent-based models over the CA as Agent-Based Cellular Automata (ABCA) are developed. A key aspect for the integration is to synchronize models at appropriate spatial and temporal resolutions. Thus, for managing the spatial component for different simulations, a Geo-Spatial Analyser is proposed. The detailed functionalities of this are discussed in the subsequent chapters. The utility of High Level Architecture (HLA) developed by the United States Department of Defence, which provides a time-variant framework for distributed simulations, is incorporated in the simulation framework. The HLA, which is also an IEEE standard (IEEE 1516) uses a federation approach and facilitates interoperability of distributed simulators. The current research addresses the synchronization of the different individual agent-based models in both space and time over a CA. The general simulation framework is shown in Figure 1. In such a framework, the final transition rules of a CA would be derived from the interactions among different agents. Thus the agent-based model would operate on top of a CA layer where agents' respond or initiate state transitions in the CA. The research approach is outlined in Figure 2. The dynamics of urban sprawl is considered as the use case. In case of urban sprawl, the external driving process like infrastructure development process can be an agent.

1.5. Structure of the Document

The report is organised in six chapters. The second chapter describes the theoretical background of the CA models and the agent-based models in detail. In the third chapter, the framework for the integration of CA models and agent-based models outlining the conditions requisite for the integration are addressed along with the concepts of Geo-Spatial Analyser and the HLA framework. This chapter includes the descriptions on the functionalities of the proposed GSA and the HLA architecture. The application of such a framework for the use case of urban sprawl is the fourth chapter. This chapter also describes the study area in question and the drivers considered as agents. A generalized CA model is developed in the simulation framework. The fifth chapter describes the results and discussion after implementing the framework at the prototype level. The calibration and validation of such simulations is also taken up for discussion. The sixth chapter concludes with the scope for further research in the area.

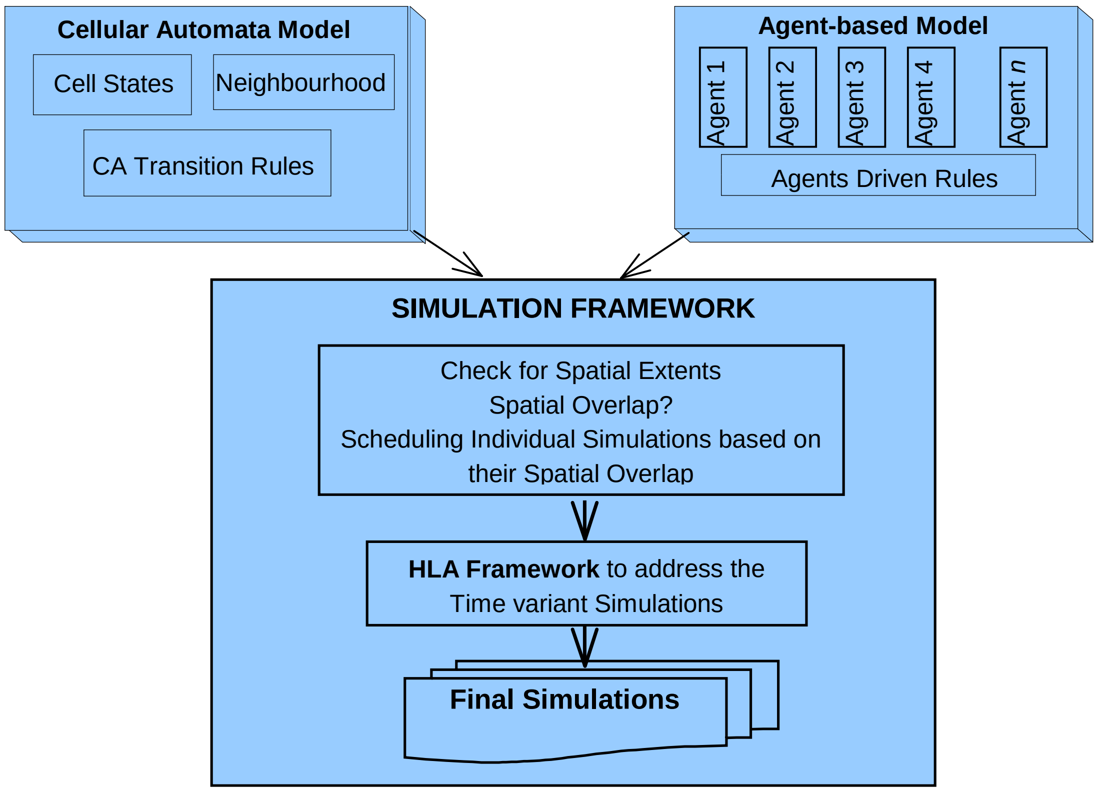


Figure 1.1: Framework for the Simulations Integrating the CA and Agent-based Models

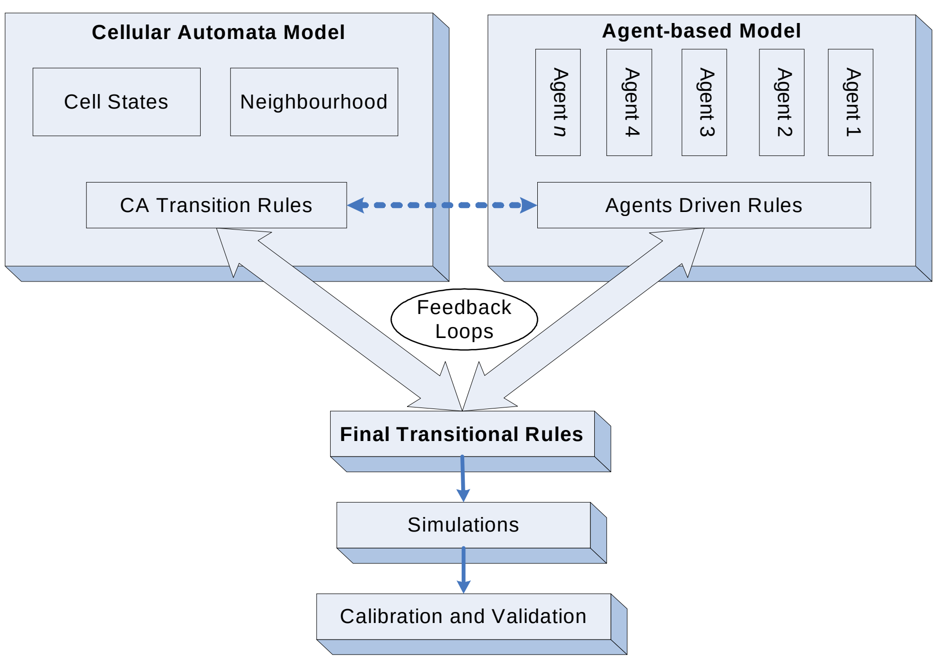


Figure 1.2: Outline of the Research Approach

2. Cellular Automata and Agent-based Models - A Theoretical Framework

2.1. General

Cellular Automata (CA) plays in important role in modelling and simulation of spatio-temporal processes. CA was developed by Ulam in the 1940s and soon used by von Neumann to investigate the logical nature of self-reproducible systems (White and Engelen, 1993) and extensive experiments were done by Wolfram (2002). Later on researchers applied CA to model the geo-spatial domain, especially on urban systems (Couclelis, 1997; Batty and Xie, 1994). However, most models of spatial dynamics rest with land cover and land use change studies (Yang and Lo, 2003) and urban growth models (Batty and Xie, 1997; Batty, 1998; Clarke, et al., 1996; Couclelis, 1997; White and Englen, 1993; White, Englen and Uljee, 1997; Jianquan and Masser, 2002).

2.2. Overview of Cellular Automata

In the geo-spatial domain, CA has been applied to urban systems with fervour and has been used to explore research questions in urban applications (Torrens, 2000). Essentially CA is a cell-based approach to model processes in a two–dimensional space. Benenson and Torrens (2003) define an automaton as a discrete entity, which has some form of input and internal states. These states can change over time according to a set of rules to determine a new state in a subsequent time step. A typical CA system comprises of four components namely, cells, states, neighbourhood and rules. Cells are the smallest units of the system having adjoining neighbours. Any cell can represent a theme by its state. Like in land use map, a cell can have a state of built-up or vegetation or water bodies etc. However in CA the state of a cell can change only based on the transition rules, which are defined in terms of neighbourhood functions (Li and Yeh, 2000). Transition rules are the real engines of change in a CA (Torrens, 2000). These rules control the transformation of a cell state to another cell state over the specific period of time depending on the neighbourhood of the cells. The notion of neighbourhood is central to the CA paradigm (Couclelis, 1997).

The integration of CA with GIS has enabled a flexible framework for modelling, simulation and dynamic visualization of urban systems. A simple transition rule in a cellular automaton model is given by Li and Yeh (2000) as,

$$s^{t+1} \approx f\left(s^{t}, N\right) \qquad \text{… Equation 1}$$

with $s \in S$ (S set of all possible cell states like, built-up or vegetation or open land, etc.). **N** is the neighbourhood of the cell, which acts as inputs for the transition rules. The function ***f*** defines the transition rules from time period **t** to **t + 1**. The adjacent neighbours are defined by the cells formed by the co-ordinates $\{x \pm 1, y \pm 1\}$ as in Moore's neighbourhood, with, x=0, y=0 as the centre cell.

Typically in a CA, the neighbourhood and the transition rules play an important role for the automaton to initiate state transitions in a cell. In most CA models, the automata are influenced only within the Moore's neighbourhood, which are the eight adjacent neighbourhood cells. Few researchers like Wolfram (2002) and White, Engelen and Uljee (1997) have explored the influence of cell states beyond the eight neighbourhood cells on the automata. Figure 2.1 depicts the notion of neighbourhood ascribed by von Neumann and Moore.

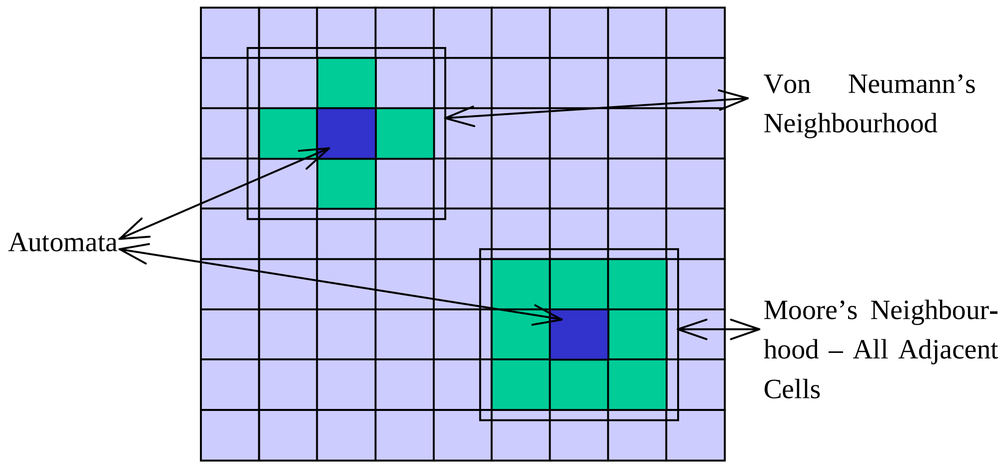


Figure 2.1: Cells and Neighbourhood

An important notion of the CA paradigm is the geometry in two-dimension space as regular tessellated structures notably as the cells. As any typical regular tessellations, the cells are of the same size and shape, and the value attributed to the cell corresponds to the whole region bound by the cell. Other different cell geometries for regular tessellations are triangular and hexagonal cells. Even though uniformly regular spaced square cells are used as in the case of classical CA, they are unreal for realistic representation of reality (Torrens, 2000). Torrens (2000) argues that most objects in reality are not regular and hence they are not square in shape. To counter such situations irregular lattice structures are being introduced in the CA framework. In order to have congruity with the geo-spatial data obtained by remote sensing which are also in regular tessellated square or rectangular cells, the scope of the present research confines to the geometry of the cells used in CA as regularly tessellated square or rectangular cells.

Several advances are being made to explore the possibilities of the CA technique. Notable among them are calibration approaches for constrained CA models, which were applied to model geographic process (Li and Yeh, 2000; Straatman, White and Englen, 2004). These models were constrained by causal factors driving the geographical processes such as urban sprawl, wherein the availability of land and proximity to city centres and highway are without any interaction with the causal factors. According to Li and Yeh (2000), the constraints used are mainly related to land suitability according to accessibility that affects land development probability, such as cost distance to city centres, roads and railways. In their case constraints for sustainable urban development are classified into three types: local, regional and global constraints. Local constraints contain detailed spatial information for each cell, but regional constraints have only aggregated or partial-spatial information. Global constraints, however, are characterized by temporal or non-spatial information. Such an approach of defining the constraints and operating in a CA was the next step of achieving some linkages with the causal factors. However once the simulation is executed, the constraints would be static over time all through. In reality, there are situations wherein the constraints are dynamic dependent on the situations. And so it would require an ideal framework to address issues concerning the dynamics of these causal factors and constraints. Further, in such models these constraints would not behave as an event-based system with a cause-effect relationship.

Typically, the geo-spatial models need to account for the external drivers that are not accounted in the transition rules of the CA. Certain externalities can be system wide or specific to certain locations, for which the CA models have to evolve to address such requirements. Further there are possibly also different significant processes that take place in the region in question, apart from those represented in the CA model. CA models are yet incapable of representing the external factors responsible for driving the change dynamics as the transition rules account only for the states and neighbourhood. Even though certain approaches of using the constrained and state-based CA framework were suggested, these methods are not supportive of individual spatial interactions and linking dynamically directly with the externalities and constraints over time. Given that the land use / land cover change dynamics are subjected to the drivers at various scales, this multi-scale dynamics are ineffectively handled in CA. To counter such paradigm, different approaches are being suggested. Among them is the integration of agent-based models over a CA framework, as agent-based models can be constructed to represent the externalities driving the processes. Thus the current research is approaching towards the integration agent-based models (multi-agent systems) with the CA models, such as in the case of modelling the dynamics of urban sprawl by incorporating different drivers as agents involved enabling the individual spatial interactions by defining the spatial and temporal relationships to these agents.

