

**CENTRE FOR ADVANCED
SPATIAL ANALYSIS
Working Paper Series**



Paper 65

**AGENTS, CELLS
AND CITIES: NEW
REPRESENTATIONAL
MODELS FOR
SIMULATING
MULTI-SCALE
URBAN DYNAMICS**

Michael Batty



Centre for Advanced Spatial Analysis
University College London
1-19 Torrington Place
Gower Street
London WC1E 6BT

[t] +44 (0) 20 7679 1782
[f] +44 (0) 20 7813 2843
[e] casa@ucl.ac.uk
[w] www.casa.ucl.ac.uk

http://www.casa.ucl.ac.uk/working_papers/paper65.pdf

Date: September 2003

ISSN: 1467-1298

© Copyright CASA, UCL

Agents, Cells and Cities: New Representational Models for Simulating Multi-Scale
Urban Dynamics

Michael Batty

Centre for Advanced Spatial Analysis, University College London,
1-19 Torrington Place, London WC1E 6BT, UK

m.batty@ucl.ac.uk

Agents, Cells and Cities: New Representational Models for Simulating Multi-Scale Urban Dynamics

Michael Batty
m.batty@ucl.ac.uk

Centre for Advanced Spatial Analysis, University College London,
1-19 Torrington Place, London WC1E 6BT, UK

September 2003

Abstract

New forms of representation at a fine spatial scale, where units of space are conceived as cells and populations as individual agents, are currently changing the way we are able to simulate the evolution of cities and related systems. In this paper, we review progress to date in this field. We show how these new approaches are consistent with traditional urban models that have gone before with the emphasis no longer being on spatial interaction but on the dynamics of development and local movement. We first introduce a generic structure for urban simulation based on ideas about spatial evolution as reaction and diffusion, and then show how problems conceived in terms of cells, or agents, or both enable new implementations of this generic model. We sketch the rudiments of cellular automata (CA) which emphasises rules of development, and agent-based models which focus on how agents respond to attributes of their environment often encoded in cellular landscapes. We develop various exemplars based on residential location to impress the way these approaches work. Three applications are then presented at very different spatial scales: first pedestrian movement at the building scale, then the evolution of systems of cities at a country scale, and finally urban growth at the city scale. In developing these approaches, we show how cellular and agent-based models have the potential for explicitly incorporating spatial interaction and transportation which is their current weakness. We conclude with proposals that formal policy analysis in this domain should always be informed by more than one approach.

Acknowledgements

The ESRC Nexus Project (L326-25-3048) has provided partial support for this research. An early version of this paper was presented to the Conference on *Framing Land Use Dynamics*, held at the University of Utrecht, The Netherlands, April 16-18, 2003.

From Pattern to Process

The earliest mathematical models of cities, developed and implemented in the 1950s and 1960s, were focussed entirely on simulating patterns of land use and transport at a single cross-section in time. These models assumed that cities could be conceived of as structures with distinct causes and effects, that what was observed in cities could be factored into different spatial patterns with unique roles as either ‘independent causes’ or ‘dependent effects’. There was some logic to this separation. In the industrial city at least, residential patterns and related services were clustered around more basic export-orientated employment, while transport routes determining the relative pattern of accessibility, provided differential competitive advantage which in turn dictated where the best locations for various activities might lie.

What gave this approach credence was the idea that cities reflected a spatial and structural equilibrium which was relatively unchanging in time. If we could find strong enough correlations within spatial structure to enable robust predictions to be made from one pattern to another, then it mattered little how the transformation from independent to dependent variables was actually accomplished. In short, the process of moving from one pattern to another could be as simple as possible if the association between them was significant and robust enough. Thus models came to be fashioned around relatively simple associations at an aggregate scale where variations in spatial patterning were least. Parsimony was the goal with models being constructed to meet the traditional canons of scientific method: first the reproduction of existing urban spatial structures, ideally in many different places using the same model, and thence if their performance was good enough, their subsequent use in prediction.

As we now know, this rather superficial view reflects the fact that at an aggregative enough scale, all the volatility and dynamism of land use change in the city is ironed out, smoothed away. When we dig under this surface, this apparent equilibrium is far from being the stable and well-behaved system that we once assumed. In fact, most traditional urban models did not suffer from the problem of wild predictions in that their users were sufficiently wily to be able to ‘massage’ them into contexts in which

they were most useful. But there were worrying structural features in such models that could lead to untenable forecasts. Used sensibly, such models provide good short term forecasting techniques and there are many versions still in use today. They continue to form the cutting edge of operational land use-transportation modelling (Wegener, 1994).

