PROGRAM ON HOUSING AND URBAN POLICY

CONFERENCE PAPER SERIES

C01-006

NEW TOOLS FOR SIMULATING HOUSING CHOICES

By

Paul M. Torrens

May 2001

UNIVERSITY OF CALIFORNIA, BERKELEY

New Tools for Simulating Housing Choices

Paul M. Torrens, Centre for Advanced Spatial Analysis (CASA), University College London. 1-19 Torrington Place, London WC1E 6BT. *Email:* ptorrens@geog.ucl.ac.uk. *Web:* http://www.geosimulation.com.

Abstract

There are indications that the current generation of models used to simulate the geography of housing choice has reached the limits of its usefulness under existing specifications. The relative stasis in residential choice modeling—and urban simulation in general—contrasts with simulation efforts in other disciplines, where techniques, theories, and ideas drawn from computation and complexity studies are revitalizing the ways in which we conceptualize, understand, and model real-world phenomena. Many of these concepts and methodologies are applicable to housing choice simulation. Indeed, in many cases, ideas from computation and complexity studies—often clustered under the collective term of geocomputation, as they apply to geography—are ideally suited to the simulation of residential location dynamics. However, there exist several obstructions to their successful use for these puropses, particularly as regards the capacity of these methodologies to handle top-down dynamics in urban systems.

This paper presents a framework for developing a hybrid model for urban geographic simulation generally and discusses some of the imposing barriers against innovation in this field. The framework infuses approaches derived from geocomputation and complexity with standard techniques that have been tried and tested in operational land-use and transport simulation. As a proof-of-concept exercise, a micro-model of residential location has been developed with a view to hybridization. The model mixes cellular automata and multi-agent approaches and is formulated so as to interface with meso-models at a higher scale.

Keywords: geocomputation, urban geography, urban simulation, multi-agent systems, residential location, housing choice.

1. Introduction

As the field of urban simulation moves into a state of maturity, it is noteworthy that the pace of change in model development appears relatively sluggish. Models in practical use today do not seem much changed from those in use ten or even twenty years ago (with the exception, of course, of models developed in academic circles; but, even there, there is much room for improvement). There are signs that the current generation of urban models has reached the limits of its usefulness under existing specifications. This proposition is unremarkable when we draw comparisons with other fields; new avenues of exploration dwindle, leaving little room for innovation. It is surprising, however, in the field of urban simulation, where cities are evolving and adapting at a pace that outstrips our capacity to study them in theoretical terms, let alone to model them. In short, the time is ripe in urban systems simulation for the infusion of new ideas.

The relative stasis in urban modeling stands in marked contrast to simulation efforts in other disciplines (ecology, environmental science, biology, physics, economics) where techniques, theories, and ideas drawn from computation and the burgeoning field of complexity studies are revitalizing the ways in which we conceptualize and model real-world (and hypothetical) phenomena. Many of these concepts and methodologies are appropriate for application to urban systems simulation, and particularly to the modeling of housing choice. Indeed, in many cases, ideas from computation and complexity studies—often clustered under the collective term of geocomputation within geography—are ideally suited to the simulation of urban processes and the patterns that those processes drive. The conditions to support the proliferation of geocomputation models in urban studies are, to a certain extent, already there. New generations of spatial data have been available for developing and validating urban simulation models at high resolutions. New data sources now exist, as do geographic information systems for managing and manipulating that data. There are new theoretical understandings of how dynamic adaptive urban systems function as complex adaptive and self-organizing systems. Computing power continues to grow in potency and fall in price. And, critically, new simulation techniques—particularly geocomputation—offer the potential for a 'revolution' in the way we model urban systems.

However, there exist several significant barriers to the successful use of these new tools in urban simulation. If ignored, these obstacles could doom these new ideas to a fate reminiscent of earlier waves of large-scale urban modeling (Torrens & O'Sullivan, 2001). And importantly, 'traditional' urban simulation models still have a great deal to offer.

This paper describes a relatively new approach to urban simulation; it describes a hybrid geocomputation model designed to support the exploration of 'what-if' scenarios for urban planning, urban management, and public policy formation, with particular emphasis on modeling residential location dynamics within the context of an interactive urban system. The hybrid approach fuses 'traditional' simulation methodologies that operate at macro- and meso-levels with a 'new wave' of geocomputation methodologies at a micro-scale. To demonstrate some of the practicalities of building hybrid models, a prototype residential location simulation is developed, fusing cellular automata and multi-agent systems at the micro-scale and designed to interface with meso-models at higher scales.

2. 'Traditional' urban models and 'traditional' housing choice models

'Traditional' urban models, developed in the style of the spatial interaction (and, to a lesser extent, the spatial choice) model, were pioneered in a time in which the field of urban simulation—and our ideas about how cities worked—was radically different from current manifestations. Computing power was also relatively less ubiquitous and sophisticated than it is today and detailed data sets were not widely available to 'feed' these models. The 'traditional' generation of urban simulation models has come under heavy criticism (Lee, 1973; Sayer, 1979; Lee, 1994). Many of these criticisms overlook some of the successes achieved by those models (Batty, 1979; Harris, 1994). However, we can identify several key weaknesses of 'traditional' models that still remain, particularly when contrasted with newer models currently being developed in academic contexts.

In most instances, the criticisms of urban models in general are equally applicable to models of housing choice, particularly geographic models. Traditionally, there are two main methodologies for simulating the geography of residential location: aggregate and what we shall term—for want of a better classifier—*choice* models. Aggregate geographic models of residential location were developed in the tradition of the Alonso model (Alonso, 1960) and

subsequent notions of bid-rent theory (DiPasquale & Wheaton, 1996). Other aggregate models of residential location were developed as spatial interaction models (Fotheringham & O'Kelly, 1989). What we have termed as choice models extended the work in economics on housing search, following in the style of disequilibrium approaches and stopping-rule models (Clark & Flowerdew, 1982), essentially adding location as a new variable to the search procedure. Additionally, some more spatially explicit models have been developed to explore the role of geography in housing search, including area-based search and anchor point search models (McCarthy, 1982).

Some important criticisms have been leveled against urban simulation in general (Lee, 1973; Sayer, 1979), and in many instances these complaints also ring true for models of housing choice. Aggregate models have been accused of an overly centralized approach, while the models in general have been censured for weak treatment of dynamics, poor attention to detail, lack of usability, flexibility, and realism. Importantly, many models of housing choice present work on individual behavior, without modeling relocating households as entity-level individuals!

3. A 'new wave' of urban models

In recent decades, geographers, economists, and sociologists have begun to work with a new class of simulation techniques that open up new and exciting possibilities for simulating systems of all descriptions, and in particular the simulation of behavioral processes and the structures that they generate. These models are in their relative infancy as applied to urban modeling and constitute a new class of simulation tools that borrow heavily from developments in geographic information science, artificial intelligence and artificial life, complexity studies, and simulation in natural sciences and social science outside of geography. While the use of computers and computation in urban simulation is by no means new, the geocomputation approach—modeling systems at the scale of individuals and entity level units of the built environment—is particularly innovative from an urban simulation standpoint and offers some significant advantages for the simulation of housing choice.

Advances in geographical information science

Within the geographical sciences, geocomputation models have been supported by a flood of detailed geographic information that has become easily attainable in recent years. This data has been made available in a variety of media and covering phenomena that would not have been possible a relatively short time ago, e.g., multi-spectral and fine-scale resolution remotely sensed data on land-use and land cover change in urban areas. The provision of these data has been directly responsible for addressing some of the weaknesses we have just mentioned: a lack of detail, for example. Also, it has had indirect impacts on urban simulation by supplying new insights into how urban systems operate, thereby allowing us to develop better-informed simulations. Furthermore, geographic information systems (GIS) have been developed to store, manipulate, and display spatial data. There is now a rich tradition of use of these systems in operational contexts.