2.3. Agents from Artificial Intelligence

Agents, have their origins in software engineering and artificial intelligence where they are used in networking, communications and many more applications. The aim of agent design is to create a program, which interacts with its environment. The term ‘agent’ is usually applied to describe self-contained programs, which can control their own actions based on their perceptions of their operating environment. A significant definition is that, an agent is considered as a self-contained program capable of controlling its own decision-making and acting, based on its perception of its environment, in pursuit of one or more objectives. The most general way in which the term agent is used is to denote a hardware or (more usually) software-based computer system that enjoys the following properties (Wooldridge and Jennings, 1995):

- *Autonomy*: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state;
- *Social ability*: agents interact with other agents (and possibly humans) via some kind of *agent-communication language*;
- *Reactivity*: agents perceive their environment, (which may be the physical world, a user via a graphical user interface, a collection of other agents, the INTERNET, or perhaps all of these combined), and respond in a timely fashion to changes that occur in it;
- *Pro-activeness*: agents do not simply act in response to their environment; they are able to exhibit goal-directed behaviour by *taking the initiative*.

There are a number of points about this definition that require further explanation. Agents are (Jennings, 2000):

- Clearly identifiable problem solving entities with well-defined boundaries and interfaces;
- Situated (embedded) in a particular environment—they receive inputs related to the state of their environment through sensors and they act on the environment through effectors;
- Designed to fulfil a specific purpose—they have particular objectives (goals) to achieve;
- Autonomous—they have control both over their internal state and over their own behaviour;
- Capable of exhibiting flexible problem solving behaviour in pursuit of their design objectives—they need to be both reactive (able to respond in a timely fashion to changes that occur in their environment) and proactive (able to act in anticipation of future goals).

Although, the origins of agent-based models have been in the artificial intelligence, they are also developed in social sciences extensively. Amongst other application domains agent-based models are now also used for studying the urban dynamics (Portugali, Benenson and Omer, 1997; Sanders et al., 1997; Benenson 1998; Batty, 2003) over a GIS environment. Agents can be considered as a special case of an automaton, having all features of the general automaton, with a distinction that these agents are mobile and they can represent the external drivers responsible for the processes (e.g. socio-economic, population, etc.). The idea is to treat each of the individual drivers as agent-based automata enabling the spatial and temporal relationships. Supposing in the case of urban sprawl dynamics, economic activity, infrastructure availability/development, population development process, etc. are the different drivers of the process operating over the region at their respective scales in both space and time, each of these is defined as an agent-based model. There can be as many agent-based models as the number of externalities identified driving the processes at appropriate scales. Such processes take place at specific locations and are generally not system wide. While a classical CA transition rule is system wide, such agent-based models would only be specific to certain locations only. These agent-based models are to act in conjunction with the regular transition rules of the cellular automata. This integrated agent-based cellular automaton is discussed in Chapter 3.

The essential concepts of agent-based computing are agents, high-level interactions and organizational relationships (Figure 2.2). It can be seen that there can be numerous agents wherein specific agents can interact amongst themselves and/or have the same sphere of visibility and influence. There can also be agents who act independently without any interaction of other agents and unique sphere of influence.

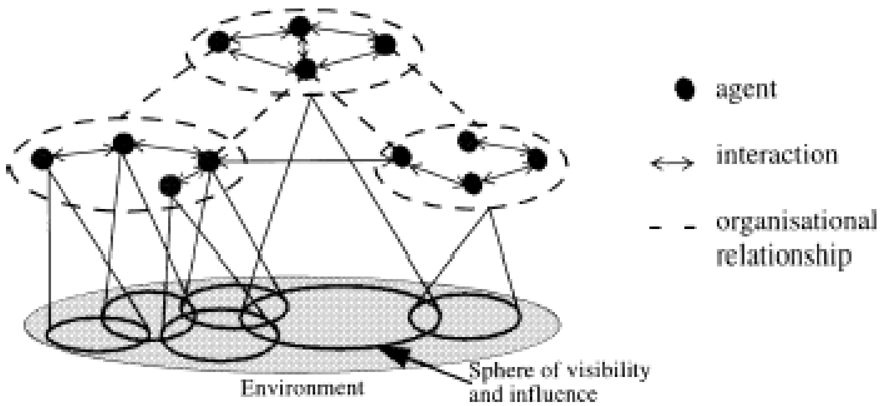


Figure 2.2: Canonical View of an Agent-based System (Jennings, 2000)

2.4. Classification of Agents

The agent ontology has been attempted in a number of ways by different investigators as per their objectives. The classifications are made based on the characteristics (autonomy, reactivity, social ability and pro-activeness) and applications of the agents. Nwana (1996) classified agents in five broad categories.

The first type of classification concerns with the mobility of these agents and so, the agents are classified as static and mobile. The next type of classification addresses the ways in which these agents are modelled. Batty and Jiang (1999) note that the ways in which the agents 'sense' and 'act' in their environment are central determinants of the behaviour which is to be modelled. Hence in this type of classification, agents are categorised as 'reactive' and 'deliberative'. The reactive agents are those agents, which is autonomous and react to its environment or to other agents. The deliberative or cognitive agents are those, which behave according to specific protocols and predefined rules. These agents need not wait for any responses but can act independently, thus emanating the autonomous characteristic of the agents.

In the third classification, agents are classified along several ideal and primary attributes, which agents should exhibit. The three characteristics autonomous, cooperation and learning are used to derive the four types of agents, viz., collaborative agents, interface agents, collaborative learning agents and smart agents (Figure 2.3).

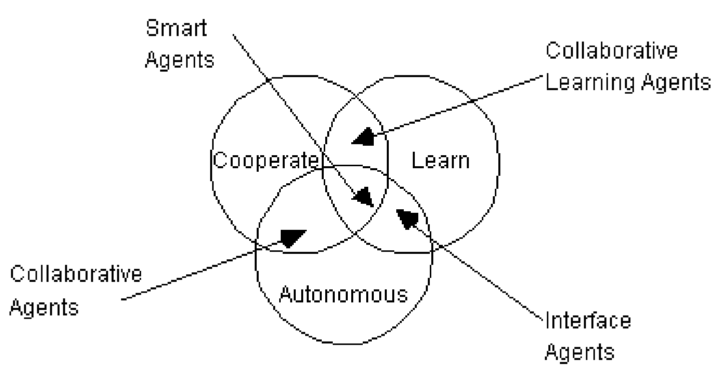


Figure 2.3: Agent Typology (Nwana, 1996)

In the fourth type of classification, the agents are classified by the roles they play like in the case of World Wide Web (WWW) as information agents. This type of agents is used in internet search engines as web crawlers and spiders. This type of agents helps in managing large amount of information over the internet. These information agents may be further classified as static, mobile or deliberative. Lastly, the combination of two or more type of agent behaviours in a single agent is termed as *hybrid* agents.

A significant utility of the agents in the geo-spatial modelling and simulation is that these agents can come handy where individual spatial interactions at the local level are modelled while collectively initiating the global actions, which were not well handled by the traditional CA based approaches. Few or all or a combination of the above mentioned agent-types (heterogeneous agents) can be used in the geo-spatial domain to effectively model and simulate the multi-scale dynamics at the local levels and the global level.

2.5. Tools for Agent-based Modeling

With the intense research in the realm of agent-based modelling under the distributed artificial intelligence domain, scores of tools are developed for building ABMs, in particular by making use of the programming languages C, Objective C and Java. The development of these ABM tools came up with the academic and research institutions for applications in social simulations and studying complex behaviour. Among the earliest tool is the StarLogo, developed by the Massachusetts Institute of Technology's Media Labs. After the StarLogo, then came the SWARM, StarLogoT (a variant of StarLogo), REPAST, ASCAPE and NetLogo (from the makers of StarLogoT). The industrial circuit also has actively taken part in the developments of these agent-based tools. Notable among them are the research in International Business Machines (IBM) Corporation Limited, British Telecom and the open source project – ECLIPSE. The following gives a brief description of the some of the abovementioned tools.

StarLogo: Developed by Massachusetts Institute of Technology (MIT) Media Laboratory (http://www.media.mit.edu/starlogo/), StarLogo is a programmable modelling environment for exploring the workings of decentralized systems that are self-organizing and self-coordinating. A central notion of the StarLogo is that it consists of three elements – turtles as agents, patches as cells and worldview as observer. With StarLogo, it is possible to model (and gain insights into) many real-life phenomena, such as bird flocks, traffic jams, ant colonies, and market economies. StarLogo is a specialized version of the Logo programming language. With traditional versions of Logo, it is possible to create drawings and animations by giving commands to graphic "turtles" on the computer screen. StarLogo extends this idea by allowing the user to control numerous graphic turtles in parallel with “patches" makes the turtles' environment.

StarLogoT: A variant of StarLogo, developed by the Centre for Connected Learning and Computer-Based Modelling, Northwestern University, USA (http://ccl.northwestern.edu/cm/starlogoT/). StarLogoT is a programmable modelling environment for building and exploring multi-level systems. It is one of a class of new "object-based parallel modelling languages" (OBPML). Currently, StarLogoT is a Macintosh-only program that lets the user to explore simulated environments. StarLogoT allows controlling the behaviour of thousands of objects in parallel. This allows one to model the behaviour of distributed and probabilistic systems, often systems that exhibit complex dynamics. StarLogoT was developed at the Tufts University Centre for Connected Learning and Computer-Based Modelling, which has since relocated to Northwestern University. It is an extended version (a superset) of StarLogo, which was developed by the MIT Media Laboratory.

SWARM: Developed by the Santa Fe Institute (http://www.santefe.edu), SWARM (http://www.swarm.org/) is a software package for multi-agent simulation of complex systems. The basic architecture of Swarm incorporates the collections of concurrently interacting agents. Swarm is essentially a collection of software libraries, written in Objective C, for constructing discrete event simulations of complex systems with heterogeneous elements or agents. Some lower-level libraries, which interface with Objective C, are also written in Tk, a scripting language that implements basic graphical tools such as graphs, windows, and input widgets.

REPAST: Stands for REcursive Porous Agent Simulation Toolkit (http://repast.sourceforge.net/), which is an open source, agent-based simulation toolkit for creating agent-based simulations using Java (1.4 or higher). RePast provides a library of classes for creating, running, displaying and collecting data from an agent based simulation. In RePast simulations the running of the simulation is divided into time steps or "ticks." Each tick some action occurs using the results of previous actions as its basis. RePast has much of the functionalities that are borrowed from the Swarm simulation toolkit and could be termed as "Swarm-like".

ASCAPE: Developed by Centre on Social and Economic Dynamics (CSED), Brookings Institution (http://www.brook.edu/es/dynamics/models/ascape/default.htm), Ascape (Agent-Landscape) is a research tool to support agent-based modelling and simulation. A high-level framework supports complex model design, while end-user tools make it possible for non-programmers to explore many aspects of model dynamics. It is written entirely in Java, and run on Java-enabled platform. Models developed within it can be easily published to the web for use with common web browsers. Sugarscape models using the ASCAPE (Epstein and Axtell, 1996: In, Batty and Jiang, 1999) demonstrate how agents are used to move over a cellular space.

NetLogo: This is a cross-platform multi-agent programmable complexity modelling environment also developed by the Centre for Connected Learning and Computer-Based Modelling, Northwestern University, USA (http://ccl.northwestern.edu/netlogo/). NetLogo comes with a large library of sample models and code examples that help beginning users get started authoring models. NetLogo is in use by research labs and university courses across a wide variety of domains in social and natural sciences.

Intelligent Agents Project at IBM T.J. Watson Research: The mission was to develop intelligent agent technology that is highly reusable and easy to integrate with a broad spectrum of networked applications. The research also contributed to company-wide efforts in strategy and in common architecture, e.g., for inter-agent knowledge-level communication and interoperability. The reusable intelligent agents technology is embodied as an extensible structured class library, called *RAISE* (Reusable

Agent Intelligence Software Environment). RAISE is object-oriented in design and is implemented in C++. RAISE also features dynamic plugging ability of user-authored rule sets (including easy merging and updating), and development-time plugging ability of reasoning engines.

ABLE: The *Agent Building and Learning Environment (ABLE)*, a project made available by the IBM T. J. Watson Research Centre (http://www.alphaworks.ibm.com/tech/able), is a Java framework, component library, and productivity toolkit for building intelligent agents using machine learning and reasoning. The ABLE framework provides a set of Java interfaces and base classes used to build a library of JavaBeans called AbleBeans. The AbleBeans library contains reading and writing text and database data, data transformation and scaling, rule-based inferences (using Boolean and fuzzy logic), and machine learning techniques (such as neural networks, Bayesian classifiers, and decision trees). Developers can extend the provided AbleBeans or implement their own custom algorithms.

ZEUS: The ZEUS toolkit, which provides a library of software components and tools that facilitate the rapid design, development and deployment of agent systems. The three main functional components of the ZEUS toolkit are: the agent component library, agent building tools and the visualisation tools. The components of the agent component library implement the different aspects of agent functionality. The agent building tools provide an integrated development environment through which the agents are specified and generated. The visualisation tools enable applications to be observed and, where necessary, debugged.

JADE: Java Agent DEvelopment Framework (JADE) is a software framework to develop agent-based applications in compliance with the Foundation of Intelligent Physical Agents (FIPA - http://www.fipa.org) specifications for interoperable intelligent multi-agent systems (http://jade.tilab.com/index.html). The goal is to simplify the development while ensuring standard compliance through a comprehensive set of system services and agents. JADE can then be considered as an agent middleware that implements an Agent Platform and a development framework. It deals with all those aspects that are not peculiar of the agent internals and that are independent of the applications, such as message transport, encoding and parsing, or agent life cycle.

The abovementioned tools are not the exhaustive ones, while a detailed description of all the tools is beyond the scope of this research. A detailed description of the software and tools related to the agent-based modelling has been given by Torrens (2004) - http://www.geosimulation.org/abm/, Testfatsion (2004) - http://www.econ.iastate.edu/tesfatsi/acecode.htm, SWARM Wiki site (2004) - http://wiki.swarm.org/wiki/Tools_for_Agent-Based_Modelling.

Although there is significant number of tools for building agent-based models, these tools are yet to evolve for applications in geo-spatial simulations. A main reason for this is that the spatial relationships or the topology and geometry have to be defined in these tools for ensuring them to handle the geo-spatial databases. The prevalent tools for building ABM are of significance only while dealing without any spatial relationships. In other sense, these agents are not bound by the geo-spatial data models. Consequently the research is attempting to define the spatial relationships for the agents modelled so that they can be used in the geo-spatial domain. In the current research NetLogo was used to develop a prototype of a hypothetical model for demonstrating the urban sprawl phenomenon in radial direction. NetLogo was favoured because it was user-friendly and supported extensive documentation for building models. The description of the prototype demonstration is dealt in Chapter 5.

2.6. CA and Agent-based Models as Distributed Simulation Systems

A simulation is a computer program that models the behaviour of a physical system over *time*. The program variables or the state variables represent the current state of the physical system. Simulation program modifies state variables to model the evolution of the physical system over time. A key paradigm in the conventional simulation systems is the treatment of time. The notion of time is as per Cellier (1991) as continuous-time models, discrete-time models and discrete-event models.