The main criticism of these models however does not revolve around their equilibrating structure and aggregative nature *per se*. It is more that they do not address the concerns of contemporary planning and policy analysis, now strongly orientated to questions of regeneration, segregation, polarisation, economic development, and environmental quality. Urban sprawl and transportation are still within their ambit but these new problems exist at a scale that these models do not reach. To develop models which simulate how much finer scale actions take place requires significant disaggregation often to the point at which individuals and certainly groups need to be explicitly and formally represented. And it is a consequence of such disaggregation that temporal change comes much more centrally onto the agenda. There is no avoiding the fact that simulating individual behaviour at the finest spatial scales must take place over time.

So in the quest to address different kinds of urban issues, we come full circle again to the need for simulating urban dynamics; in short, we need to simulate the processes that lead to the spatial patterns that we observe as if cities are in equilibrium. None of this is very new for there have been many attempts at making static models dynamic (Forrester, 1969; Batty, 1971) as well as a series of attempts to ground these traditional models in the new nonlinear mathematics of chaos, catastrophe, and bifurcation (Wilson, 2000). But all these explorations have failed to yield models that are significantly different from their aggregate predecessors. What has happened is that a new generation of thinking, based not on aggregative, equilibrium seeking assumptions but without any formal assumptions about dynamics whatsoever has emerged, consistent with models of how activities cooperate and compete in producing emergent social structures from the bottom up (Epstein and Axtell, 1996). These are the models that we will review in this paper. We will show how new ways of representing change in urban systems through rule-based decision processes can be

articulated through new ways of representing urban elements as individuals which we call ‘agents’ and ‘cells’ defining their location.

We will begin outlining principles for urban simulation which tie these newer styles of model to the old, and then we will define new ways of representation. To link the old to the new, we present three exemplars: first a traditional urban model in which one spatial pattern – residential location, is immediately predictable from another – accessibility, with this transformation based on an arbitrary allocation or assignment not matched to the reality of how this actually happens; second, a model predicting the same but with the spatial assignment subject to an evolution at the most local level where what has already happened dictates, to an extent, what will happen; and third, a model in which this evolution is given active form through agents that compare different places in their quest to optimise their residential location. These three exemplars set the scene for our review of new models at three different spatial scales: first we simulate the local dynamics of movement in buildings and streets at the very small scale where agents are walkers or pedestrians; second we move to the very large scale using these agents to grow a system of cities, evolving a landscape where the migration of agents generates urban agglomerations; and third we generalise these agents to cells, simulating the evolution of a metropolitan area at the meso- or intermediate scale. Our exemplars and applications provide us with clear conclusions on the problems and opportunities this new perspective offers.

Generic Structures for Urban Simulation

All urban models can be written in functional form as some convolution of independent and dependent variables, parameters, and random errors which encapsulate noise or uncertainty in data and behaviour. Assuming K dependent variables $Y_i^k, k = 1, 2, \dots, K$, M independent variables $X_i^m, m = 1, 2, \dots, M$, $K + M$ parameters $\lambda^\ell, \ell = 1, 2, \dots, K + M$, and a composite random variation term ε_i^k where the subscript i refers to a location, a census tract, a cell, a parcel and so on, the generic model can be specified as

$$Y_i^k = f\{\mathbf{Y}_i, \mathbf{X}_i, \mathbf{\Gamma}, \varepsilon_i^k\} \quad . \quad (1)$$

The **bold** symbols define appropriate vectors of variables and parameters. This structure contains two features which are often missing in specific models – the positive feedback effect associated with the dependent variables and the error term absent from deterministic models. Feedback is relevant to systems of equations such as (1) where several different dependent variables are being simulated as, for example, in econometric models which are usually specified in linear simultaneous form. Equation (1) thus implies a process which is temporal in an implicit sense for it is inconceivable that such causal effects take place instantaneously.

The simplest model from equation (1) scales one spatial pattern $\{X_i^1\}$ to another $\{Y_i^1\}$ as $Y_i^1 = \lambda X_i^1$ where the causal effect is a proportionate one. In fact, this structure was used quite widely for the first land use-transportation models in the 1950s. Hansen (1959) proposed his residential location model in these terms as

$$P_i \sim \sum_j \frac{E_j}{d_{ij}^\alpha} \quad , \quad (2)$$

where P_i is defined as population in i , E_j as employment in j , d_{ij} the travel distance (or travel cost) between i and j , and α a parameter reflecting the friction of distance. Equation (2) is the well-known measure of potential or accessibility (Stewart and Warntz, 1958), in this case the accessibility of residential location i to all employment locations j . In applications of this model, equation (2) is usually scaled to ensure that the total population sums to a predetermined total $P = \sum_i P_i$. When an increment of population ΔP is to be allocated, the summation is $\Delta P = \sum_i \Delta P_i$ where ΔP_i rather than P_i is predicted from equation (2).