Object-oriented programming

The treatment of discrete entities of urban systems; e.g., land parcels, buildings, administrative zones, households, and individuals; as objects has several advantages from a simulation standpoint. There are benefits associated with object-oriented programming (OOP) that remedy some of the deficiencies of 'traditional' models that we have already mentioned, particularly flexibility, usability, and realism. Object-oriented software has the advantage of being more realistic in terms of representing cities, housing, and households. The basic unit in OOP is the object (as opposed to the statement or the expression in procedural software). The conceptualization of pieces of inanimate code as objects mimics the way that we think of real world objects such as dwellings and the people that inhabit them: as discrete units with associated attributes and behaviors. Indeed, in OOP data and behavior are integrated (unlike the case in procedural software, where they are separate). This has the advantage of allowing model developers to focus on the program as a *simulation* rather than as a piece of software.

Complexity studies

Complexity studies are closely related to chaos theory (Gleick, 1987). The main idea in complexity is that of emergence. In emergent systems, a small number of rules or laws, applied at a local level and among many entities are capable of generating surprising complexity and often ordered patterns in aggregate form. Additionally, these systems are

dynamic and change over time without the direction of a centralized executive. Complex patterns manifest themselves in such a way that the actions of the parts do not simply sum to the activity of the whole (Holland, 1998). Essentially, this means that there is more going on in the dynamics of the system than simply aggregating little pieces into larger units.

Examples of emergent systems abound. For example, the liquidity of water is more than a simple extrapolation of characteristics that can be attributed to individual water molecules, which have no liquid quality in isolation (Krugman, 1996). Many urban systems are also complex in this sense. From the local-scale interactive behavior (commuting, moving) of many individual objects (vehicles, people), structured and ordered patterns emerge in the aggregate, such as peak-hour traffic congestion (Nagel, Rasmussen, Barrett, 1996) and the large-scale spatial clustering of socioeconomic groups by residence (Benenson, 1998). In urban economics, large-scale economies of agglomeration and disagglomeration have long been understood to operate from local-scale interactive dynamics (Krugman, 1996). Also, cities exhibit several of the signature characteristics of complexity, including fractal dimensionality and self-similarity across scales, self-organization, and emergence (Batty & Longley, 1994; Allen, 1997; Portugali, 2000).

Complexity studies offer the potential to shed new light on our thoughts regarding the inner workings of cities and subsystems such as the residential location market, and have already had profound impacts on our approach to urban simulation and modeling housing choice. Complexity studies point to a need for detailed, decentralized, and dynamic views of urban systems. The ideas also suggest that the answer to questions of the form, 'How do housing markets work?' might find new answers among the myriad and evolving interactions of individuals and the urban spaces that they inhabit. This is a much more *generative* approach than the reductionist view that is traditionally adopted in urban studies. Simply dissecting cities may not provide all the answers; on the contrary, there may be a need to build them up from the bottom and in doing so we may learn a lot about how they work. This may have some direct analogies in urban simulation also; indeed there are modeling techniques in geocomputation that work exactly on these principles, chiefly cellular automata (CA) and multi-agent systems (MAS).

Cellular automata and multi-agent systems

In terms of urban simulation, CA are perhaps best used to represent the dispersal of activity and characteristics between discrete spatial units of urban infrastructure, such as housing units or parcels of land. MAS may be more suited to simulating urban population as collectives of individuals or households with associated behaviors and traits and the capacity for spatial mobility and communication.

Cellular automata were originally pioneered in computing (Sipper, 1997) but have since seen uses in a wide variety of fields, including urban studies (Batty, Couclelis, Eichen, 1997; Torrens, 2000a). A cellular automaton is a finite state machine (an engine of sorts) that exists in some form of tessellated cell-space. The term automaton refers to a self-operating machine, but one of a very distinct nature: "An automaton is a machine that processes information, proceeding logically, inexorably performing its next action after applying data received from outside itself in light of instructions programmed within itself." (Levy, 1992, p.15) Additionally, CA are *parallel* automata: more than one automaton is active at any given instance. CA are comprised of five components. The *lattice* of CA is the space in which they exist. This might be considered equivalent in a residential urban context to an environment, a landscape, or a territory. The lattice can also be generalized to represent urban spatial structures, networks of accessibility, the physical structure of the city, etc. CA *cells* represent the discrete confines of individual automata. They are the elemental building blocks of a CA, just like individual land parcels or buildings in a city. CA cells are, at any time, in a particular *state*. The cell state offers a flexible framework for encoding attributes of a city into an urban simulation model, e.g., land-use, density, land cover, rental value, etc. *Neighborhoods* are the localized regions of a CA lattice (collections of cells), from which automata draw input. Neighborhoods in an urban CA might represent spheres of influence or activity, e.g., market catchment areas, commuting watersheds, etc. The real driving force behind CA are *transition rules*. These are simply a set of conditional statements that specify the behavior of cells as CA evolve over time. The future conditions of cells are decided based on a set of fixed rules that are evaluated on input from neighborhoods. CA rules can be devised to mirror how phenomena in real cities operate. Additionally, we might discern a sixth component to CA—*time*—that is generally discrete and proceeds in iterative steps.

CA offer a range of advantages for general urban simulation as well as for housing choice modeling, and in several ways they remedy particular deficiencies of 'traditional' models. CA can be designed with attention to detail. They are inherently spatial and decentralized. They are dynamic, as well as being intuitively useful and behaviorally realistic. Additionally, they have a "natural affinity" with raster data and GIS (Couclelis, 1997), as well as OOP. CA also provide a mechanism for linking micro- and macro-approaches and for connecting patterns with the processes that produce them.

While CA are most suitable, in urban simulation contexts, for representing infrastructure, MAS are better used to model population dynamics. MAS also have origins in computer science, although their development post-dates that of CA by some years. Most commonly, MAS are used in computing as artificial intelligence systems or artificial life forms (Kurzweil, 1999). Additionally, there are 'species' of agents that serve as network bots, webcrawlers, and spiders (Leonard, 1997). Network agents are used to navigate computer information networks, to 'mine' data, retrieve it, and return it to human users. There is also a tradition of using software agents to explore entomological behavior (Bonabeau, Dorigo, Theraulaz, 1999) and the actions of agents in economic systems and markets (Luna & Stefansson, 2000).

Agents are quite similar to automata in their formulation but have less well-defined characteristics. They constitute pieces of software code with certain attributes (states) and behaviors (rules) (see Ferber, 1999 for a general introduction to intelligent software agents). They differ from CA in their spatial mobility: agents can be designed to navigate (virtual) spaces with movement patterns that mimic those of humans, while CA are only capable of exchanging data spatially with their neighborhoods. Additionally, agents can be given functionality that allows them to evolve over time, altering their attributes and behavior with the help of genetic algorithms (Mitchell, 1998).

MAS are excellent tools for representing mobile entities in urban environments, e.g., people, households, vehicles, etc. They have been used in urban contexts to simulate pedestrian movement in dense urban environments (Schelhorn, O'Sullivan, Haklay et al., 1999; Dijkstra, Timmermans, Jessurun, 2000) and relocating householders (Benenson, 1998). However, their application to urban studies has not been as widespread as that of CA, despite offering

the advantages for urban simulation. Like CA, MAS are easily programmed in OOP environments, as well as offering advantages in terms of detail, flexibility, dynamics, usability, and behavioral realism.

4. The need for hybrid models

Even though CA and MAS are very suitable to the simulation of urban systems and despite the fact that they offer significant advantages over 'traditional' models, there are simply some things that they cannot represent well, most notably systems that operate from the top-down. In urban contexts—and housing markets in particular—there are several systems and mechanisms that operate in this manner, including constraints such as planning restrictions and global level phenomena such as socioeconomic shocks. In light of these and other considerations, there is a convincing argument for developing *hybrid* models for real-world urban planning and management and the formation of public policy, as well as for academic inquiry.