In reality, the models of urban sprawl dynamics are discrete-event and discrete-time. A discrete-event simulation refers to the computer model for a system where the state variable changes at discrete points in simulation time. The essential ingredients of a discrete-event simulation are the system state or state variables and the state transitions or the events itself. A discrete-event simulation can be viewed as a sequence of event computations, with the computation of each event is assigned to a time stamp or simulation time. Thus each event computation can change the state variables or initiate new events. A discrete-event simulation system refers to the model of the physical system together with the simulation executive comprising of the event management list and simulation time advancing mechanisms.

An execution of the collection of such discrete simulation systems simultaneously makes the distributed simulation systems. A key aspect of a distributed simulation systems is the execution of various simulations simultaneously to optimise the usage of processors, memory, input and output devices, as well aiming at the synchronization of these different simulations. Typically distributed simulation refers to the technology concerned with executing computer simulations over computing systems containing multiple processors which may be using tightly coupled multiprocessor systems or workstations interconnected via a network (e.g., the Internet) and so on. The key advantages for opting distributed simulation systems are to minimize the model execution time, integrate simulation running on different platforms (interoperability and reuse), enable geographical distribution, for scalable performance, and fault tolerance ultimately enabling a higher performance computing. The execution of the different agent-based automata involve the automaton to change as per the transition rules as event lists with specific time advancement mechanisms in iterations. And so, the collection of all such agent-based models with the CA model is treated as distributed simulation systems.

The forthcoming chapter discusses the general framework for the integration of the agent-based models with the CA model.

3. Development of the Integration Framework for CA and Agent-based Models

3.1. The Agent-Based Cellular Automata (ABCA): Combining CA and Agent-based Models

Based on available approaches for modelling geo-spatial phenomena using traditional cell-based CA techniques, the future of geo-spatial simulations is being aimed at integrating agent-based modelling techniques with dynamic capabilities to handle spatio-temporal phenomenon for better and efficient decision-making. Batty and Jiang (1999) illustrate the correspondence of these ideas in Figure 3.1 showing the progression of cell-based approaches to CA to agent-based models.

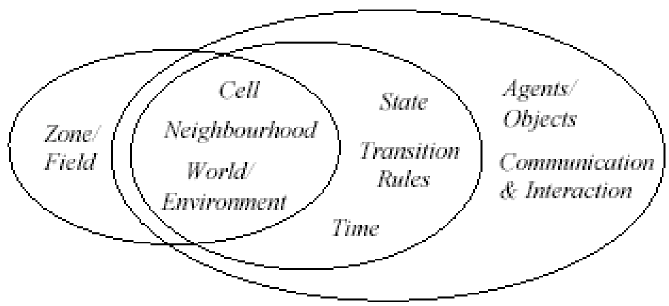


Figure 3.1: Relations between Cell-Based GIS, CA Modelling, and MAS (Batty and Jiang, 1999)

In this backdrop, an Agent-based Cellular Automata (ABCA), which combines CA and agent-based models, is proposed. In the ABCA framework, the object-oriented approach to cells is coupled with the transition rules defined by the models as automata. The agent-based models defining the transitional rules are termed as agent-automata.

3.2. Agent-based Models as Geographic Objects

A topology is the definition of spatial relationships between features. From earlier discussions (Chapter 2), it is clear that even though there are significant tools for building agent-based models, they lack the spatial relationships or the topology definitions without corresponding geometry to the agents thus inhibiting them to operate in the geo-spatial domain. A key aspect of these agent-automata is that the spatial relationships are defined with respect to certain geometry to these and so they are conceived to act over a geo-spatial domain. In the geographic perspective, the topology or the spatial relationships are defined with respect to Euclidean space, wherein every point in the space can be represented by a set of coordinates. The representations of geographic objects within the context of topological features are the simple geometric shapes in respective dimensions are called 'simplices'. These are point

(0-simplex), line (1-simplex), polygon (2-simplex) and tetrahedron (3-simplex). Since the present research is confined to the two-dimensional Euclidean space, the definitions of objects are restricted to points, lines and polygons. Typically these representations of geographic objects refer to the vector data types. However, the main focus of this research is on to simulate the agent-based models over a CA model, which is essentially a cell-based model. Thus, it would be inconclusive for simulation of agent-based models, defined as geographical objects (with spatial relationships) only, with the CA model. And so, it is suggested that these agents as geographical objects would be designed to associate to the cell resolution corresponding to that of the CA model. In the current research, this association of objects as cells is assumed over the resolution of the CA model. However a detailed study can point out the implications of scale for such association of objects as cells. In the present context, these agent-based models are attributed the spatial relationships defined over a cellular space representing discrete entities of the phenomenon. Such agents defined over a cellular geographic space are also conceived with the typical properties of the agent-based models of being deliberative or reactive and importantly autonomous. The diffusion of agent action on cellular space is discussed in the subsequent paragraphs.

3.3. Agent-based Models and Simulation Time

Apart from the definition of a spatial relationship to these automata, the notion of time perceived by the simulations with respect to the agent-based models is discussed. Normally, the physical system or process in question is modelled with respect to the physical time. In this context, if the urban sprawl is modelled to predict the sprawl annually, the physical time considered is in terms of 'years'. A simulation time is defined as a totally ordered set of values where each value represents an instant of time in the physical system being modelled. On the same lines, if sprawl is modelled from 1990 to 2000 as the physical time, the same in simulation time would be in terms of t_i where, *i* =1 to 10. In the CA paradigm, if the transition rules are defined for a certain period (say, $t_1 \bullet t_i \bullet t_{10}$), the simulation takes place in iterations with respect to the simulation time of $t_{1,}$ $t_{2,}$ $t_{3\ldots}$ t_{10}. The agent-based models would be defined based on the process of interest. In reality, the behaviour of such models need not be modelled for the same simulation time of the CA. Thus, in this research the simulation time for the CA and agents are assumed to have a single time advancement mechanism (Figure 3.2). However, each of the agent-based models could be defined separately with different time advancement mechanisms depending upon the events for which they are defined, as discrete-event models. Then a significant issue would be to address the synchronization of these models with different time advancement mechanisms, this however is beyond the current scope of research.

3.4. Formalism of the Agent-based Cellular Automata

A formal definition of the ABCA is deduced from the traditional CA transitional rule mentioned in Equation 1. It should be noted that Equation 1 accounts for the state, *s*, and neighbourhood, N, of the cell at time, t, to define the transition to the state at time, t+1. Equation 1 accounts for the discrete time-stepped simulation of the entire region from time t to t+1. The agent automata that are suggested are those, which can be as many in number for the region with varied spatio-temporal characteristics. The variation in space is to denote the sphere of activity or influence of the agent automata in question and the temporal variability indicates the discreteness of the agent automata and the different start and end time of the agent automata. In this regard, the agent automata are considered as distinct discrete-time simulation types. On the lines of a CA transition rule, an agent automaton is defined:

$$AA \approx \left[K \bullet \left\{f(A) \bullet \mathrm{E}_{x,y,t}\right\}\right] \quad \text{...Equation 2}$$

where,

K is the type of agent

$f(A)$ is the agent-based model driving the transition of the cell

$\mathrm{E}_{x,y,t}$ is the extents of the agent-based model spatially and temporally

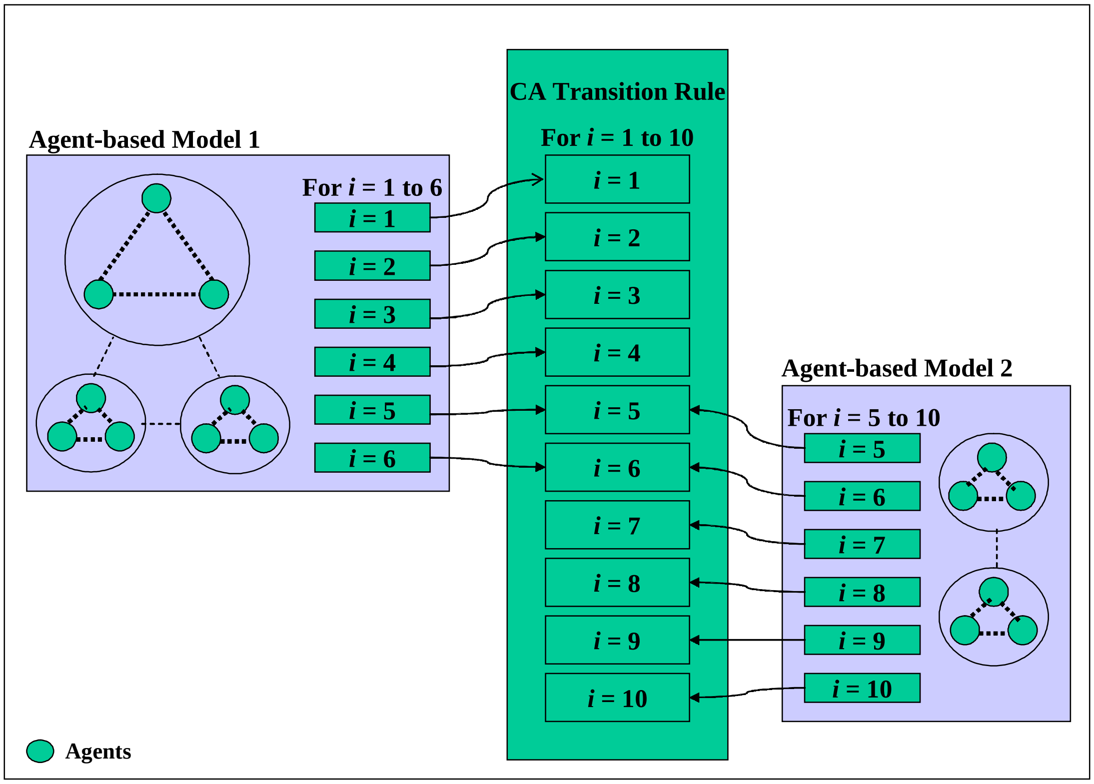


Figure 3.2: Time Advancement Mechanism for CA and Agent-based Models

The type of agent governs the behaviour of the automata accordingly whether the agent is reactive or deliberative. This condition of the automata would enable it to act autonomously. Further, the function $f(A)$ defines the behaviour for the agents in the agent-based model. This behaviour will be responsible for initiating or responding to the transitions of cell states. The extents of the agent automata are defined to streamline the specific sphere of activity or influence of the agent automata within that space and time.

In the real sense, only those agent automata types would be defined the modeller is interested in and so there can be a case where in certain cell space and over a certain time there might not be any agent automata accounted. And so, even in such cases, the CA transition would any way account for such instances. However for effective overall process simulations, both CA and agent automata have to be combined effectively to simulate each discrete time-stepped models at their respective scales of space and time. In this context the unification of CA and agent automata is termed as Agent-Based Cellular

Automata (ABCA). A formal definition of an ABCA is given as a set of all the individual agent automata and the general CA by the following:

$$ABCA \approx \left[f\left(s^t, N\right) \bullet \left\{ (AA)_1 \bullet (AA)_2 \bullet (AA)_3 \bullet (AA)_4 \bullet \Lambda \ n \right\} \right]$$

$$\Rightarrow ABCA \approx \left[f\left(s^t, N\right) \bullet \prod_{x=i, y=p, t=m}^{x=j, y=q, t=n} (AA) \right] \qquad \text{...Equation 3}$$

where,

$(AA)_1 \approx \{K \bullet f(A) \bullet E_{x,y,t}\}_1$; $(AA)_2 \approx \{K \bullet f(A) \bullet E_{x,y,t}\}_2$; $(AA)_3 \approx \{K \bullet f(A) \bullet E_{x,y,t}\}_3$; etc.

3.5. Agent Action in Cellular Space and Discrete Time

The kind of agent-automata described are those, which can initiate, respond and react to the dynamic conditions prevalent in the system in question. It is assumed that a CA transition rule would act globally over the region under question. However, the utility of engaging the agent-automata is to enable the local spatial interaction of the drivers of the systems, initiate transitions based on system variables attaining certain levels of threshold, and respond to the interactions among the different agent-automata. In such scenarios these agent-automata would diffuse across the cells during that simulation time for which the agent- automata are defined. Thus if the cell state at C (x_i, y_j) in the system is to be updated iteratively by the CA, then according to equation 1:

$$C^{t+1}\left(x_i, y_j\right) = f\left(s^t, N\right)_{i,j}$$

In the instances of agent-automata defined for the system in question, these agent-automata are to diffuse into these transition rules to initiate, interact and respond to the cell state transitions. Assuming the agent-automata are defined for the system in question, in the ABCA framework the updating of the cell states would be as follows:

$$C^{t+1}\left(x_i, y_j\right) = f\left(s^t, N\right) \bullet \prod_{i=1, j=1}^{i=m, j=n} (AA)$$

The agent-automata would act at those cell spaces or extents and during the time for which they are defined. In the absence of any activity for the agent in certain spatial extents and over certain simulation time, the actions of agent-automata would act as an identity element (in effect there would be null updates from the agent-automata). The set of all the agent-automata and the CA can be part of the simulation application, while the simulation executive would execute these automata for the simulation time.

With the unification of agent automata and CA, the imminent problem foreseen is to achieve the synchronized update of the cell states, initiated by both of these. For achieving the synchronization of the different individual automata and the CA, the concept of having a geo-spatial analyser is suggested. The description of the functionalities of the geo-spatial analyser (GSA) is dealt in the subsequent sections.

3.6. Scope and Limitations

The unification of agent automata and CA offers opportunities for geo-spatial modelling and simulation. With the available spatial and temporal data, satellite remote sensing and the capabilities of geo-information processing through GIS, the technique of modelling spatial phenomena using CA was already mastered as can be inferred from the various investigations. The limitation of CA being not able to respond to drivers and to various externalities dynamically are now being addressed through the agent-based approach which upon unification with CA would now enable the geo-spatial modellers to plug-in any process model as an agent and apply it within the ABCA framework. The ABCA is limited by the consistency of the input data sets and the type of relationships, which are modelled amongst agents. The key conditions for the integration of these agent-based models with CA models are that the spatio-temporal extents of these models/processes in question are to be predefined. By ensuring such a measure of predefining the spatio-temporal characteristics of the models, GSA would schedule these automata while the HLA offers synchronization of different models. It is believed that this framework is more robust to tackle, analyse, test and evaluate the different geo-spatial processes dynamically at discrete space and time than the Geographic Automata System (GAS) framework (Benenson and Torrens, 2003).