The largest number of land use transportation models developed to date have been based on accessibility equations of this kind. The most widely used structure involves a generalisation of this to two dependent activities – population and employment as

$$\left. \begin{aligned} P_i(z+1) &\sim \sum_j \frac{E_j(z)}{d_{ij}^\alpha} \\ E_j(z+1) &\sim \sum_i \frac{P_i(z)}{d_{ij}^\beta} \end{aligned} \right\} , \quad (3)$$

where the iterator z simply represents the way positive feedback and simultaneity enter the computational process. Iterating in this way until convergence (which is usually guaranteed) provides a solution to the nonlinear equations in (3). Note that there are now two parameters on distance, α and β , which ensure the system can be calibrated. Scaling parameters to enable correct orders of magnitude to be predicted can also be introduced to provide an appropriate computable form.

The first model based on equations (3) is due to Lowry (1964) although it usually involves another independent variable – basic employment – which is added as a driver to the second (employment) equation in the system. Several well-known versions of this model exist (see Batty, 1976) in which the simultaneity between employment and population is conceived of as a multiplier process, thus breaking the positive feedback cycle proposed by Lowry. The structure as stated does not compute spatial interaction *per se* although this is implicit in the definition of accessibility. A key interest in this structure relates to the process used to map the two patterns of population and employment into one another. This process is implicitly dynamic in a somewhat artificial way in that the model is started with estimates of population and employment often taken from observed data, and then driven to solution through computer time. This iterative process can be thought of as having a parallel in real time which has been exploited in some temporally dynamic versions of these models (Batty, 1984) although the usual way of making these models dynamic is to simply compute the increment or decrement of activity through time using the same structure.

A more appropriate temporal extension grounded in the dynamics of physical processes involves its specification in reaction-diffusion equations. Here we must introduce scalars directly into the model for these structures, unlike those on which traditional land use-transportation models have been built, deal with both growth (and

decline) of activities in time as well as their distribution across space. We will only state the model for the population equation for others follow by analogy. Then

$$P_i(t+1) = \lambda P_i(t) + \psi \sum_j \frac{E_j(t)}{d_{ij}^\alpha} + \varepsilon_i^P(t) \quad , \quad (4)$$

where the postscripts t and $t+1$ refer to time instants, the scaling factors λ and ψ enable appropriate magnitudes to be grown, and the error term $\varepsilon_i^P(t)$ is simply indicative of need to provide some noise within the dynamic process in contrast to the previous spatial allocation models where random error is not usually considered. In equation (4), the first term on the RHS is the action/reaction while the second – the accessibility term – is the diffusion element. In this sense as population changes through time, it is always a function of population at the previous time period (the positive feedback effect) and employment in other locations which acts to diffuse population around its centres. It is worth noting that structures such as these when made operational, are often tempered against sets of constraints which enable discontinuities and thresholds to be met. For example, the Lowry models in equations (3) above, are usually subject to capacity constraints on the density of cells while reaction-diffusion equations like (4) are often operationalised using cellular automata (CA) methods which are discrete, often in binary form as we will illustrate in the next section.

Finally we need to show how these structures are consistent with movement through the diffusion effect. In a way, all the models we have stated so far have interaction between locations implicit in their form through the definition of accessibility. If we write the accessibility term in equation (2) as

$$p_{ij}(t) \sim \frac{E_j(t)}{d_{ij}^\alpha} \quad , \quad (5)$$

where the term $p_{ij}(t)$ is now the explicit interaction/diffusion between i and j , we can state a more complete interaction relation adding the reaction and noise terms as

$$P_{ij}(t+1) = P_i(t) + p_{ij}(t) + \varepsilon_i^P(t) \quad . \quad (6)$$

If we now use spatial interaction accounting to examine the population in i at $t+1$, we derive the conventional reaction-diffusion equation in (4) above as

$$P_i(t+1) = \sum_j P_{ij}(t) \sim P_i(t) + \sum_j \frac{E_j(t)}{d_{ij}^\alpha} + \varepsilon_i^P(t) \quad . \quad (7)$$

In fact there is really a rather strong tie-up to traditional spatial interaction theory and gravitational modelling of city and transport system using these forms. If we add the reaction term into the product of the diffusion term and forget the noise, then equation (6) becomes

$$P_{ij}(t) = P_i(t) \frac{E_j(t)}{d_{ij}^\alpha} \quad , \quad (8)$$

and with appropriate scaling, total population and total employment can be computed in the same way as constraints are handled in Wilson's (1970) family of spatial interaction models. In short, it is quite feasible in dynamic model structures such as these to incorporate relevant conservation laws with respect to the total activities generated.