An approach that is based purely on CA or MAS is weaker than a more combined effort would be. Urban 'cells' do not simply mutate like bacteria in a lab experiment (O'Sullivan & Torrens, 2000); the characteristics of the urban infrastructure change over time because of human intervention within and around them. Similarly, cities are more than the people that inhabit them; there is a built environment that they influence and are, in turn, shaped by. Also, there are phenomena that operate above the scale of individuals and the urban fabric, such as regional economics, national geopolitical systems, weather, etc. CA and MAS are not well equipped to model these macro-level systems.

To focus purely on a 'new wave' of urban models would ignore a rich history and methodology of 'traditional' models that have been developed and applied to cities over many years. CA and MAS are new ideas and have not been fully tested in real-world contexts. Additionally, there is the problem of 'legacy' systems: very many planning agencies have elaborate and expensive systems in operational use already, formulated under the influence of 'traditional' methodologies. A 'new wave' of models could not hope to simply sweep the existing simulation infrastructure aside, nor would that be prudent. It would be

much better to work within existing simulation infrastructures, to interface with 'traditional' models and supplement them rather than supplanting them.

5. A conceptual design for a hybrid geocomputation model

With the foregoing considerations in mind, we now present a conceptual framework for a model designed as a hybrid geocomputation environment for generating what-if scenarios to help in public policy formation—you might consider them as tools to think with. The framework merges approaches from geocomputation (CA and MAS) with 'traditional' simulation techniques, offering a suite of tools for modeling urban systems. Macro-scale dynamics that operate from the top-down are handled by 'traditional' land-use and transport models, while micro-scale dynamics that work from the bottom-up are delegated to geocomputation models. The two methodologies are fused in a modular fashion using a system of constraining feedback mechanisms. In section 6 a prototype model for simulating residential location dynamics is presented within the framework of a fully interactive urban model, demonstrating how geocomputation models can be designed with this sort of framework as a consideration.

Hybrid models are not new to urban simulation. Most operational urban models are hybrids consisting of separate modules for handling land-use (location decisions, development and redevelopment, market-clearing) and transport (potential demand and trip generation, trip distribution, modal split, trip assignment; figure 1). Moreover, hybrid geocomputation models are not new! White, Engelen, and colleagues have developed a comprehensive hybrid simulation environment using CA and more 'traditional' simulation techniques for operational uses in the Netherlands and elsewhere (White & Engelen, 1997; White & Engelen, 2000).

Figure 1. Specification of a 'traditional' land-use and transport model.
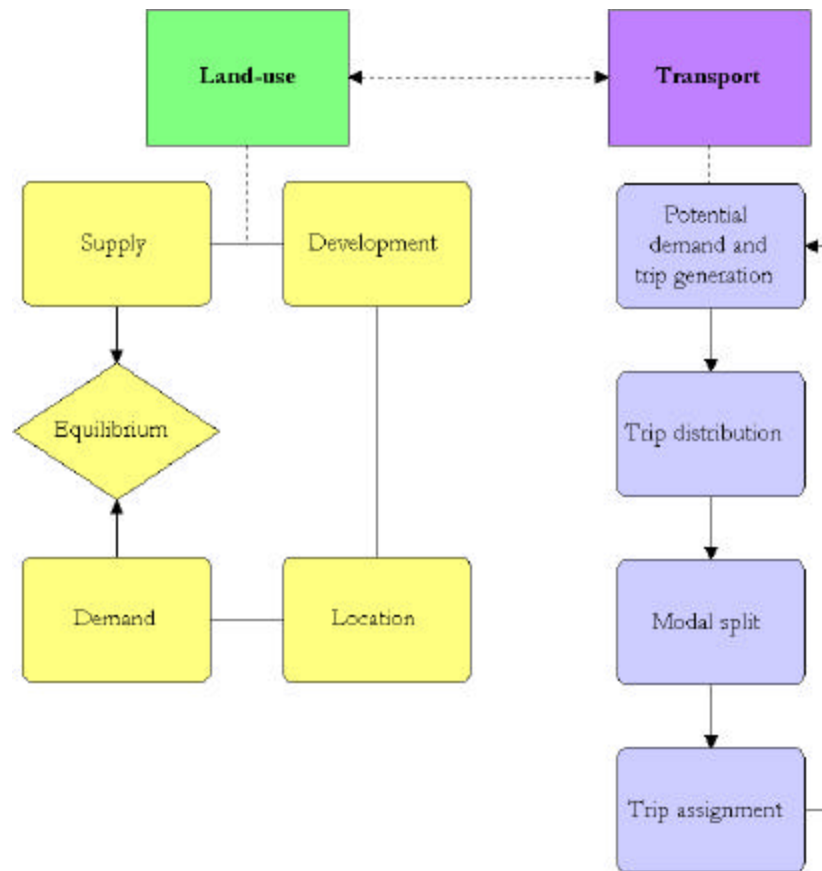
So, how does our conceptual design differ from that of related work? Essentially, our model is designed to do mostly the same things, and goes about it in a roughly similar fashion (figure 2). There are some important differences however. Our model is formulated so as to interface with systems that are already familiar in planning agencies. The micro-scale models we are developing can be viewed as a logical extension of the 'traditional' model design. This interface could, conceptually, constitute a simple exchange of data between models, a set of constraints operating from the top-down or from the bottom-up, or the connection could be more tightly coupled through integrated modeling or feedback mechanisms. Our design uses MAS at the micro-scale, closely merged with a CA environment. Individuals in this design are represented explicitly as agents, while sites are modeled as CA. The algorithms that drive dynamics at the micro-scale are also designed so as to be as compatible as possible with existing systems commonly in real-world use in many planning agencies. Wherever feasible we use methodologies already tried and tested in operational simulation, particularly ideas

from urban economics and decision theory while adding unique behaviors to the model. The goal is to make the connection with 'traditional' models as seamless as possible, while retaining innovation.

The model is designed in a highly modular fashion and as such has the potential to be highly flexible. Modeling of land-use and transport is separated (although the two approaches are linked via feedback mechanisms) because the two systems require quite different treatment, both in a theoretical sense and in terms of designing simulations. For the purposes of this discussion, we will focus on the land-use component of the model. The 'traditional' tool for transport modeling is the four-stage model (figure 1), but there are quite a rich range of methodologies for microsimulation of transport (Ben-Akiva & Bowman, 1998) and there are several innovative geocomputation approaches to traffic simulation (Nagel, Beckman, Barrett, 1999).

The land-use component of the simulation environment is divided into three sets of models: those dealing with macro-level, meso-level, and micro-level subsystems (figure 2). We are not necessarily concerned with building models at the macro- and meso-scales as there are several such models currently in existence and in operational uses in urban planning and management (Torrens, 2000b). However, it is important to consider such systems when developing interface tools that operate at the micro-scale. Standard regional science models are used to establish 'seed' conditions for the model at the macro-scale. Generally, such models are split between simulating economic and demographic transition (Isard, 1975). This section of the model operates at very coarse levels of spatial and socioeconomic resolution. Geographically, it deals with large metropolitan regions, or perhaps with collections of such regions. On a socioeconomic level, employment and economic activity is divided into only a few key sectors, while demographics are handled at the level of a few household types. At the meso-level, the simulation is divided by activity. Land and real estate development is modeled on the demand and supply sides, with market-clearing mechanisms to reconcile the two. A land-use transition model simulates the dispersal of activity in the urban infrastructure. The location decisions of households, office employment activities (finance, real estate, and insurance), and (non-service) industry are handled by meso-scale location models. The meso-scale models simulate at an intermediate level of spatial and

socioeconomic resolution. Geographically, the lowest level of detail is that of the TAZ or local economic submarket (a neighborhood or district within a city, for example).