3.7. The High Level Architecture

Processes in environmental systems take place at their own respective scales in both space and time. The geo-spatial community has for long until now attempted to model and simulate these different processes on the desktop aided by the advances in the computational capabilities. A significant challenge faced by the geo-spatial community is the need of dynamic systems for modelling and simulation of the spatio-temporal and thematic databases in real-time. Further, as there are numerous processes taking place in space and time within an environmental phenomenon in question, a typical GI system should be capable of integrating or synchronizing the different models of environmental processes at their respective scales.

In the modelling and simulation domains of computer science and engineering numerous advances are made with respect to distributed and parallel simulations. A key distinction is that the modelling and simulation community is able to achieve the synchronization of time-variant simulation models by using a standardized simulation framework, the High Level Architecture (HLA) initially developed by the United States Department of Defence and now an IEEE standard (1516) (Dahmann, Fujimoto and Weatherly, 1998; IEEE, 2000). Wilcox, Burger and Hoare (2000) present a review of the history, evolution, current and future developments concerning the advanced distributed simulations (ADS) notably by the emergence of HLA.

The HLA is the architecture for reuse and interoperation of simulators. The intent of the HLA is to provide a structure that will support reuse of capabilities available in different simulators, ultimately reducing the cost and time required to create a synthetic environment for a new purpose and providing developers the option of distributed collaborative development of complex simulation applications. Thus the HLA offers a federation approach and addresses the interoperability and reuse of individual components. In order to facilitate interoperability, each member (federate) of a distributed simulation (federation) is equipped with appropriate interfaces to interact via a run-time infrastructure. The HLA does not prescribe a specific implementation, nor does it mandate the use of any particular software or programming language. Figure 3.3 shows the functional view of a HLA federation. A detailed description of the HLA is given in Appendix 1.

Carothers et al., (1997) present the design and algorithms used to implement the HLA Time Management Services in the Run-Time Infrastructure (RTI) component of the HLA. Fujimoto (2003) addresses how the two key categories of time management algorithms, conservative and optimistic synchronization are supported in the HLA. The paper discusses in detail the issues of time management in the distributed event simulations concerning the synchronization of computations on different processors.

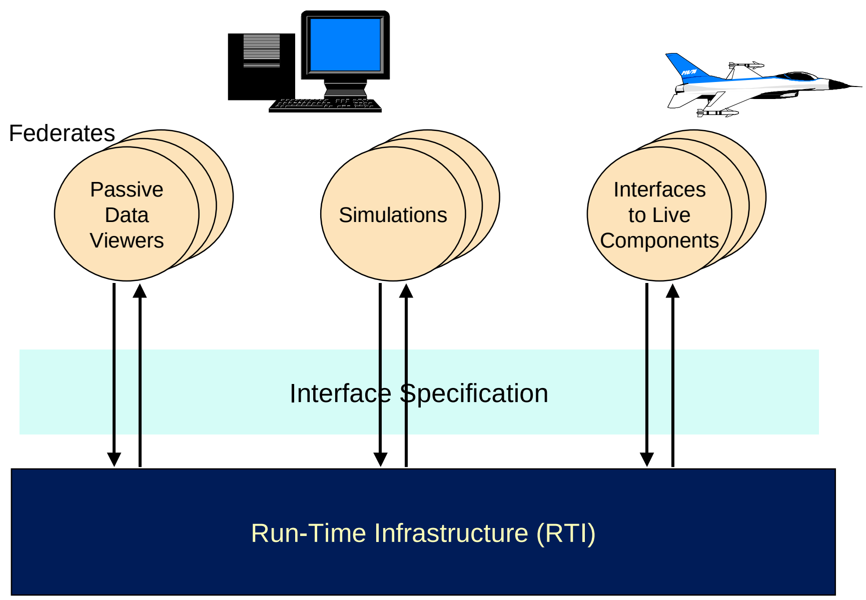


Figure 3.3: Functional View of an HLA Federation (Dahmann, Fujimoto and Weatherly, 1998)

On similar lines with respect to time synchronization, Tacic and Fujimoto (1998) describe a mechanism to realize properly synchronized data distribution in distributed simulations using logical time. This is achieved using a connection look-ahead approach that allows dynamic network topology changes where federates can advance further away from each other, yielding the better performance.

With the development of HLA and subsequent notification of the same as an IEEE standard (1516), the simulation framework has seen tremendous scope in the modelling and simulation community moving from the defence sectors to application areas of public domains. Borshchev, Karpov and Kharitonov (2002) develop approaches to model and simulate hybrid systems in a distributed manner, where hybrid state machines are used to model complex interdependencies between discrete and continuous time behaviours. Lindenschmidt, Hesser and Rode (2004) present an interesting application of using the HLA to integrate various water quality models, in which the simulations of a hydrodynamic model, eutrophication model and sediment and micro-pollutant transport model are interlinked and coordinated by the HLA RTI environment.

Since the development of HLA, the modelling and simulation community are using the framework extensively to achieve interoperability amongst heterogeneous simulators. However a notable drawback with these is that they lack the spatial component in them (Schulze et al., 2002). A possible bridging of the concepts of HLA into the geo-spatial domain is attempted recently with a view of providing a set of respective web services. In this regard, Simonis, Wytzisk and Streit (2003) suggest and describe an interoperability framework to bridge the High Level Architecture and the Open Geospatial Consortium (OGC) specifications. The thrust is on providing simulation-born spatio-temporal data in an interoperable manner facilitating benefits for a wider user community at large.

Apart from the application of HLA for integration of different simulation models elsewhere in the simulation community, within the GIS another framework is suggested based on the emerging Open GIS standards, which will allow the integration of parallel computing technology such that it becomes a viable component of a new generation of geographical information system software (Dowers, Gittings and Mineter, 2000). Ammerlhan et al., (2000) address the integration of geographically distributed functions with interactive human control where most of the components are distributed which is achieved through a coordination of the simulation through a new parallel discrete-event simulation system. Their solution is provided through a Java-based distributed discrete event simulator, called infrastructure for distributed enterprise simulation (IDES).

Thus, numerous advances are made within the modelling and simulation community with little focus on enabling the spatial aspects for synchronisations of different models. The direction of the current research is to fill the gap of spatial synchronisation for these models for effective simulations within the geo-spatial domain retaining the HLA framework for synchronising time variant simulations.

3.8. The Spatial and Temporal Synchronization

3.8.1. Geo-Spatial Analyser (GSA)

Spatial scheduling refers to the ordering of simulators for execution based on the spatial extents defined. To enable the spatial scheduling of the various simulations, it is suggested to incorporate a "**Geo-Spatial Analyser"**. Each of the discrete entities of the phenomenon being modelled as an agent-based model is considered as a discrete time-stepped simulation. These models would have specific spatial and temporal extents. A spatial extent for a specific model would imply that the agents would be acting in that specific spatial extent, as well as a temporal extent would also be necessary. In such cases, the duration of the process would be required for executing different automata over time. Mostly, the temporal synchronization of distributed simulators could be addressed by the HLA. In the instances, where there are two or more simulators that need to be executed over different time, HLA offers to synchronize these simulators. But, if the simulators would be acting over the same spatial extent, at the same time, now synchronized by HLA, there can be a conflict for which simulator execution would be first required to have the implication on the spatial extent. With a GSA, this can be resolved. Thus, the key functioning of GSA would be to facilitate different simulators to take place at their specified locations based on the spatial extent of influence of the agents. To enact such a functioning, the GSA would be made up of two components, a Simulation Control Service (SCS) and the Geo-scheduler. The functionalities of the SCS is similar the one suggested by Schulze et al. (2002).

The SCS would act as an interface between the different simulations at two levels. In the first level, SCS would act as the interface for the *geo-scheduler* to schedule the different simulations and then in the next level the SCS would actually be controlling different distributed simulations in terms of initiating and ending different federations.

During the execution of the model for every simulation time, the *geo-scheduler* would analyze the spatial extents of the different agents to ascertain which agent is influencing over which spatial extents. In the case of overlapping of the spatial extents (say a process taking place at the same location), the *geo-scheduler* would then evaluate the spatial extents of these agents and schedule accordingly based on the agents with least spatial extent first followed by the agents with higher spatial extents. Suppose there are several agent automata defined for a geographic phenomenon as in equation 3. The key functionality of the GSA would be to schedule the automata as given by the following during the iteration:

$$GSA \Rightarrow \lfloor (AA)_{i,}, (AA)_j, (AA)_k, \ldots (AA)_n \rfloor \; \forall i \in x, y \,\&\, i < j < k < \ldots < n$$

These scheduled simulators would then be directed to the SCS. Based on the evaluation, supposing few agents influencing at the same place and time, these federates would run interacting with the run-time infrastructure to undertake the overall simulation.

Functionalities of the GSA

- When simultaneous simulations are initiated, all simulations are directed to the Simulation Control Service (SCS), which would act as an interface with the *geo-scheduler*.
- The SCS in the first level would address all the simulations to the *geo-scheduler*.
- The *geo-scheduler* analyses each of these simulations at runtime for any spatial overlaps evaluating the spatial extents of individual simulations.
- Based on the spatial extents, the *geo-scheduler* of the GSA schedules the simulations into different *federates* through the SCS of the second level.
- The SCS in the second level deploys these federates by starting, controlling and destroying the federates.
- The GSA assumes a bottom-up approach, wherein the underlying assumption that all individual local interactions lead to the global variations. Thus the scheduling of different simulations would be based on the least spatial extent first followed by the higher spatial extents.
- When two or more simulations are waiting to be run within the same spatial extent, it can also be that certain agent-based models are to obtain inputs based on certain specific earlier simulations in time. The agent type defined in the agent automata and the function specified will be able to handle such scenarios.
- The scheduled simulations which form the federates through the SCS of second level that are synchronized for different time-variant simulations by the runtime infrastructure.

The overall framework for the simulations is shown in Figure 3.4.

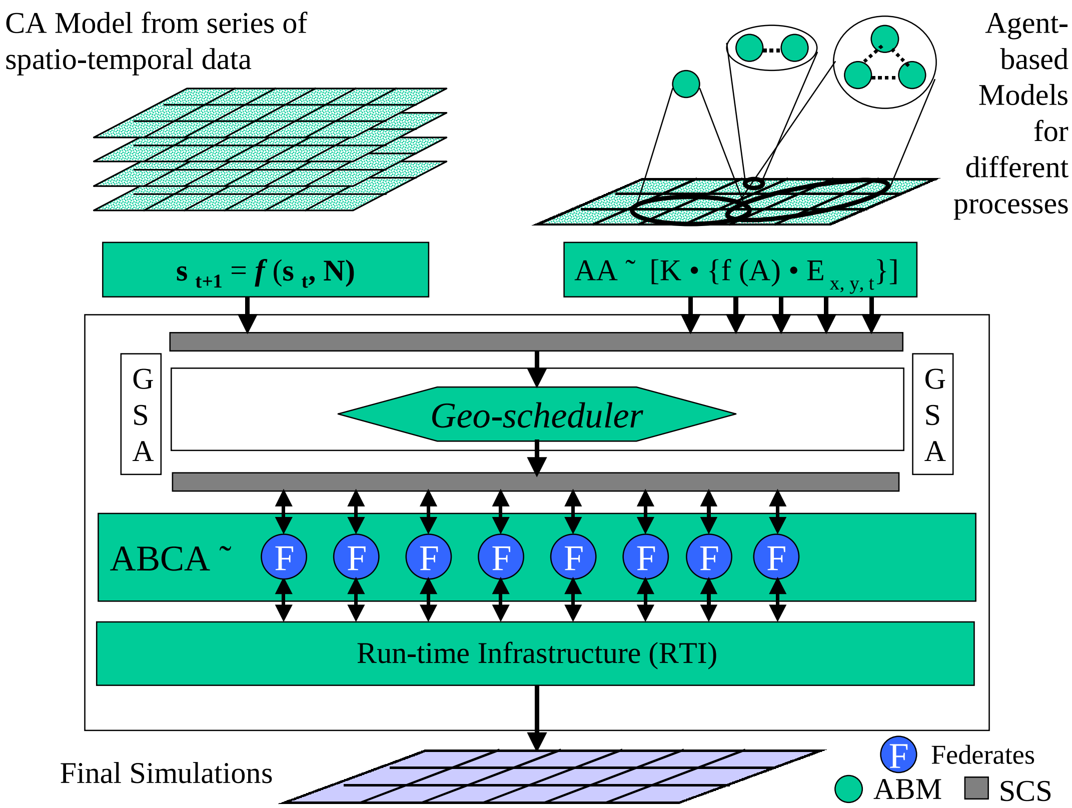


Figure 3.4: Framework of the overall simulations integrating the CA model with the agent-based model using the HLA framework and the GSA

ABM: Agent-based Model; SCS: Simulation Control Services

3.9. Application Scenario

To illustrate the applicability of this framework, an example from the urban sprawl dynamics is considered. Considering a city with some spatial extent, the land-use change for a period of time is modelled using CA. Further different processes like the infrastructure and population growth can take place over the region. An agent-based model can be thought of to represent the addition of new infrastructure, like the construction of a new ring road. This, when considered as an agent-based model, would have agents in it to initiate cell transitions along the specified areas (5 km – linearly) representing the road over duration of time, such as 6 years. Along with it is another agent-based model, the population development process, acting in the same region also during the same decade. The agents of the agent-based model (population) would initiate transitions like increasing the built-up area for every 'n' number of people increased and 'm' number of density crossed. To illustrate, how the interaction with the CA can take place, the same scenario of ring road development and population growth is described with reference to the time advancement as shown in Figure 3.2. If you consider the agent-based model 1 in the Figure 3.2 as the ring road development and agent-based model 2 as population development process, the interactions among these would be as follows. The agents of the ABM 1 would initiate the transitions for the change in state in the specified region from the other land use class to road from the iteration 1 to 6 sequentially over the years. Consequently, wit the development of a ring road, other processes like the establishment of industrial areas, commercial places or the residential areas can come up around the vicinity of the region. This would ideally start during the iteration 5 and go on until 10. The agents of these can represent the industrial layouts, commercial establishments, or residential layouts, implying that in the cell states surrounding ring road would eventually be built-up. With in the sub-model of an industrial layout, different agents can represent each industry in that layout. Even though the cell states pertaining to the layout might eventually be converted in to built-up from other land uses in iterations 5 to 10, certain agents representing individual industries can initiate these transitions from iterations 5 to 10; there by collectively the land use is converted into developed. With the development of ring road and other development activities, the population would increase; the agents representing the same can efficiently do this task. Each of the agent action is one simulation of the model. However, only with the establishment of each industry first, the industrial layout is complete. It would be inappropriate representation for the agents to first represent the industrial layout and then the industries as such. In such scenarios, the GSA will play a significant role. The same analogy holds good for the houses and the entire residential layouts. In such scenarios, the role of GSA would be to evaluate the spatial extents of these two agent-based models and then utilize the HLA framework to synchronize the overall simulations, and ultimately facilitating the synchronisation of the individual processes in the overall simulations.

3.10. Interoperability of the ABCA Framework

The ABCA framework is a combination of the domains of GIS, simulation (HLA) and the artificial intelligence (agent-based modelling) for ensuring the geo-spatial simulations of dynamic physical processes. An ideal framework for simulation should be reusable and interoperable for different use cases and multidisciplinary research teams thereby minimizing the cost of effort and time for modellers to rebuild a simulation framework specific to their use case or research problems.

To promote the interoperability of solutions and applications for geo-spatial services, data and applications, the Open Geospatial Consortium has established the necessary standards and specifications (OGC, 2003). For the simulation applications, the High Level Architecture developed by the US Department of Defence is now established as an IEEE standard (1516) (Dahmann, Fujimoto and Weatherly, 1998; IEEE, 2000).The HLA is the architecture for reuse and interoperation of simulators.

Similarly in the agent-based modelling domain, the Foundation for Intelligent Physical Agents (FIPA) has been actively developing the specifications and standards to promote the interoperability of the agents and the services they can represent. Accordingly, FIPA has specifications categorized into few categories based on the Applications, Abstract Architecture, Agent Communication, Agent Management, and Agent Message Transport. The abstract architecture specification (FIPA, 2002) defines the architectural elements and relationships, guidelines for the specification of agent systems in terms of particular software and communications technologies and specifications governing the interoperability and conformance of agents and agent systems.

Even though, since the development of HLA, the modelling and simulation community are using the framework extensively to achieve interoperability amongst heterogeneous simulators, these lack the spatial component in them (Schulze et al., 2002). Very few have addressed the need for interoperability among the simulation models in geo-spatial domain and the time management of simulations bridging the OGC and the HLA (Schulze et al., 2002; Simonis et al., 2003). A possible bridging of the concepts of HLA into the geo-spatial domain is attempted recently with a view of providing a set of respective web services. Simonis et al. (2003) suggest and describe an interoperability framework to bridge the High Level Architecture and the Open Geospatial Consortium (OGC) specifications.

An important paradigm here is that there are separate architectures offering interoperability in the respective domains of GIS, simulation and agent-based modelling. However, for the proposed ABCA framework to promote interoperability, it has to comply on these architectures separately and not collectively. In this regard, Schulze et al., (2002) suggest the integration of two standardization techniques, the OGC and HLA, for offering features and services accessible to both basic architectures and provide a robust interoperable architecture for distributed web-based spatio-temporal simulation. Accordingly, Schulze et al. (2002) propose the DALI-Architecture (Distributed spAtio-temporaL Interoperability-Architecture). The DALI-Architecture is ideally suited for web-based simulations.

A comparison of these different architectures gives an outlook on their usability in the ABCA framework. The ABCA is inherently intended to comply with the OGC and HLA. An overview of the comparison of the different architectures is provided in Table 3.1. While the OGC has sense of space, the HLA has the sense of time and the agents the sense of the behaviours there in. The key aspect of the availability of the services is that for the OGC it is permanent and for HLA and FIPA they are during the runtime.

Table 3.1: Key Comparisons of OpenGIS, HLA and FIPA (Adapted from Schulze, et al., 2002)

Criteria	OpenGIS	HLA	FIPA
Domain	Space	Time	Agent-based Modelling
Applications	GIS	Simulation	Modelling and Simulation
Standardization	OGC	DoD, IEEE	FIPA
Temporal Awareness	No	Yes	Partial
Synchronization / Time Management	No	Yes, Extensive Interoperability	Not Exclusive
Spatial Awareness	Yes	No	No
Availability of Services	Permanent	During Federation Runtime	During Runtime

In the current research, integration of agent-based models over CA has been attempted in the geo-spatial domain and subsequently a framework has been proposed for ensuring the geo-spatial simulations. Thus, it is suggested that an integrated architecture combining the standards and specifications of OGC, HLA and FIPA would help in the realization of the overall goal of achieving the interoperability across the domains.

4. Application of Framework for Simulating Urban Sprawl

4.1. Dynamics of Urban Sprawl

Economic activities coupled with new infrastructure to meet the requirements of unprecedented population growth and migration into urban centres has resulted in urbanization. More and more towns and peri-urbans are blooming with a change in the land use along the highways and in the immediate vicinity of the city due to adhoc approaches in decision-making. This dispersed development outside of compact urban and village centres along highways and in rural countryside is defined as sprawl and is devoid of basic amenities consequent to the adhoc approaches in planning. Hence, dynamics of sprawl needs to be understood in terms of the rate and pattern of growth to plan basic amenities.

Urbanization is a form of metropolitan growth that is a response to often bewildering sets of economic, social, and political forces and to the physical geography of an area. Some of the causes of the sprawl include - population growth, economy, patterns of infrastructure initiatives like the construction of roads and the provision of infrastructure using public money encouraging development. The direct implication of such urban sprawl is the change in land use and land cover of the region.

Sprawl generally infers to the increase in built-up and paved area with impacts such as loss of agricultural land, open space, and ecologically sensitive habitats. Also, sometimes sprawl is equated with growth of town or city (radial spread). In simpler words, as population increases in an area or a city, the boundary of the city expands to accommodate the growth; this expansion is considered as sprawl. Sprawl also takes place on the urban fringe or the peri-urban region, at the edge of an urban area or along the highways.

The sprawl results in the growth of villages into peri-urban areas, peri-urban areas to towns, towns into cities and cities into metros. However, in such a phenomenon of development to have basic infrastructure, regional planning requires an understanding of the sprawl dynamics. However, in majority of the cases there are inadequacies to ascertain the nature of uncontrolled growth. Sprawl is considered to be an unplanned outgrowth of urban centres along the periphery of the cities, along highways, along the road connecting a city, etc. Due to lack of prior planning these outgrowths are devoid of basic amenities like water, electricity, sanitation, etc. resulting in inefficient and drastic change in land use affecting the ecosystem (Sudhira et al., 2004a).

4.1.1. Forms of Sprawl

Sprawl development consists of three basic spatial forms: low-density (radial) sprawl, ribbon and leapfrog development (Barnes et al., 2001). Radial sprawl is the consumptive use of land for urban purposes along the margins of existing metropolitan areas. This type of sprawl is supported by piecemeal extensions of basic urban infrastructures such as water, sewer, power, and roads. Ribbon sprawl is development that follows major transportation corridors outward from urban cores. In this case, lands adjacent to corridors are developed, but those without direct access remain in rural uses/covers. Over time these nearby "raw" lands maybe be converted to urban uses as land values increase and infrastructure is extended perpendicularly from the major roads and lines. The leapfrog development is a discontinuous pattern of urbanisation, with patches of developed lands that are widely separated from each other and from the boundaries, albeit blurred in cases, of recognised urbanised areas (Barnes et al., 2001). This form of development is the most costly with respect to providing urban services such as water and sewerage. Figure 4.1 shows the different forms of sprawl.

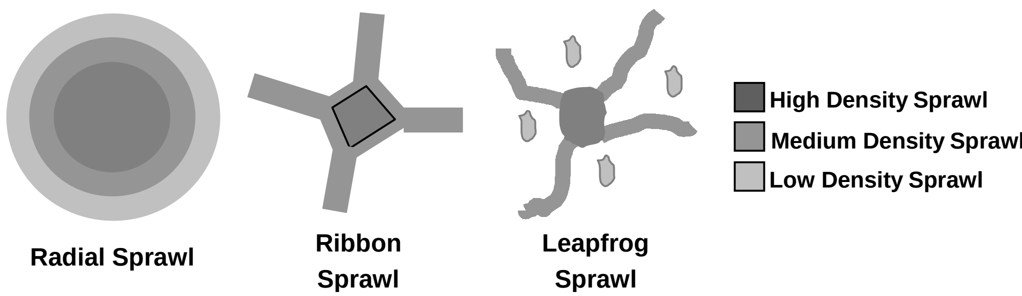


Figure 4.1: Forms of Sprawl

In industrialized countries the future growth of urban populations will be comparatively modest since their population growth rates are low and over 80% of their population already live in urban areas. Conversely, developing countries are in the middle of the transition process, when growth rates are highest. The exceptional growth of many urban agglomerations in many developing countries is the result of a threefold structural change process: the transition away from agricultural employment, high overall population growth, and increasing urbanization rates (Grubler, 1994).

Geo-informatics is very useful in the formulation and implementation of the spatial and temporal development strategies, which are essential components of regional planning to ensure the sustainable development. The different stages in the formulation and implementation of a regional development strategy can be generalized as determination of objectives, resource inventory, analysis of the existing situation, modelling and projection, development of planning options, selection of planning options, plan implementation, and plan evaluation, monitoring and feedback (Yeh and Xia, 1996). The geospatial techniques are evolving for implementation of such a proposed strategy. The spatial patterns of urban sprawl on temporal scale is analysed and monitored using the remotely sensed satellite imageries. The image processing techniques are also quite effective in identifying the urban growth pattern from the spatial and temporal data. These help in delineating the growth patterns of urban sprawl such as, the linear growth and radial growth patterns.

Prior visualizing of the trends and patterns of growth enable the planning machineries to plan for appropriate basic infrastructure facilities (water, electricity, sanitation, etc.). The study of this kind reveals the type, extent and nature of sprawl taking place in a region and the drivers responsible for the growth. This would help developers and town planners to project growth patterns and facilitate various infrastructure facilities.

Mapping and quantification of urban sprawl provides a "picture" of location of sprawl, type and patterns of sprawl, which helps to identify the environmental and natural resources threatened by such sprawls. Analysing the sprawl over a period of time will help in understanding the nature and growth of this phenomenon and thereby visualizing the likely scenarios of future sprawl (Sudhira et al., 2004a).

In the recent years, a lot of thrust in this field has been to understand and analyse the urban sprawl pattern. Various analysts have made considerable progress in quantifying the urban sprawl pattern (Theobald, 2001; Torrens and Alberti, 2000; Batty et al., 1999). However, all these studies have come up with different methodologies in quantifying sprawl. The common approach is to consider the behaviour of built-up area and population density over the spatial and temporal changes taking place and in most cases the pattern of such sprawls is identified by visual interpretation methods.

Defining this dynamic phenomenon with relative precision and accuracy for predicting the future sprawl is indeed a great challenge to all working in this arena. The study of urban sprawl (Sierra Club, 1998) is attempted in the developed countries (Batty et al., 1999; Torrens and Alberti, 2000; Barnes et al., 2001, Hurd et al., 2001; Epstein et al., 2002) and in developing countries such as China (Yeh and Li, 2001; Cheng and Masser, 2003) and India (Jothimani, 1997; Lata et al., 2001; Subudhi and Maithani, 2001; Sudhira et al., 2003). However there are very few studies on modelling urban sprawl are attempted (Subudhi and Maithani, 2001; Sudhira et al., 2004b). Although the notion of developed and developing countries is not crucial, the impacts of urban sprawl are ultimately on the regions land use mainly resulting in loss of prime agricultural lands and water bodies.

Although it is often considered endemic the phenomenon has impacts on the structure and growth of any city or town. Development of suburbs as a consequence of increased population growth and infrastructure facilities around the cities is a well-established reasoning for urban sprawl. Batty et al. (1999) have demonstrated the urban sprawl phenomenon as a spatially aggregate model. Various analysts have worked on urban growth modelling considering the spatial and temporal analyses of land use / land cover changes (LUCAS, GIGALOPOLIS, RESACC). However few analysts have made significant contributions on the urban sprawl dynamics (Batty et al., 1999; Torrens and Alberti, 2000).

The phenomenon of urban sprawl is very dynamic in nature. A complex of activities involving the economics, infrastructure addition, population growth and so on mainly attributes the urban sprawl dynamics. In this case, the different drivers are modelled as agents. The different drivers or the agents for inducing sprawl can manifest the sprawl by the complex of interactions and responses amongst them. The interactions among the agents of sprawl like the population or infrastructure can be complimentary to each other. Subsequently the reactions of these agent manifestations can lead to fuel further sprawl. Further, these agent-behaviours are not same with space and time; in effect these agents have an impact over system dynamically both in space and time. Visualizing such multi-scale space-time dynamic phenomenon like the urban sprawl is still not well handled by the traditional GIS (Batty, 2003).

Till date CA has been applied for simulating urban growth considering some of the driving forces responsible for sprawl. However some of the issues like the impact on socio-economic factors and policy matters are to be evaluated effectively while addressing issues like urban sprawl. For an effective simulation of the urban sprawl, modelling has to be attempted in both spatial and non-spatial domain. Most of the work in modelling the urban sprawl in the non-spatial domain is by the application of statistical techniques while CA models are known for modelling in spatial domain. In the recent years, the agent-based models in conjunction with the CA are used to represent the dynamics responsible for simulating scenarios of future urban sprawl (Chapter 2).

In this study the urban sprawl for a region is studied and modelled for the dynamics driving the sprawl using CA with agent-based models in the geo-spatial domain. The framework for integration of agent-based models over a CA model is discussed in the previous chapter. Before taking up the discussions on building the agent-based models, the abstraction of CA model is discussed. The agents are then integrated over the CA model.

4.1.2. Description of Study Area – The Mangalore City

This study was carried out in Mangalore, Karnataka, India (Figure 4.2). Mangalore, being a coastal city facing the Arabian Sea and the presence of a natural harbour, this has been one of the major ports of western coast of India. This region gains importance for the numerous financial institutions and hence has brisk economic activity as evinced from the growth of cities and infrastructure developments. In the recent years (post 1999), a lot of economic activity has been evinced with the boom in the information technology (IT) in the region. With the capital of the state, Bangalore, getting saturated with the IT investments and growth of software industry as such, IT players are looking at second order cities like Mangalore and Mysore for establishing IT industries. Mangalore, situated in the coast as a natural harbour, as the headquarters of Dakshin Kannada district and the historical lineage of intense economic activities as undoubtedly attracted the IT sector. With the advent of IT sector into this city the further sprawl is seen imminent and hence, the current study takes up the region for investigation.