The micro-level models pick up where the meso-scale models have left off (figure 3). Conceptually speaking, they take constraint values from higher-level models and 'distribute' them to entity level units of the built infrastructure or individuals. Equally, they could be formulated to operate in the opposite direction, supplying constraints for higher-level models, or perhaps work in a bi-directional fashion. The micro-scale infrastructure is represented as a CA 'landscape', which we populate with life-like agents. Various components affecting land-use dynamics are modeled in the conceptual design: the supply of and demand for real estate (mediated by development agents); land-use transition; and relocating households, offices, and industries. We have developed one of the micro-scale components: a model for residential location. In section 6 we will report a prototype model that demonstrates how the micro-level modules are constructed and how they work.

6. A prototype residential location model

As a proof-of-concept exercise, we have built one component of the micro-scale simulation environment: a residential location module. The model is designed to simulate the residential location process from the standpoint of individual homebuyers and sellers, as well as the sites that they are exchanging. The model is formulated as a MAS-CA hybrid. The micro-scale model interfaces with its 'big brother'—a meso-scale residential location model (figure 2). The meso-scale model provides a set of 'seed' conditions for the micro-model. Total attribute values for a single neighborhood (which you might think of as a local residential submarket) are thus 'known' at the start of the model. At various stages in the evolution of the micro-model, we can 'feed' it more of this data, which in turn may be used to constrain the behavior of the micro-model (somewhat like checking its progress over time). (The process could potentially operate the other way around, with the micro-model serving as a constraint on the higher-level meso-model.) Essentially then, the micro-model takes output from the meso-model and assigns it to individuals and individual residences within a given local submarket. For the purposes of this discussion, we have developed a working prototype, without a meso-level interface. We have also built the model with abstract data

for a single and hypothetical submarket, although we hope to test the simulation with real data.

There are three main components to the micro-model: sites (the urban infrastructure), agents (the population inhabiting or visiting those sites), and globals (various storage bins for capturing conditions in the inner workings of the model).

Sites are formulated as a cellular automata 'landscape', however there are only a few transition rules applied to the sites and this is done simply to manipulate their state variables over time; there are no dispersal mechanisms in the model (although this may be added at a later stage, allowing the infrastructure to evolve over time, e.g., to gentrify). Each site represents a particular piece of real estate with attributes as listed in Table 1. Currently a value is assigned to a property in an abstract manner, although this could be reformulated in such a way that the price of a given piece of real estate is formulated as a bundle of attributes (bathrooms, bedrooms, aspect, etc.) associated with the property: a so-called hedonic price. Additionally, for the purposes of interfacing with meso-level models, sites could have neighborhood characteristics added to their list of attributes, e.g., distance from a nearby center, accessibility to highway networks, etc.

Two types of agents are represented in the model: homebuyers ('mobile' agents) and home sellers ('residential' agents). (There is also a third, 'god' agent that is used to automate tasks within the model.) The agents are designed with various attributes as listed in Table 1. (For the sake of parsimoniousness, residential and mobile agents are designed with the same attributes, although certain values may be set to null.) Additionally, agents are entrusted with various behaviors: a set of preferences for housing as well as the capacity to move over the real estate landscape and sense their surroundings.

Figure 2. Conceptual design of a hybrid model.
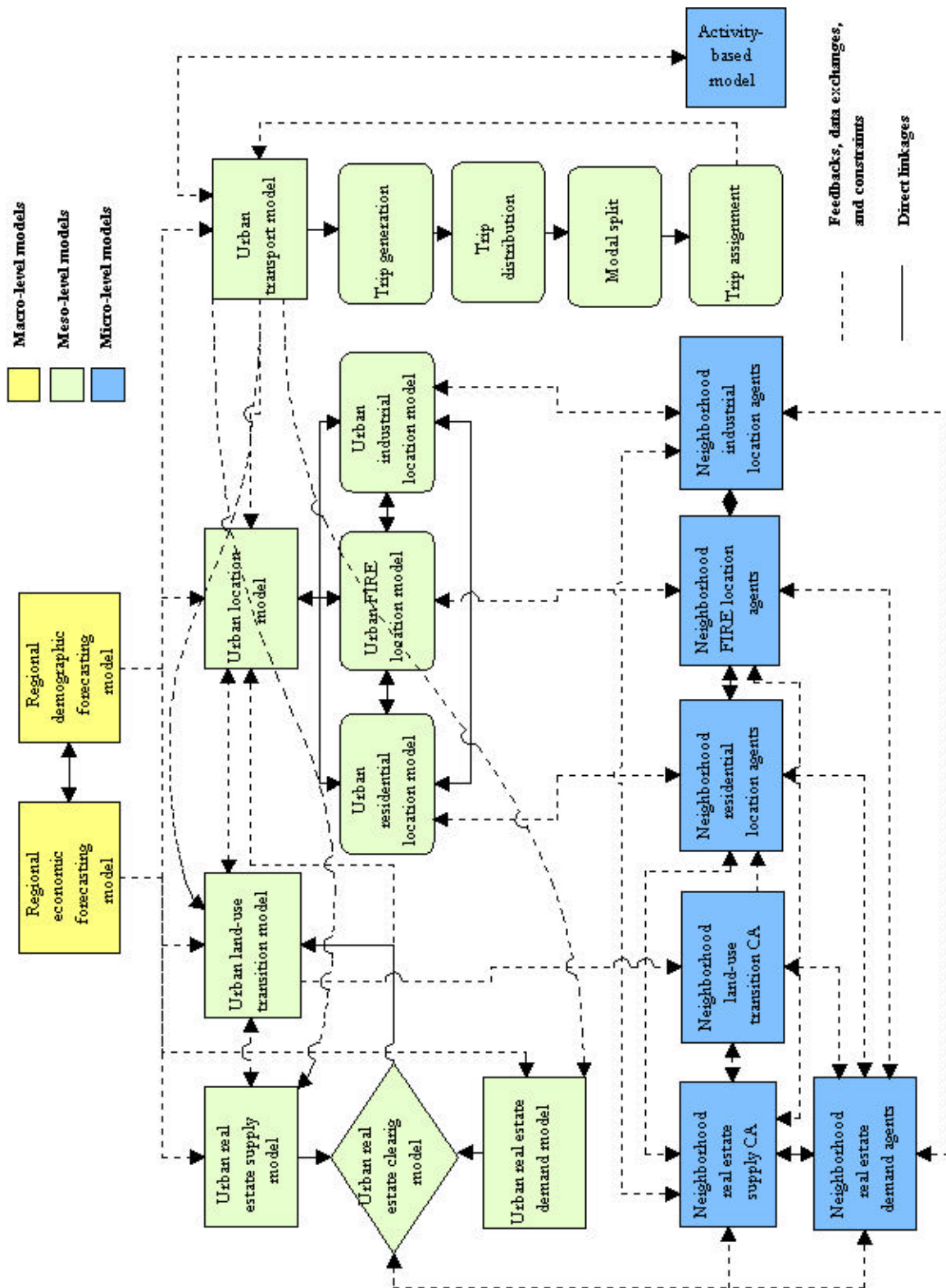
Figure 3. From macro- to micro-scales: Megalopolis to New York to Lower Manhattan to Times Square.
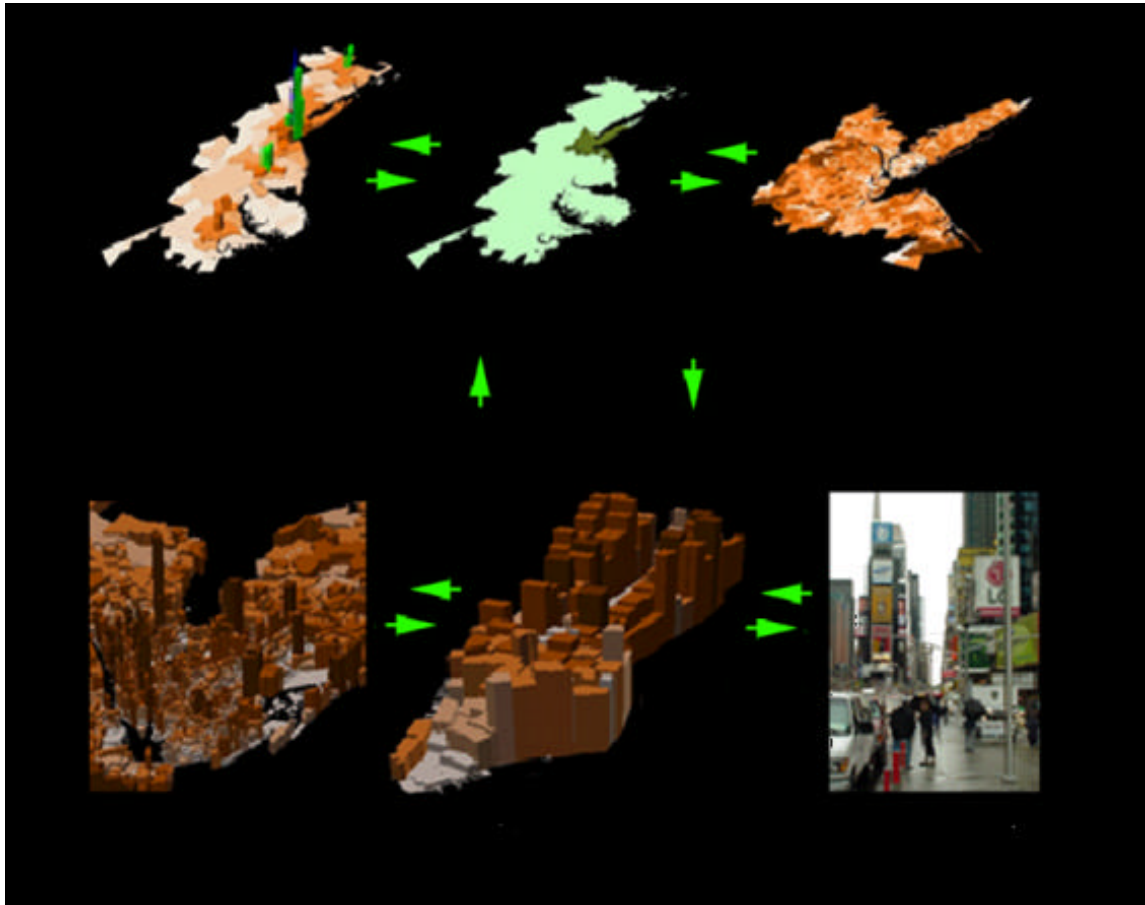
Calculating lifecycle stage and value platforms