The National Highway (NH) no. 17 passes through Mangalore. The region around Mangalore in the radius of about 8 km was chosen for thorough investigation (Figure 4.2). The total study area is 59 sq. km. The annual precipitation in this area is approximately 4242.5 mm in Mangalore. The southwest monsoon during the months of June to October is mainly responsible for the precipitation. The next round of precipitation occurs in the months of November and December due to the northeast monsoon. The relative humidity is considerably high mainly due to the proximity of the region to the coast. Mean annual temperature ranges from 18.6 ° C to 34.9 ° C (Census of India, 1981). The region of thorough investigation lies in between Longitudes 74.8224º E and 74.9155º E, Latitudes 12.8299º N and 12.9189º N.

4.1.3. Data Collection

The data collection involved the collection of multi-spectral satellite data Landsat MSS for the year 1973 from the Earth Science Data Interface of the Global Land Cover Facility and the satellite data Landsat TM for the year 1987 from the National Remote Sensing Agency, Hyderabad, India. The other data collected included the demographic details from the primary census abstracts of all the villages in the study area for 1961, 1971, 1981, 1991 and 2001, from the Directorate of Census Operations, Census of India. The village maps of this region were obtained from the Directorate of Survey Settlements and Land Records, Government of Karnataka.

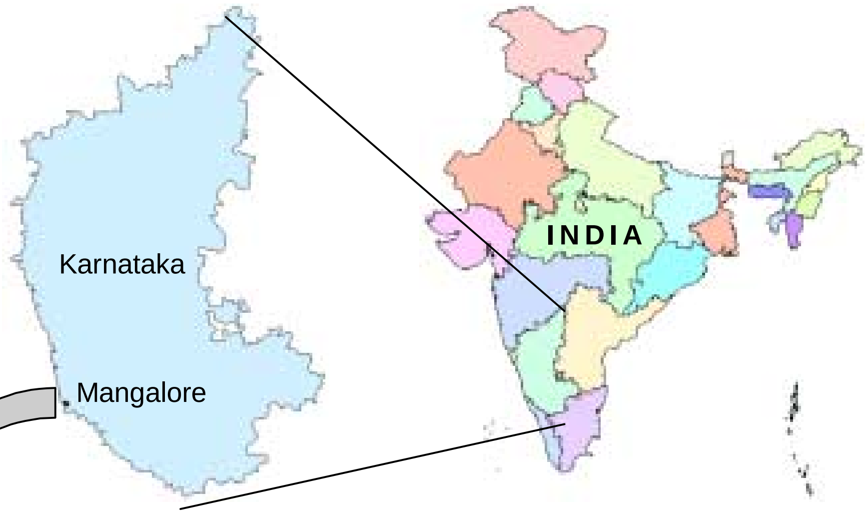


Figure 4.2: Location of Study Area, Mangalore City in India

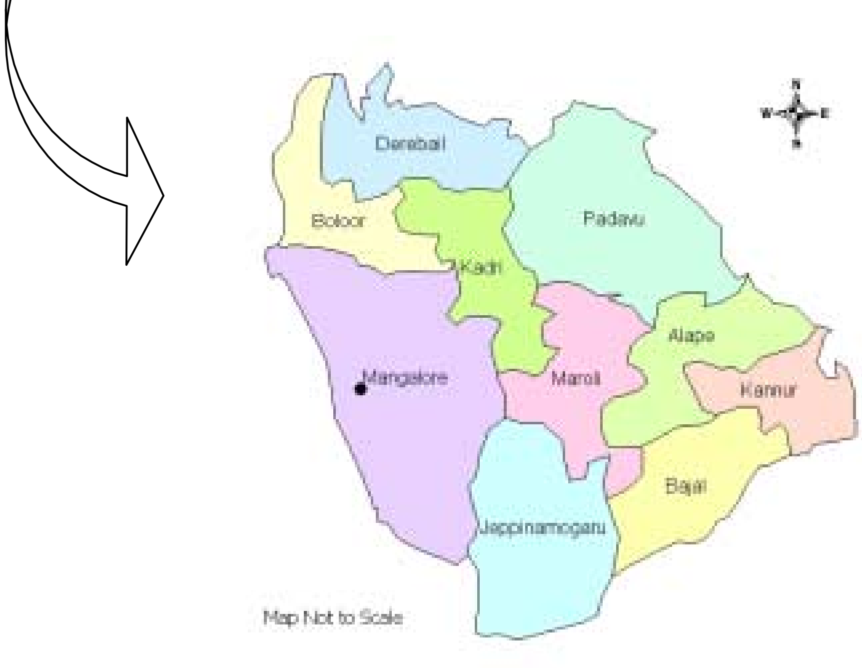


Figure 4.3: Details of Study Area

4.2. Abstraction of the CA Model for Urban Sprawl Dynamics

In order to build the transition rules for the dynamics of urban sprawl, the sprawl in the study area was analysed. The overall conceptualisation of the CA model is as shown in Figure 4.4. The urban sprawl in the region was initially quantified for the years 1973 and 1987, so as to analyse the change and the rate of growth. With the quantification of the urban sprawl from 1973 to 1987, modelling studies were undertaken considering the key driving factors responsible for the sprawl. This model was used to predict the sprawl for subsequent years, there by setting the demand for urban sprawl for the CA transition rules for allocating the different land cover classes into urban in the subsequent years.

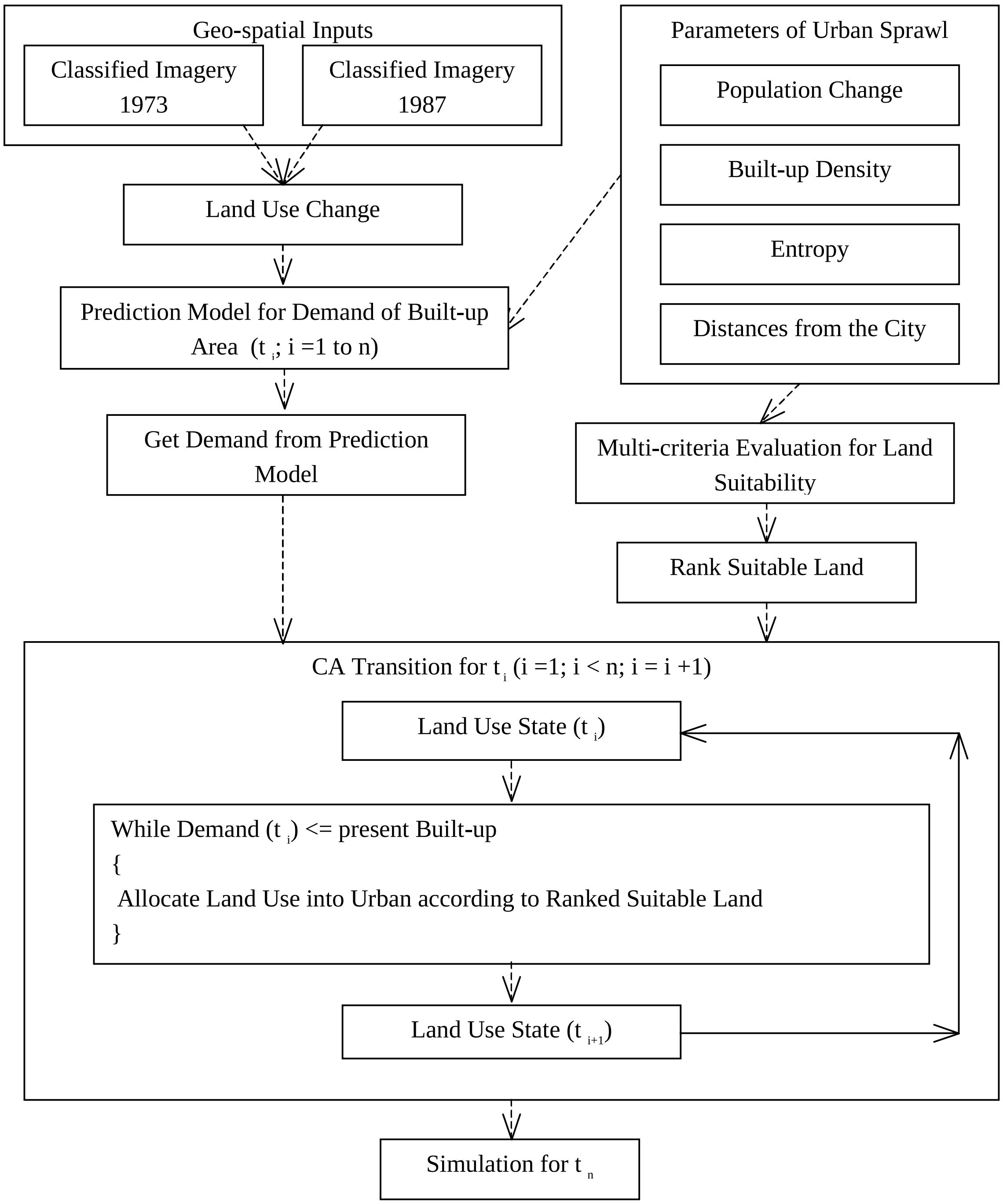


Figure 4.4: Conceptualization of CA Model

The cumulative transition rule for allocating the built-up was based on the demand set by the regression model and the suitability of the cells established by the multi-criteria evaluation and the neighbourhoods defined by the proximity of the cells to the city and the highway. The foregoing section discusses in detail on the analysis of urban sprawl and subsequent modelling of the same, establishing the suitability of the cells using the multi-criteria evaluation and the definition of neighbourhoods and constraints.

4.2.1. Analysing Urban Sprawl

Landsat MSS for 1973 and Landsat TM for 1987 of 70 m spatial resolution was used for analysing urban sprawl. The remote sensing data corresponding to Landsat MSS for 1973 was obtained from the Earth Science Data Interface (ESDI) of the Global Land Cover Facility (GLCF), University of Maryland and NASA. The Landsat TM data for 1987 was procured from the National Remote Sensing Agency (NRSA), Hyderabad. The multi-spectral data was analysed using IDRISI 32 (http://www.clarklabs.org). The image analyses included band extraction, restoration, FCC generation, enhancement and classification. Training data was obtained in the field using the GPS and accordingly training polygons were created along with corresponding attribute data, which was later verified with the composite image. Based on these signatures, corresponding to various land features, image classification was done using Gaussian Maximum Likelihood Classifier. The classified image showing the built-up of 1973 and 1987 are given in Figure 4.5 and 4.6.

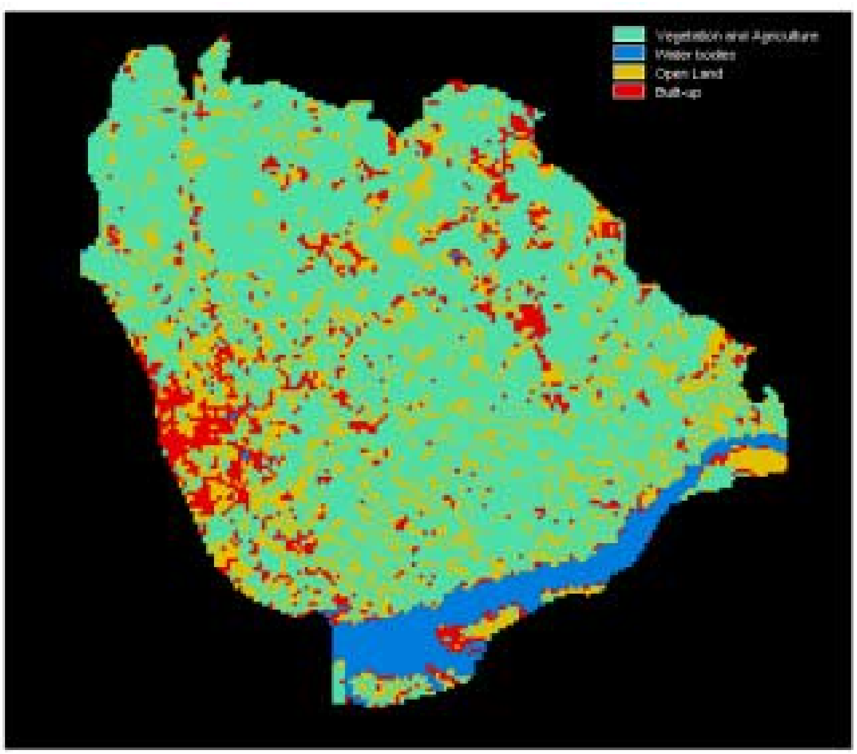


Figure 4.5: Classified Image of Landsat MSS 1973

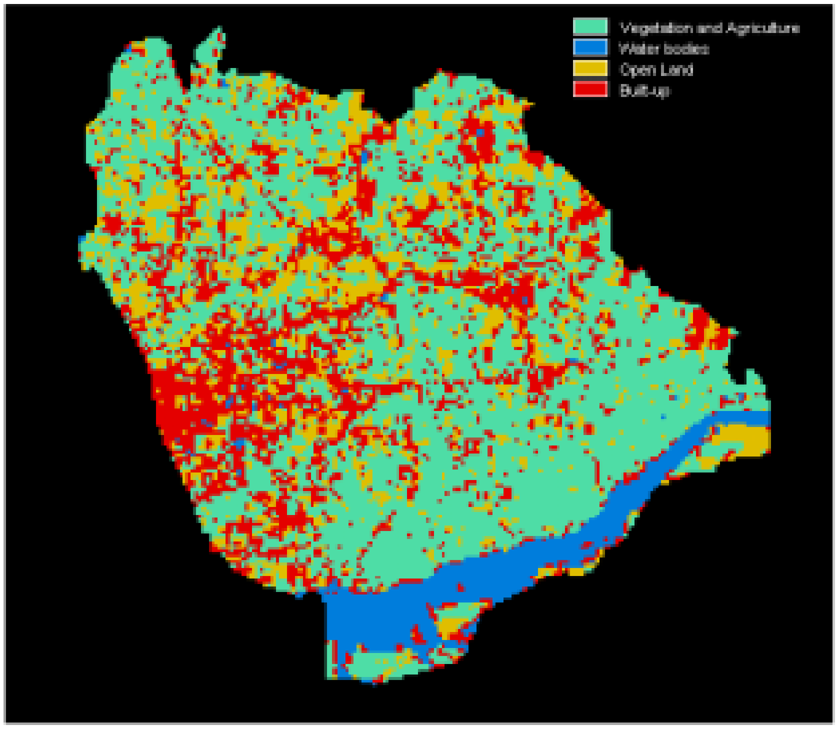


Figure 4.6: Classified Image of Landsat TM 1987

Area under built-up theme was recognized and extracted from the imagery and the area for 1973-1987 was computed. By overlaying the village/ward boundaries, village/ward-wise land use details were obtained. The distribution of village-wise land use of the study area for 1973 and 1987 are given in Figures 4.7 and 4.8 respectively.