The matching of mobile agents with sites and the decisions by residential agents regarding when to sell their properties are driven by a set of preference functions that are calculated within the model. This lends agents a set of 'likes' and 'dislikes', both for particular types of neighborhoods, other agents, and certain types of housing. Based on their preference functions, mobile agents are matched with suitable homes.

Table 1. The attributes of objects encoded within the model.

Sites	Agents
Value	Income
Housing type	Age
Lot size	Children
Housing tenure	Household size
Density	Ethnicity
Land-use	Inertia
Number of bedrooms	Residency
Rental value	Segregation preference
Discounting function	Lifecycle stage
	Tenure preference
	Housing preference
	Housing budget
	Willingness to leave submarket
	Socioeconomic preference
	Agent type

One of the key variables that determine agents' preferences for housing is their stage in the lifecycle (or family cycle). A rich literature exists for determining the role that lifecycle characteristics play in the residential location process (see Waddell, 2000 for a review), as well as a burgeoning science of geodemographics (Longley & Harris, 1999). Depending on whether individuals or households are young and/or without families or in retirement, their preferences for various types of housing or characteristics of individual properties—number of bedrooms, tenure, housing type—will change.

Currently, our model discerns three lifecycle stages: 'young', 'middle', and 'senior'. An attribute that denotes the presence of an agent in one of these lifecycle stages is added to their attribute profile. 'Young' agents are designed to represent individuals that have recently left the family home and are striking out on their own for the first time. They may be studying or working in their first full-time jobs. In the context of the geography of urban

location, individuals at this stage in their lifecycle may well demonstrate a preference for central locations close to entertainment facilities. Also, we can identify certain housing-specific preferences; individuals at this stage in the lifecycle are more likely to favor apartment living than a house.

'Middle' agents represent individuals that are at a stage in their lives where they may be beginning to start a family, or may already have started a family. Such individuals are bound to have different residential location geographies when compared to other lifecycle groups. One factor that they may find desirable, but which would be unlikely to feature highly in the preferences of other groups, is the presence of good schools in a suburban location, for example.

'Senior' agents correspond to those individuals entering into retirement age, either without children or with children that have left home. We might consider these agents as representing 'empty-nesters'. This is a tricky demographic group to model. Income variations may well influence the residential location behavior of 'senior' groups more than in other groups. Some may own multiple homes with quite different characteristics, e.g., a house in one location and a condominium in another.

Currently, the calculation of lifecycle stage is performed quite simply in the model as a set of conditional statements based on age (although the potential to expand that calculation to incorporate other factors and perhaps to link it with family cycles, along with the potential of diasaggregating the groupings further, is there). If agents are between the ages of 22 and 35 they are assigned a 'young' tag; between 35 and 65 they are assigned a 'middle' tag; and over 65 they are regarded as 'senior' (agents under the age of 22 are not represented in the model).

Another important variable that needs to be calculated and assigned as an agent attribute is a 'value platform': the amount of money that an agent can spend per month on rent or mortgage payments. Currently, value platforms are calculated by simply dividing an agent's income by 12. However, this could potentially be reworked as a more complicated calculation relying on other agent attributes such as number of children, employment, and age.

Variables for lifecycle stage and value platform are used in conjunction with other agent attributes (income, age, presence of children, size of household to which the agent belongs, ethnicity, inertia, and period of residency) as ingredients for the derivation of a set of preference functions. These preference functions—coupled with a set of transition rules, the capacity for spatial mobility, and the ability to 'sense' their surroundings—govern the behavior of agents as the model evolves.

Establishing preference functions

6.1.1.   *Site specific preferences*

Agents are assigned a set of preferences in the model, both for specific attributes of sites and for the neighborhoods in which individual properties are situated. A preference for housing types (apartments or houses) is assigned to each agent. Housing preference is one of the methods that rely heavily on an agent's lifecycle attribute. Depending on an individual's stage in the lifecycle, she is likely to have a strong preference for a house or an apartment (regardless of whether she can afford it). Preference for housing is assigned to agents in the model, principally based on lifecycle stage. If an agent is 'young' its preference is for apartments. Individuals with families are likely to prefer houses, all other things being equal. 'Middle' agents with children are given a preference for houses, while those without are assigned preferences for apartments. 'Senior' agents are also assigned a preference for houses.

Preferences for housing tenure (rent or own) are also assigned to agents in the model. 'Young' agents are assumed to give preference to rental accommodation, while 'middle' and 'senior' agents have a preference for owner-occupation.

6.1.2.   *Neighborhood preferences*

In addition to preferences for site-specific attributes of housing, agents are also assigned neighborhood-level preference functions. The implication here is that homebuyers and home sellers factor certain conditions of the local residential submarket into their location decisions, principally ethnicity and socioeconomic factors. (We could also add some other indicators representing the quality of the built environment or the availability of neighborhood-scale amenities such as recreation, retail, and entertainment.)

Socioeconomic preferences are currently calculated only for mobile agents (although this could be extended to embrace residential agents also, factoring into their decisions regarding whether to begin a housing search). Upon entering the local submarket, an agent assesses whether the neighborhood is too cheap or too expensive for its budget. If so, the agent moves on to another submarket; if not, the agent begins to evaluate individual properties in the submarket. This preference is calculated as follows:

$$S_n = f(c,e) \text{; where } c \in \{0,1\} \text{ and } e \in \{0,1\} \tag{i}$$

$$c = 1 \text{ if } \left(V_{\min} \geq \left(\frac{I_m}{2} \right) \right), \text{ otherwise } c = 0 \text{;} \tag{ii}$$

$$e = 1 \text{ if } \left(I_m < V_{\min} \right), \text{ otherwise } e = 0 \tag{iii}$$

Where S_n is the socioeconomic preference for neighborhood n; c is an evaluation of whether a submarket is too cheap and e is an evaluation of whether a submarket is too expensive. $V_{\min}$ is the minimum value of housing in the neighborhood and I_m is the income of mobile agent m.