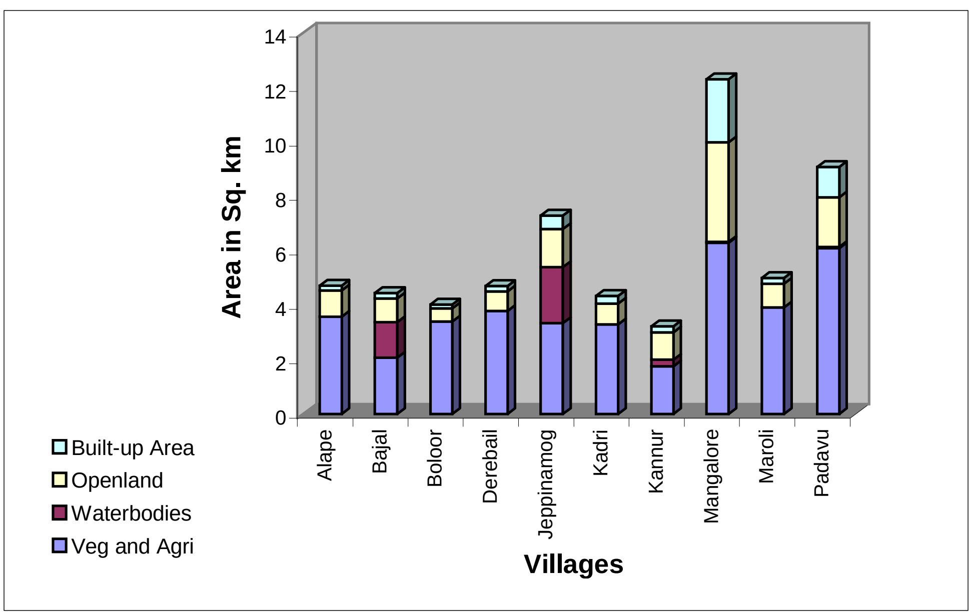


Figure 4.7: Land Use of Mangalore Region in 1973

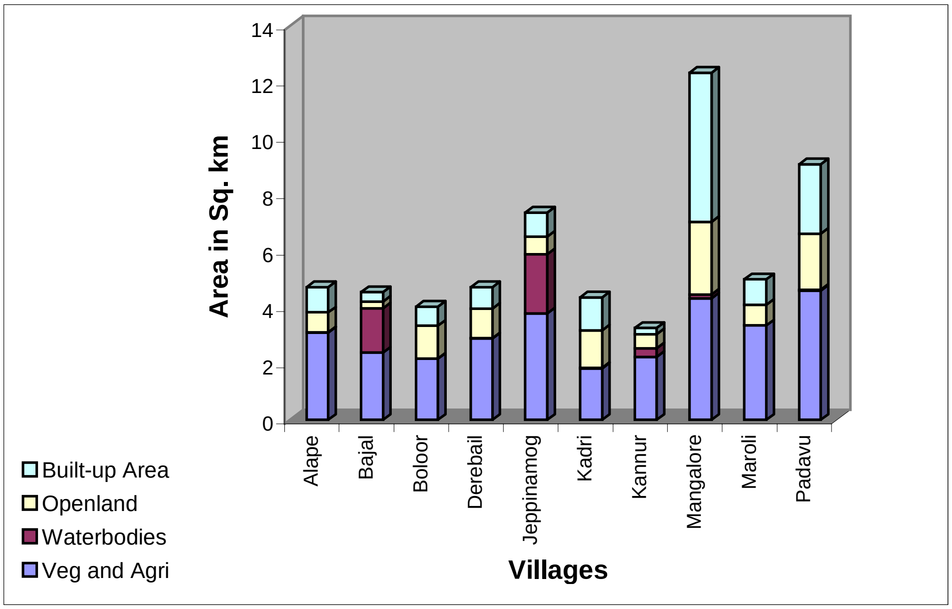


Figure 4.8: Land Use of Mangalore Region in 1987

The corresponding land use change in the region with respect to each of the land use classes are shown in Figure 4.9.

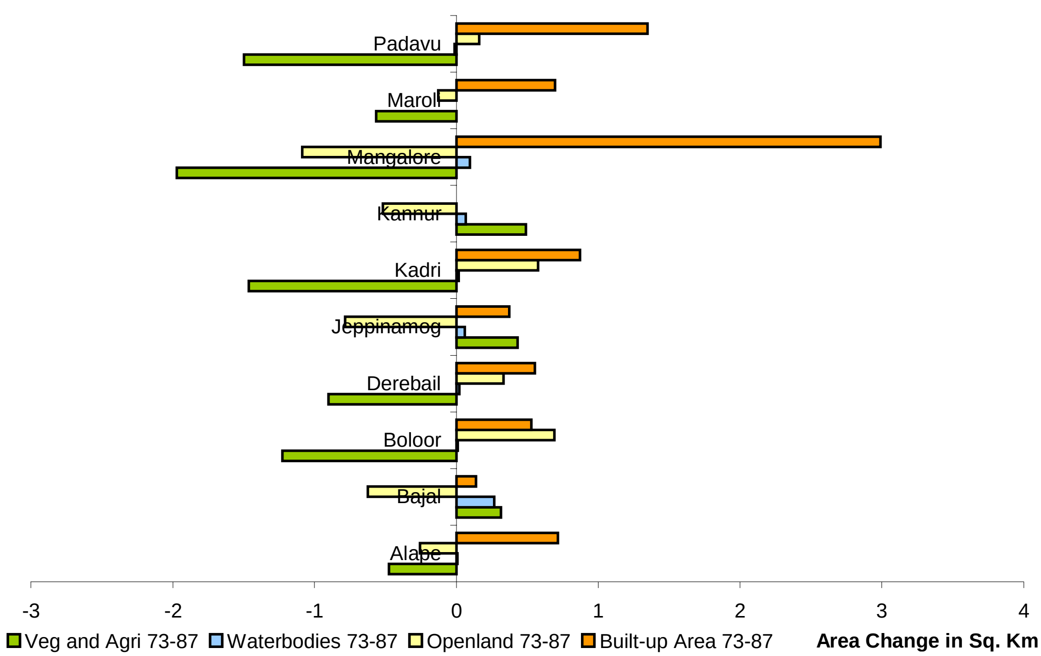


Figure 4.9: Land Use Change in Mangalore Region from 1973 to 1987

4.2.2. Population Growth and Built-up Area

From the analysis of the rates of development in the Mangalore region it was found that the amount of increase in built-up was nearly three times to that of the population growth. Thus, implying that even though the rate of population growth was around 50 %, the corresponding change in the built-up area was about 150 %. This indicated that more land was consumed per person, than what was being consumed earlier. The change in population in the region was found to be 70706 in 1987 as against 47424 in 1973, thereby attributing a percentage change of 49.09 %. However, when the built-up area was computed for the same duration, it was found that the built-up area had increased from 5.3473 sq. km to 13.5595 sq. km attributing almost 153.57 % change from 1973 to 1987 (Table 4.1). In sum, it indicated that for every increase in population by 1 %, the built-up increase was almost 3 %.

Table 4.1: Percentage Change of Built-up Area, Population, and Built-up Density from 1972 to 1987

Study Area	1973	1987	Percentage Change
Built-up Area (sq. km)	5.3473	13.5595	153.57 %
Population	47424	70706	49.09 %
Built-up Density	0.09054	0.22871	152.60 %

The per capita land consumption refers to utilization of all land for infrastructure and development activities like the commercial, industrial, educational, and recreational establishments along with the residential establishments per person. Since most of the activities pave way for employment and livelihoods, the development of land is seen as a direct consequence of region's economic development and hence it can be concluded that the per capita land consumption is inclusive of all the associated land development.

4.2.3. Modelling Urban Sprawl

The land use analysis of the region and the preliminary analysis on the population growth for the same time period as discussed in the previous section indicated some significant relationships with the change of built-up area to the population growth, and the population-built-up density. Population has been for long accepted as a key factor of urban sprawl. The spatial organization of the built-up area can be obtained by computing the Entropy for the built-up density. Yeh and Li (2001) and Sudhira et al. (2003) have documented the application of entropy in urban sprawl studies. Further, the urban sprawl phenomenon is influenced by the distance of the land from the city centre as evinced from following regressions model.

Application of regression techniques for analyzing urban growth has been attempted for establishing the relationships between urban growth and its drivers. Cheng and Masser (2003) report the spatial logistic regression technique used for analysing the urban growth pattern and subsequently model the same for a city in China. Sudhira et al. (2004b) also report the multivariate linear regression technique for analyzing the urban sprawl. In order to quantify and test for relationships with respect to the urban sprawl, the different drivers, which were considered for regression analyses, were:

- Population Change
- Built-up Density
- Entropy considering Built-up Density
- Distance from the City Centre

Initially, various regression analyses (linear, quadratic, exponential and logarithmic) were carried out to ascertain the nature of significance of the causal factors (independent variables) on the sprawl, quantified in terms of built-up. The linear, quadratic (order = 2), and logarithmic (power law) regression analyses were tried. All these regression analyses revealed the individual contribution by the causal factors on the sprawl. However, to assess the cumulative effects of causal factors, stepwise regression analysis considering multivariate was done. In the multivariate regression it is assumed that the relationship between variables is linear. The multivariate regression gives the cumulative relationship among the variables.

In order to explore the probable relationship of built-up (dependent variable) with causal factors of sprawl (population change, built-up density, entropy, etc.), multi-variable regression analyses were carried out to assess the cumulative effect of causal factors. The multivariate regression analyses reveal that all causal factors have a significant role in the sprawl phenomenon. The probable relationship for the built-up in subsequent year was:

Built-up = 0.08353 * Population Change + 24.3115 * Built-up1/VillageArea + 0.23898 * Entropy - 0.21114 * Mangalore Distance - 0.2153 … Equation 4

$r = 0.9664$; $p < 0.01$; $SE_{y\ estimate} = 0.5224$

4.2.4. Prediction of Urban Sprawl

It is evident form the above relationship that the subsequent built-up has a significant relation with the drivers identified. Subsequently, the likely increase in the built-up for each year up to 1999 was predicted using the Equation 4. For example, to project built-up for 1999, corresponding population was computed considering annual growth rate based on the historical population data of 1961-2001 and the same was used in the equation to obtain the built-up area for the corresponding year.

The built-up area so obtained established the demand on land for conversion into built-up every year. This demand is utilized in the transition rules of the CA to allocate the land use into built-up in the subsequent years based on the suitability and constraints.

4.2.5. Multi-criteria Evaluation for Land Suitability

White et al., (1997) suggest the utility of assigning suitability of any cell for being allocated into built-up in the CA transitional rules. This concept has also been subsequently applied in various other CA models (Singh, 2003; Sun, 2003). On the similar lines, the multi-criteria evaluation was carried out for evaluating suitability of land cover to be allocated into built-up. The availability of suitable land is another prime factor for urban sprawl. The chief land use / land cover classification carried out for the satellite imagery includes, already developed (built-up) area, water bodies (sea, rivers, streams, tanks, lakes, etc), area under vegetation (agriculture, plantation, paddy, etc) and area under open land (barren, uncultivated, unused open land). The themes of already developed areas and water bodies' act as constraints for any further development. However, the probability of open land, barren land and uncultivated land for future built-up is very high. Similarly the possibility of area under vegetation to be converted into built-up as a result of sprawl cannot be ruled out. Apart from the prevalent land cover classes, the distances from the city centre and the roads are other important factors. Closer the distances from the city centre and the roads, the land has higher probability of becoming developed. Thus a sigmoidal monotonically decreasing fuzzy membership is assigned to weigh the factors of road and city centre distances.

Thus, each of the factors is evaluated for the suitability based on the criterion to each of the factor. As in the case of ordered weighted averaging, the suitability is defined by:

$$S = \sum w_i * x_i \qquad \textbf{…Equation 5}$$

where, S = suitability; w_i = weight of factor i (and sum of the factor weights = 1); and x_i = criterion score of factor

The key factors and constraints considered are given in Table 4.2. Thus, a final suitability image is obtained wherein each cell has the suitability value based on the above factors and constraints defined.

Table 4.2: Factors and Constraints

Factors	Land Use: Open land Distances from Roads Distance from the City Centre Land Use: Vegetation and Agriculture
Constraints	Built-up Water bodies

From the suitability images so obtained, the image is ranked according to the probability that which cells are having the higher probability to be converted from other land use classes into built-up.

4.2.6. The CA Transition Rules

The CA transitions are defined for the allocation of new built-up into other land use classes. The allocation is determined by the demand established by the statistical model. Further the allocation is also based on the suitability of the cells which area ranked after the multi-criteria evaluation for the suitability of the land. In one-iteration of the CA model, the simulation time is set to one year and so the model would allocate the land use into urban for the amount determined by the prediction model for demand also established for one-year simulation time. An underlying assumption is that only those land use classes other than urban would be considered for new urban areas implying that the urban growth will only increase and would never recede. Thus, when the simulation is executed, a set of output for the respective years would be depicting the future urban growth.

Thus the CA model for simulating urban growth is configured for the area under study by considering some of the drivers. With the basis of such a CA model, it can be seen that although the model can simulate for the future scenarios, the model is yet incapable of directly interacting with the drivers dynamically. For example, in the event that there is a sudden upsurge in population in one of the regions, the present configured model is incapable of addressing such scenarios. The transition rules of the CA model, so defined are inept to reflect the dynamic changes of the drivers by not interacting with them directly. Further the transition rules are system wide and so much of the details are generalized in the over all model. In order to address these inabilities of the CA model, the utility of the agent-based models in these scenarios is the crux of the research.

4.3. Agent-based Models in Simulating Urban Sprawl

In view of the above-mentioned shortcomings of the CA based models, the integration of agent-based models is attempted in this research to demonstrate that the agent-based models can act a synthetic interface to the dynamic drivers of the system and the simulation framework. Noting the key characteristics of the agent-hood, that they are autonomous, have social ability, are responsiveness, and pro-active, the same is infused into the framework discussed in the previous chapter. These agents are conceived to diffuse into the CA model and initiate transitions, respond, exchange and report accordingly to the properties associated with each of the agent-based models. Such properties are essential for these agents to facilitate the human-decision making involved in the simulation in more realistic sense. Thus agents of an agent-based model to initiate transitions would diffuse into the CA transition to enact such functions. Likewise, agents of those agent-based models to respond or react would act accordingly. On the similar lines, the agents can be conceived to represent the realistic decision-making undertaken apart from interacting with the drivers dynamically. The interactions among the drivers by these agents are simulated by means of a feedback loops with the CA transitions. The agent manifestations into the CA transition rule resulting the final transitions based on the feedback loops is as shown in Figure 4.11. The feedback loops are associated with the different agents and their behaviours for which they are modelled. Subsequently, the final transition rule gets the update from these agents before updating the cell state in the subsequent iterations apart from the inherent CA transition rules.

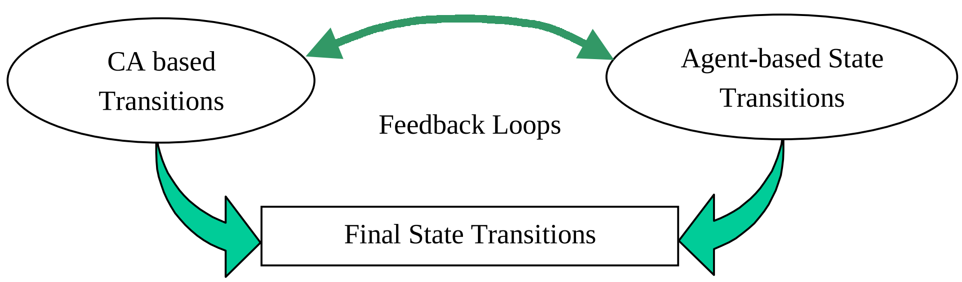


Figure 4.10: Feedback Loops between CA Transitions and Agent-based State Transitions enabling the Final State Transitions

An essential aspect to be addressed involving these feedback loops in a geo-spatial discrete-time models is the capabilities of these different models to represent the dynamics and respond to them at the respective spatial and temporal scales of the models. In sum, within the dynamics of urban sprawl, there are so many drivers that are system-wide and specific to certain localities, implying the variation in scales of their activity in both space and time. Each of the agent-based models representing the drivers is considered as a discrete-time stepped models, while the general CA being another discrete-time stepped model with similar time advancement mechanisms. Thus, while dealing with these different space and time variant models adequate care has to be ensured for the synchronization of such models. And so, for the current simulation, the framework developed and discussed in Chapter 3 is used to ensure the applicability of the agent-based cellular automata systems.

4.3.1. Agents in the Simulation Framework

Specific agent-based models are conceived for the use case of simulating the dynamics of urban sprawl. Among the key factors responsible for the dynamics of urban sprawl is the population of the region. In such case, the population development process can be considered as an agent that is driving the sprawl phenomenon. In the conceptualised CA model (Figure 4.4), even though the population of the region under investigation is considered, the model is not so dynamic enough to accommodate for any change in the population already fed in the model to be reflected in the CA transitions. Further, the population considered for the conceptualised CA model is computed annually based on the growth rates of the previous years. In spite of a built-up prediction model established to reflect the population of the villages in the study area, the typical CA model does not facilitate any interaction of these drivers. Consequently, the idea is to utilize the properties of the agent-based models as the representation of these drivers so that they can better reflect these properties.

Apart from the population development process driving the sprawl, the other drivers can be the infrastructure initiatives. The infrastructure initiative can be considered as an addition of a ring road around the city or demarcation of an area under different land cover as a residential layout. Such infrastructure initiatives are very typical in the fast and rapidly growing cities like Mangalore, also common in most of the similar cities of the country. The agent-based model conceived here representing the infrastructure initiative is for the ring road development process. In the typical instances of the ring road development around a city, the actual construction of the roads takes about two to three years. By the time the ring road is near completion, other allied developments like the establishments of institutions, industries, residential layouts in the vicinity are all imminent. Thus, an over all implication of such initiative in the due course of time, say about five years from the completion of the ring road, the region in the vicinity of the ring road will be more or less developed or the land use will be converted into urban. This above scenario is of an infrastructure initiative is conceptualised as an agent-based model which has multiple events resulting out of such activity and the pattern of growth varying over time. The initial creation of a ring road, which can be thought of as an addition of a poly-line feature around the city, results in the suitability of the land becoming into built-up with the completion of the ring road. Thus, the features (land use class) surrounding the poly-line (ring road), gets a higher suitability values over a period of three years also based on the prevalent land use classes and constraints. Eventually, all the land use classes surrounding the new ring road would become urban, subject to the constraints of the system. This process itself can represent one such agent-based model to initiate state transitions over time depending on the situations in relation to the prevalent conditions. In the same situation, different agents can also be defined to report the nature of growth taking place over time. With the possibility of defining all these processes as agent-based models, the agent-automata described for this scenario over a specific period of time, can be coupled in over all simulation, which would enable in the visualization of such ‘what if’ scenarios more effectively. The utility of involving the agents in the simulation of urban sprawl can be thought of as much as those only limited by the conceptualisation of the modeller and other resource constraints.

In this research the general scheme for building the agents in the geo-spatial domain suggested is discussed in the foregoing section while the implementation of the same is taken up in the next chapter. An alternate way of introducing the agents in the geo-spatial domain would be to enable the prevalent tools for agent-based modelling like the NetLogo, StarLogo, SWARM, etc. to interact with the geo-spatial domain. This may be achieved by programming these tools to communicate and exchange the geo-spatial data with appropriate interfaces within the simulation framework of the respective agent-based modelling tools.

4.3.2. Agents in the Geo-spatial Domain

In light of the above discussion of involving the agents in the geo-spatial domain, it becomes consequential that the spatial relationships or the topologies and specific geometry to these agents are defined so that they can be enabled to act in conjunction with the CA model. An ensuing approach in this regard was the creation of an interface for building the agents as geographic features, such as point, poly-line or a polygon thereby defining the spatial extents of these. The scheme of activity for such utility is as shown in Figure 4.12. In the next step the temporal dimension is attached to these agent-features so defined within the interface. Subsequently, the agent-behaviour in question or the model for the agent activity is established. Thus each of the agent-based models conceived for the region under investigation can be created on the fly before the simulation is executed. However, the agents are until now represented as geographic features, need to be acting in the cellular space, for which these agent-features are rasterised at appropriate scales. The scales at which they are rasterised needs to be calibrated so as to ensure effective synchronization and represent the realistic behaviour of the system in question. This interface to build agents along with their extents, duration and defining the behaviour is to act in concurrence with the geo-spatial data representing the different land use classes.

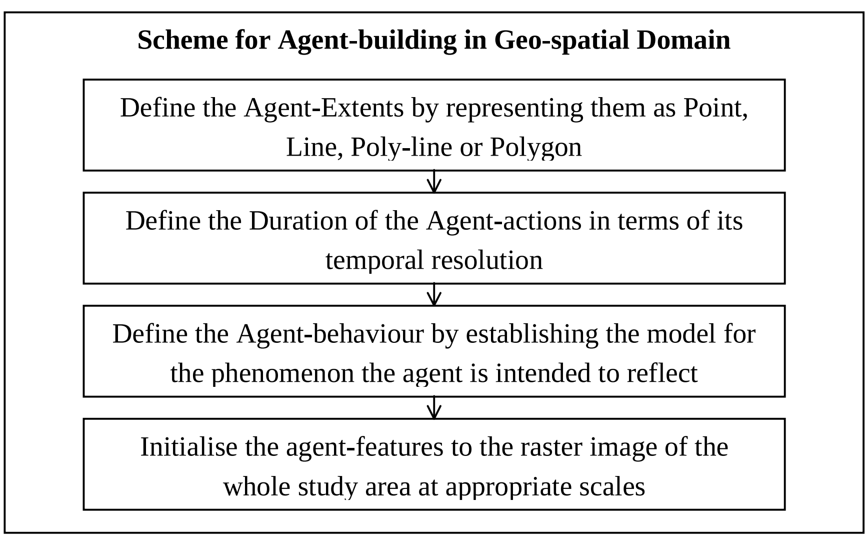


Figure 4.11: Scheme for Agent-building in Geo-spatial Domain

Each of the agent-based models thus defined are treated as one discrete time-stepped simulation system. The behaviour of different agent-automata over the simulation time is shown in Figure 2.2. The imminent task for ensuring the simulations is by properly scheduling the different models based on the spatial extents and synchronizing these time-variant models appropriately in the overall framework.

5. Implementation, Calibration and Validation of Simulations

5.1. Implementations of Agent-based Models for Simulating Urban Sprawl

Simulating a dynamic phenomenon like the urban sprawl using the agents in conjunction with the CA models was attempted in this research. The framework for ensuring the simulations developed in the Chapter 3 and the dynamics of urban sprawl studied and modelled using CA and agents in the Chapter 4 are implemented. Accordingly, there are two agent-based modelling approaches for simulating urban sprawl. In the first approach, much before researchers started to use the agent-based models in the geo-spatial domain, architecture and tools for building the agent-based models were developed (Chapter 2). Hence, much of the development and applications was restricted to the software engineering domain, until user-friendly tools for building agent-based models emerged in research and industry. However, with the application of agent-based modelling for geo-spatial processes, few researchers have attempted to fuse the agent-based models over a GIS (Batty, 2003; Benenson and Torrens, 2003). Using the first approach, NetLogo (Wilensky, 1999) (tool for building agent-based models) was used to demonstrate the capabilities of agent-based modelling techniques for studying urban sprawl dynamics. In this approach no actual geo-spatial data was used. Just a prototypical situation of a city exhibiting urban sprawl in radial direction was demonstrated. In the second approach, agent-based models operated on geo-spatial data and were applied for a region experiencing a high rate of urban sprawl. The ensuing sections discuss on the implementations of the same.

5.1.1. Agent-based Modelling Using the Tool – NetLogo

The agent-based modelling tool NetLogo developed by the Centre for Connected Learning and Computer Based Modelling, Northwestern University, USA was used to demonstrate a prototypical simulation of the urban sprawl in radial direction. In the NetLogo parlance the agents are conceived as ‘turtles’, the sense of state is through ‘patches’ and the worldview through the ‘observer’ (Wilensky, 1999). Patches are similar to the notion of cell in CA with the regular lattice structure. Each cell in the CA terminology corresponds to the patch in the NetLogo parlance. Thus, the notion of space is based on regular lattice structures of square cells and agents are simulated to move over a cellular space.

A prototype city was created by patches in the centre of a worldview of 29 x 29 cells. Several types of agent-hood (Jennings and Wooldridge, 1996), autonomy, social ability, responsiveness and pro-activeness are attributed to the turtles to operate in these patches. The developed – called ‘built-up’ – is attributed as one type of agent or turtle in this case. The other agents include available land for development as ‘land’. The built-up agent is represented by the graphic shown in red colour as a house, while the land agent is represented by the graphic shown in brown colour as a box over the patches. In order to establish a hypothetical city exhibiting radial urban sprawl, the model is initialised having a city with built-up as core area at the centre and the land with the scope for development spread surrounding the core region. For setting up the model or creating this instance of having a city in the centre and the available land around it, the agents in this context, built-up and land were defined appropriately. The built-up agents are programmed to populate in the centre of the monitor depending upon the size of the city. A monitor is the space where all the patches are visualized, which in this case is 29x29 matrices of cells. The size of the city is controlled by its ‘radius’, which is scaled from 1 to 5 cells from the core region. For testing the rate of development that may take up depending upon the various sizes of the city and the nature of growth, the amount of land to built-up percentage can also be initialised by the user. An option to set the rate of development as ‘built-up growth’ in terms of percentage growth for each simulation time can also be initialised. There are adequate monitors in the prototype model to indicate the advancement of simulation time, number of land agents and number of built-up agents. An initialised model for simulation is as shown in Figure 5.1.

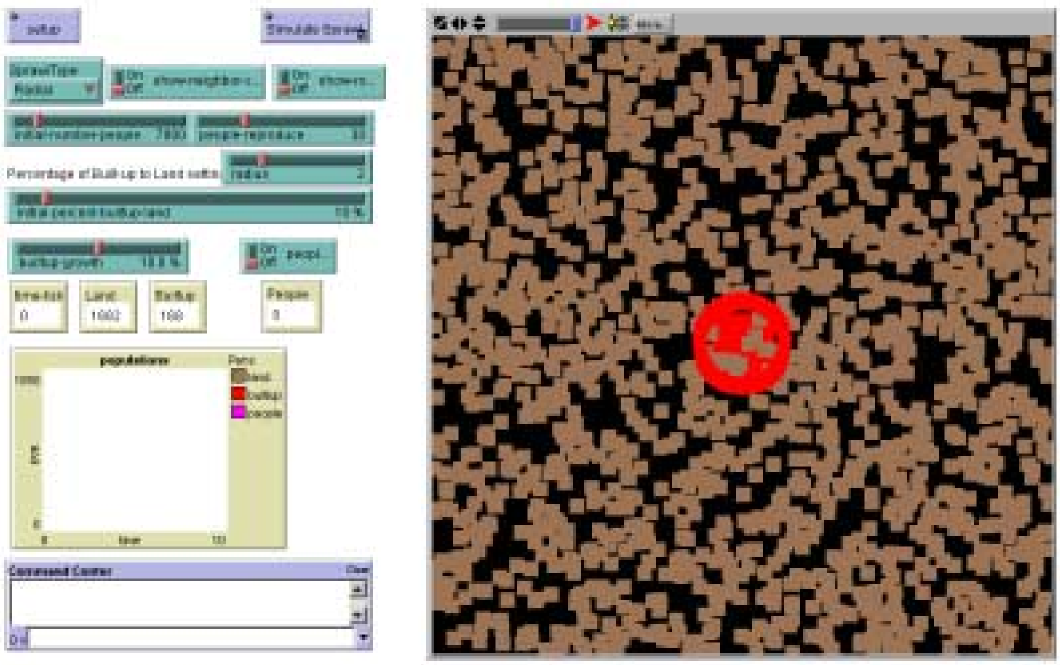


Figure 5.1: Initialised Model for Simulating Radial Urban Sprawl

In this model, the built-up agents are programmed so that they would populate or ‘breed’ in the agent-parlance if there exists a patch in its immediate neighbourhood with the ‘land’ agent occupied in it. The user can also set the radius of the neighbourhood from 1 to 5 cells using a 3x3 to 7x7 kernel. Within this predefined neighbourhood these built-up agents would be prioritised depending on whether the patch is already occupied by an agent. If a built-up agent has already occupied the patch, the agent would look for another patch. Thus the simulation of this model would go on until all the land is consumed in the neighbourhood.

The simulation output of the model demonstrating the radial urban sprawl is shown in Figure 5.2. From the model execution, various parameters like the nature of sprawl taking place and the rate of consumption of land over time are computed. From the model interface, it is evident that the agent actions can be continuously obtained in terms of their numbers over time. Further, for testing different scenarios and the implications of such scenarios can be easily visualised and quantified for various