Socioeconomic preferences are also calculated for residential agents, although they are not used as part of their decision to stay in the submarket, nor are they factored into the sale price of an agent's property. This functionality could be added, however, allowing agents to 'sense' the socioeconomic decline or gentrification of their neighborhood. Additionally, residential agents could 'sense' the socioeconomic profile of other households in the neighborhood by examining changes in the income of their neighbors.

In addition to a set of neighborhood-level socioeconomic preferences, agents are also designed with a level of bias towards the ethnic make-up of the neighborhoods that they inhabit or evaluate as a potential home. A lot of work has been done looking at the geography of segregation in the housing market. Perhaps the most famous is that of Thomas Schelling (Schelling, 1969, 1978), which looked at how large-scale residential segregation could emerge from individual biases. Also, Benenson and colleagues have developed several influential MAS for exploring the spatial dynamics of residential segregation in Israel (Portugali, Benenson, Omer, 1997; Benenson, 1998, 1999). In our model, agents are
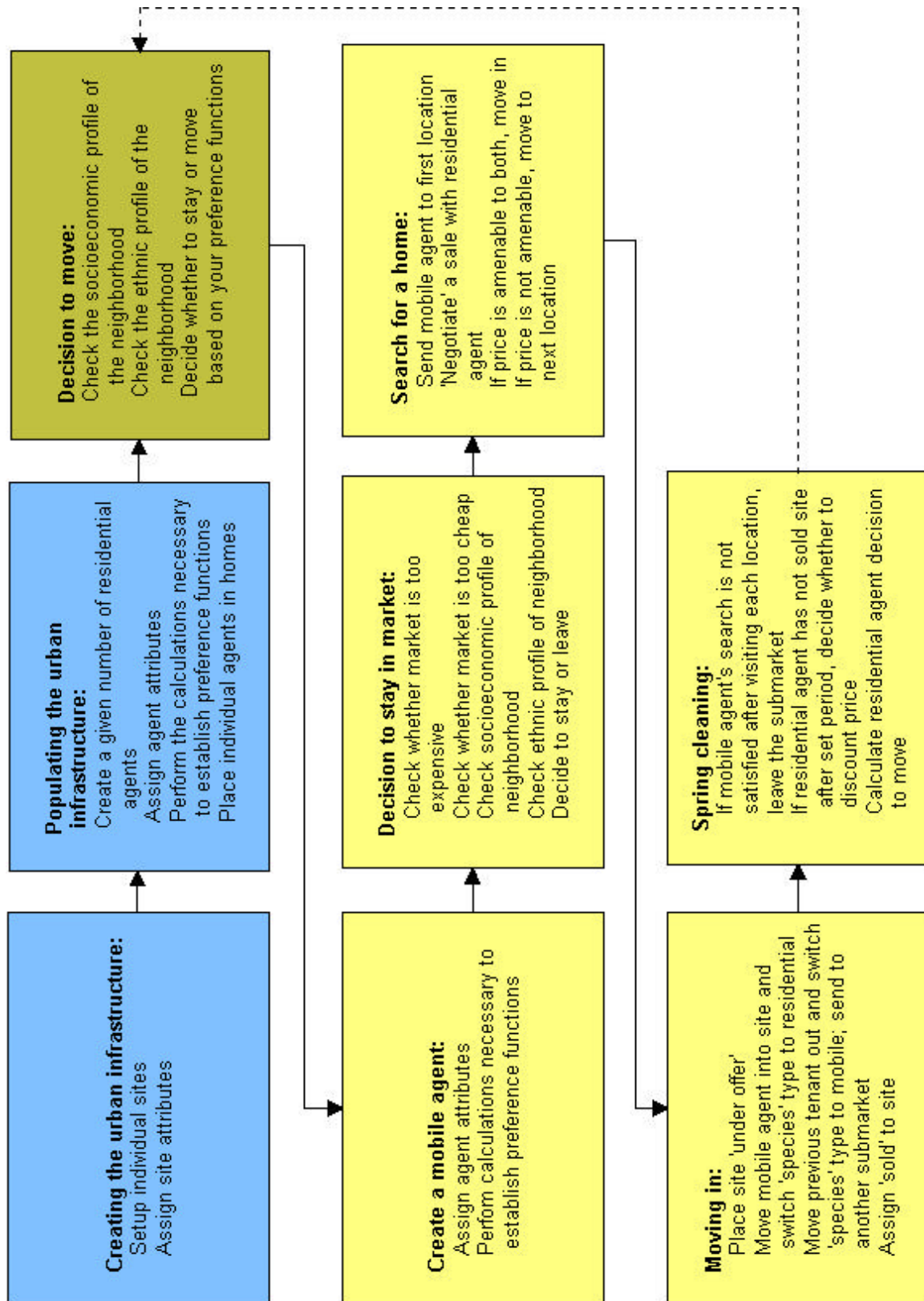
arbitrarily assigned colors (blue, red, and yellow) that we use to denote ethnicity. Agents of any given color have a certain preference for living with agents of the same or different colors. Specifically, agents are designed with a tolerance for living in neighborhoods with certain ethnic profiles. Red agents do not like to live in a neighborhood where blue agents form a majority, but are reasonably tolerant of living with yellow agents. Similarly, blue agents have a preference for living in neighborhoods where blue agents form the majority of householders. They do not like to be outnumbered by red agents and are ambivalent about the numbers of yellow agents in the submarket. Yellow agents have no bias for color. Cut-off values ('tipping balances') for these preferences are assigned as follows. Red agents do not like to live in neighborhoods where the proportion of the population that is blue exceeds 50%. Blue agents, on the other hand, will only tolerate living in neighborhoods up until the point where red agents constitute 33% of the population.

Operationalizing the model

An actual run of the model is organized as a series of events. Many sub-events within the model (such as calculations and the derivation of preference functions) occur on a parallel basis, but the main events in the model—setup and the initiation of model parameters, simulating the location process, and the updating of model parameters—occur iteratively (figure 4).

Before we introduce mobile agents into the simulation environment we must determine whether any of the residential agents would like to put their properties on the market. Some computations are performed and residential agents make a decision whether to move, based on their own conditions and their knowledge of the neighborhood in which they reside. If an agent decides to put its home on the market, the characteristics of the site variable for that particular location are updated to reflect that.

Figure 4. Schematic diagram illustrating the key events in the model.

Creating the urban infrastructure:
Setup individual sites
Assign site attributes

Populating the urban infrastructure:
Create a given number of residential agents
Assign agent attributes
Perform the calculations necessary to establish preference functions
Place individual agents in homes

Decision to move:
Check the socioeconomic profile of the neighborhood
Check the ethnic profile of the neighborhood
Decide whether to stay or move based on your preference functions

Create a mobile agent:
Assign agent attributes
Perform calculations necessary to establish preference functions

Decision to stay in market:
Check whether market is too expensive
Check whether market is too cheap
Check socioeconomic profile of neighborhood
Check ethnic profile of neighborhood
Decide to stay or leave

Search for a home:
Send mobile agent to first location
'Negotiate' a sale with residential agent
If price is amenable to both, move in
If price is not amenable, move to next location

Moving in:
Place site 'under offer'
Move mobile agent into site and switch 'species' type to residential
Move previous tenant out and switch 'species' type to mobile; send to another submarket
Assign 'sold' to site

Spring cleaning:
If mobile agent's search is not satisfied after visiting each location, leave the submarket
If residential agent has not sold site after set period, decide whether to discount price
Calculate residential agent decision to move

Now we introduce mobile agents into the simulation. Currently only a single agent visits a given residential submarket at any stage in the model, but that could be reformulated to create an environment of competitive buying, or perhaps some more complicated bidding games. A mobile agent is created, assigned attribute data, and the calculations necessary to establish its preference functions are performed. The mobile agent then goes through the process of deciding whether or not the neighborhood that it has entered is suitable, before evaluating individual sites. The mobile agent checks whether the market is too expensive for its budget, or alternatively whether it is too far below (50% of) its value platform. Then the agent scans the socioeconomic and ethnic profiles of the residential agents already residing in the submarket, and based on its biases will decide whether to stay in the submarket and evaluate sites, or move on to another submarket elsewhere.

If the agent decides to stay, it begins to search for a home. The agent moves within the model space and visits the first location for sale. Once there, it 'negotiates' a sale with the residential agent. If the price of the property is amenable to both agents (and the characteristics of the property match the preferences of the mobile agent), the mobile agent will 'move-in', otherwise it will visit the next available property. If after visiting all available properties in the model, the agent has not found a home, it leaves the particular submarket and begins its search elsewhere. However, if the agent decides to buy or rent a particular property, the property is put 'under offer'. The mobile agent and the residential agent trade 'species' tags (the mobile agent becomes residential and vice-versa); the residential agent is moved out of the submarket and the mobile agent moves into the property; and a 'sold' tag is assigned to that particular site.

The final stage in an iteration of the model is a round of 'spring-cleaning'. Dissatisfied mobile agents are sent to alternative submarkets and if a residential agent has not managed to sell its property it decides whether to discount the price of the real estate in subsequent iterations of the model. Currently, prices are discounted by 5% after four iterations of the model. The model then returns recursively to decide whether residential agents are going to move.
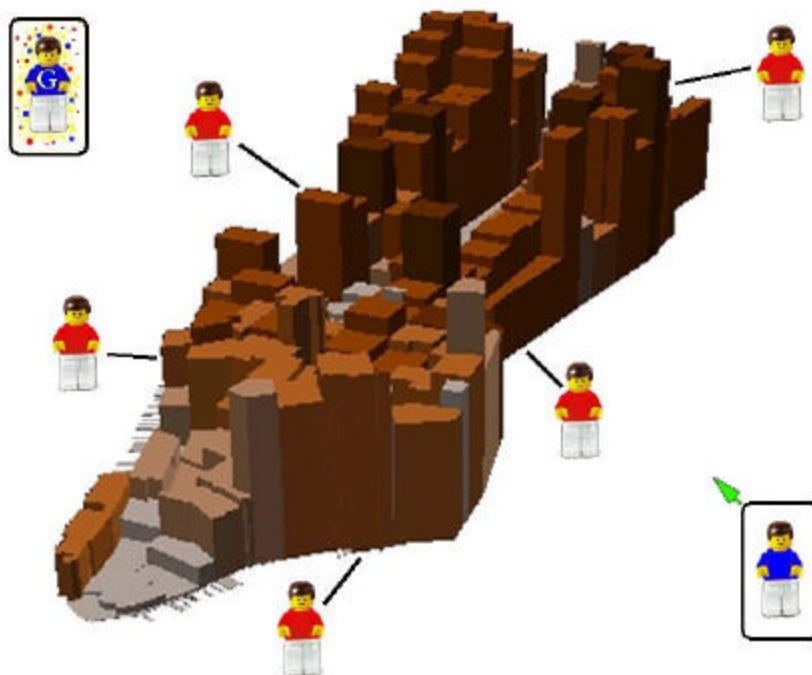
Graphic user interface

The model can be manipulated in an interactive fashion by the user through the use of a graphic user interface (GUI). Figure 5 shows the GUI for one particular stage in the run of a model. Windows for particular agents or particular sites can be called up to display the attributes of those objects at any given moment in the model. In figure 5 we have displayed windows for mobile and residential agents as well as the 'god' agent. Additionally, a window for a particular site is displayed. Also, a series of buttons and sliders are available to run particular events in the model and to vary the value of parameters that are used in model calculations, 'on-the-fly'. A main graphics window is also shown, providing information on the position of sites and agents within the model space at any given moment. Additionally, symbols in the graphics window can be programmed to alter shape and color depending on the conditions of the attributes that they represent (in figure 5 they are colored to represent the 'ethnicity' of the agents residing in those sites). The graphics window, and the artificial submarket that it represents, are designed to mimic how a residential submarket would appear in the real world (figure 6). Residential agents are situated within particular sites. Upon visiting the submarket, a mobile agent will travel to these sites and evaluate their suitability for its purposes. Additionally, we have a 'god' agent (denoted in the diagram with the letter 'G') that is active in automating tasks within the simulation, but does not partake in the residential location process.

7. Future developments

The model presented in this paper is a prototype, designed to function as a proof-of-concept tool. Several developments and additions to the model are planned. Specifically, we hope to add more attributes and behaviors into the model to make the simulation more realistic. Some of these plans call on tried and tested methodologies from 'traditional' models, such as the reformulation of preference functions as logit and spatial choice models (De la Barra, 1989). The model is currently setup in a nested fashion with the processing of events at specific cycles in the model, but the specification of functions in a probabilistic fashion at each stage in the nesting would lend the model an added degree of realism. Also, the model is quite 'old fashioned' in its characterization of residential location behavior and we would

like to explore other methodologies (marketing, spatial cognition, psychology, sociology, microeconomics, etc.) to find more suitable premises upon which we can design more life-like algorithms.

The current application of the model to a handful of sites in one particular submarket is, of course, quite simplistic. Linking several independent submarkets and facilitating the exchange of agents between them will take some further work. However, this should allow certain hypotheses about the residential location process (the dynamics of gentrification and neighborhood decline, the factors driving residential segregation, etc.) to be explored in an abstract fashion. Connecting the residential micro-model with other related micro-components such as industrial location and development modules is another task that we need to accomplish. Additionally, there is much work to be done in designing interfaces (data exchanges, constraints) with meso- and macro-scale models, as well as the design of feedback mechanisms between independent model components.

Figure 6. The graphic user interface to the residential location model.

Figure 7. Organization of the residential location model.

8. Conclusions

The discussion thus far has been quite optimistic about the potential of geocomputation techniques to revitalize operational simulation. The techniques themselves do certainly represent the possibility for a 'revolution' in the way we simulate urban systems. However, there are some imposing barriers to putting those techniques into practical use in the real world (Torrens & O'Sullivan, 2001). Ironically, computing power poses one of the most pressing limitations. The prototype that we have developed here works quite well and is efficient computationally. However, scaling that model up to represent an entire metropolitan area would require daunting levels of computing power. The only operational equivalent is the TRANSIMS model at Los Alamos National Laboratories, which relies on distributed computing clusters (Nagel, Beckman, Barrett, 1999).

Also, there are data limitations on the development of these models for practical uses. Conceptually, the idea of simulating individuals and the buildings that they inhabit is quite appealing. However, data is not widely available at the scale of the individual householder or

building. Also, there are several moral issues that arise from the use of individual-level—and often private—data in operational simulations.

Working at the micro-scale, in some cases, reveals inadequacies in the theory of how cities work. The micro-approach betrays some theoretical gaps in our understanding of the dynamics interactions that shape our urban systems. Indeed, there is some justification for a 'new urban geography' of the micro-scale.

Furthermore, micro-scale models, particularly dynamic and process-driven simulations, are quite difficult to calibrate, even if data are available. In CA research, there are some techniques for validating the *patterns* that those models generate and or matching them with real world conditions. However, *process*-based calibration techniques are not widely available (Torrens & O'Sullivan, 2001). Organizing the model as a hybrid allows the possibility of scaling up the simulation to meso-scales for validation purposes. This is a reasonable solution, but ideally micro-models would be calibrated at the scale of the entity or the individual. The likely effort required to do this is, however, a daunting prospect.

The point that we would like to convey in this paper, however, is that—at least methodologically—the techniques discussed here represent a move towards more theoretically sound, behaviorally realistic, and ultimately more *useful* simulation environments. As computer hardware develops and becomes cheaper and as detailed data become more widely available, the possibilities for applying geocomputation simulations in real world contexts grow. Certainly, these simulations can be developed as proof-of-concept tools and the methodologies can be refined in academic contexts in preparation for a day in which these tools can be used to plan and manage better cities. In the meantime, even as abstract tools, these simulations can do a lot for our understanding of how cities and housing markets work and perhaps provide new insights into how we might construct a more sustainable urban future.

References

Allen, P M, 1997, *Cities and Regions as Self-Organizing Systems: Models of Complexity*, (Gordon and Breach Science Publishers, Amsterdam).

Alonso, W, 1960, "A theory of the urban land market", *Papers and Proceedings of the Regional Science Association* **6** 149-158.

Batty, M, 1979, "Progress, success, and failure in urban modeling", *Environment and Planning A* **11** 863-878.

Batty, M, Couclelis, H, Eichen, M, 1997, "Special issue: urban systems as cellular automata", *Environment and Planning B* **24**.

Batty, M & Longley, P, 1994, *Fractal Cities*, (Academic Press, London).

Ben-Akiva, M & Bowman, J L, 1998, "Integration of an activity-based model system and a residential location model", *Urban Studies* **35** 1131.

Benenson, I, 1998, "Multi-agent simulations of residential dynamics in the city", *Computers, Environment and Urban Systems* **22** 25-42.

Benenson, I, 1999, "Modelling population dynamics in the city: from a regional to a multi-agent approach", *Discrete Dynamics in Nature and Society* **3** 149-170.

Bonabeau, E, Dorigo, M, Theraulaz, G, 1999, *Swarm Intelligence: From Natural to Artificial Systems*, (Oxford University Press, New York).

Clark, W A V & Flowerdew, R, 1982, "A review of search models and their application to search in the housing market", in *Modelling Housing Market Search*, Ed. W A V Clark (Croom Helm, London), 4-29.

Couclelis, H, 1997, "From cellular automata to urban models: New principles for model development and implementation", *Environment and Planning B* **24** 165-174.

De la Barra, T, 1989, *Integrated Land Use and Transport Modelling: Decision Chains and Hierarchies*, (Cambridge University Press, Cambridge).

Dijkstra, J, Timmermans, H J P, Jessurun, A J, 2000, "A multi-agent cellular automata system for visualising simulated pedestrian activity", in *Theoretical and Practical Issues on Cellular Automata*, Ed. S Bandini, T Worsch (Springer-Verlag, London), 29-36.

DiPasquale, D & Wheaton, W, 1996, *Urban Economics and Real Estate Markets*, (Prentice Hall, Englewood Cliffs, N.J.).

Ferber, J, 1999, *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*, (Addison-Wesley, Harlow (UK)).

Fotheringham, A S & O'Kelly, M E, 1989, *Spatial Interaction Models: Formulations and Applications*, (Kluwer Academic Publishers, Dordrecht).

Gleick, J, 1987, *Chaos: Making a New Science*, (Penguin Books, New York).

Harris, B, 1994, "The real issues concerning Lee's "Requiem"", *Journal of the American Planning Association* **60** 31-34.

Holland, J H, 1998, *Emergence: From Chaos to Order*, (Perseus Books, Reading, MA).

Isard, W, 1975, *Introduction to Regional Science*, (Prentice-Hall, Englewood Cliffs, New Jersey).

Krugman, P, 1996, *The Self-Organizing Economy*, (Blackwell, Malden, MA).

Kurzweil, R, 1999, *The Age of Spiritual Machines: How We Will Live, Work and Think in the New Age of Intelligent Machines*, (Phoenix, London).

Lee, D B, 1973, "Requiem for large-scale models", *Journal of the American Institute of Planners* **39** 163-178.

Lee, D B, 1994, "Retrospective on large-scale urban models", *Journal of the American Planning Association* **60** 35-40.

Leonard, A, 1997, *Bots: The Origin of a New Species*, (Hardwired, San Francisco).

Levy, S, 1992, *Artificial Life: The Quest for a New Creation*, (Penguin Books, London).

Longley, P A & Harris, R J, 1999, "Towards a new digital data infrastructure for urban analysis and modelling", *Environment and Planning B* **26** 855-878.

Luna, F & Stefansson, B, Eds, 2000, *Economic Simulation in Swarm: Agent-based Modelling and Object Oriented Programming* (Kluwer, Dordrecht).

McCarthy, K, 1982, "An analytical model of housing search and mobility", in *Modelling Housing Market Search*, Ed. W A V Clark (Croom Helm, London), 30-53.

Mitchell, M, 1998, *An Introduction to Genetic Algorithms*, (MIT Press, Cambridge, MA).

Nagel, K, Beckman, R J, Barrett, C L, 1999, "TRANSIMS for urban planning", LA-UR 98-4389, Los Alamos National Laboratory, Los Alamos, NM. `http://transims.tsasa.lanl.gov/`.

Nagel, K, Rasmussen, S, Barrett, C L, 1996, "Network traffic as self-organized critical phenomena", TSA-DO/SA MS-M997 and CNLS MS-B258, Los Alamos National Laboratory, Los Alamos, NM. `http://transims.tsasa.lanl.gov/`.

O'Sullivan, D & Torrens, P M, 2000, "Cellular models of urban systems", in *Theoretical and Practical Issues on Cellular Automata*, Ed. S Bandini, T Worsch (Springer-Verlag, London). `http://www.casa.ucl.ac.uk/working_papers.htm`.

Portugali, J, 2000, *Self-Organization and the City*, (Springer-Verlag, Berlin).

Portugali, J, Benenson, I, Omer, I, 1997, "Spatial cognitive dissonance and sociospatial emergence in a self-organizing city", *Environment and Planning B* **24** 263-285.

Sayer, R A, 1979, "Understanding urban models versus understanding cities", *Environment and Planning A* **11** 853-862.

Schelhorn, T, O'Sullivan, D, Haklay, M, et al., 1999, "STREETS: an agent-based pedestrian model", CASA Working Paper 9, University College London, Centre for Advanced Spatial Analysis, London.

Schelling, T C, 1969, "Models of segregation", *American Economic Review* **59** 488-493.

Schelling, T C, 1978, *Micromotives and Macrobehavior*, (WW Norton and Company, New York).

Sipper, M, 1997, *Evolution of Parallel Cellular Machines: The Cellular Programming Approach*, (Springer, Berlin).

Torrens, P M, 2000a, "How cellular models of urban systems work", CASA Working Paper 28, University College London, Centre for Advanced Spatial Analysis. `http://www.casa.ucl.ac.uk/working_papers.htm`.

Torrens, P M, 2000b, "How Land-Use and Transportation Models Work", CASA Working Paper 20, University College London, Centre for Advanced Spatial Analysis (CASA). `http://www.casa.ucl.ac.uk/working_papers.htm`.

Torrens, P M & O'Sullivan, D, 2001, "Cellular automata and urban simulation: where do we go from here?", *Environment and Planning B* **28** 163-168.

Waddell, P A, 2000, "Towards a Behavioral Integration of Land Use and Transportation Modeling", in *The 9th International Association for Travel Behavior Research Conference, Queensland, Australia*. `http://www.urbansim.org/Papers/`.

White, R & Engelen, G, 1997, "Cellular automata as the basis of integrated dynamic regional modelling", *Environment and Planning B* **24** 235-246.

White, R & Engelen, G, 2000, "High-resolution integrated modelling of the spatial dynamics of urban and regional systems", *Computers, Environment and Urban Systems* **24** 383-400.