We have presented these models in this form so that we can see how disaggregation leads to objects of interest which have a degree of discreteness and wholeness quite different from traditional conceptions of population and employment, hitherto the main working variables of traditional land use-transportation models. At much finer spatial scales, we reach a level where the zones or tracts – cells we will call them – become so small that it is appropriate to consider each to have only one state. In other words, imagine that cell i houses only one unit of population, one household. If a cell were to be developed, then $P_i(t) = 1$, if not, then $P_i(t) = 0$. At this point, we might decide that the population might be better represented as something other than a cell, for example as an independent object or agent. However the cell would still be of interest in that it would contain land on which the household resides and as such

would act as a source for attributes of the built or natural environment. This idea clearly changes the nature of the above equations but only in the way they are computed. Their generic form does not change. But to make this clear, we must first define the rules needed to work with cities which are represented by cells or agents or both.

Representations and Aggregations: Cells, Agents, Neighbourhoods and Rules

We will still define a cell as a location i but with the understanding that magnitudes associated with an activity in any cell are usually computed by adding cells within some larger neighbourhood. Most of our traditional models can be defined in terms of cells but when it comes to the definition of agents, there is no such association with particular locations. An agent k is thus an object $\{w^k\}$ which has attributes and at any point in time t is associated with a cell i . In this context, it is defined as $w_i^k(t)$; more than one agent can exist or be associated with a cell and agents can of course move between cells. This changes our conception yet again in that the magnitudes of agents associated with any cell i are computed by adding up the number of agents in i . Agent-based representation is thus much more general than the cellular in that agents usually exist on a landscape of cells whereas cellular models associate agents directly with cells: in CA, cells are agents and vice versa while in agent-based models, cells and agents are quite separate from one another.

We will deal with cellular models first. The dynamic which drives such models involves defining rules which enable a cell i to change its state over some interval of time from t to $t+1$. There is usually a limited number of states which any cell can take on, the simplest being developed or non-developed, urban or rural in which the variable $P_i(t)$ is defined appropriately as 1 or 0. Change in any cell is some function of reaction, diffusion and/or randomness as implied in equation (4) but the critical feature of cellular automata models is that the neighbourhood over which diffusion takes place is strictly limited. In traditional CA, this neighbourhood removes any ‘action-at-a-distance’ with interaction/diffusion confined to the cells immediately

adjacent to the cell which is the recipient of such change. This is based on the quite obvious notion in physical systems that when a gas or liquid diffuses, its constituents must move to adjacent locations or cells in order to travel any distance. In urban models, this is appropriate for pollution or walking, for example, but it is problematic for movement involving more manufactured kinds of technology. People can hop over intervening distances between places but particles cannot, and as such, this represents the most significant limit on the use of strict CA in urban simulation. In fact, this neighbourhood restriction is relaxed in many applications where it is more appropriate to refer to these as cell-space (CS) rather than cellular automata (CA) models (Albin, 1975; Couclelis, 1985).

The key reason why locality is so important in CA revolves around the idea that local action leads, in many circumstances, to global order, to emergent structure. Often local rules which are applied routinely lead to structures in the large that look highly ordered but cannot be predicted from any top-down command-like process or model. CA are thus excellent examples, as we will see below, of local rules which lead to surprising and possibly unexpected structures. For example, patterns of global segregation can be produced under very mild conditions of preference at the local level; if you prefer to live, say, with at least the same number of people around you who share your preferences, meaning that you would be equally at home with as many people around you with a different preference, then given enough slack in the system and some places where this is not satisfied, it is easy to show how the entire system can unravel, leading to extreme patterns of segregation where most will live with those around them sharing an identical opinion (Schelling, 1978).

To demonstrate how CA actually works, let us implement the reaction-diffusion structure in equation (4) first without noise and then with. We will structure the solution by examining reaction first and then if this test is passed, consider the diffusion. The cellular array consists of developed $P_i(t)=1$ and non-developed (empty) $P_i(t)=0$ cells. We state that a cell can only change its state – react – if it is empty, meaning that already developed cells will not change and remain developed. For reaction

$$\text{if } P_i(t) = 0 \begin{cases} \text{then } \textit{begin the diffusion test} \\ \text{otherwise } P_i(t+1) = 1 \end{cases} . \quad (9)$$

If this test is passed, we then see whether or not development will diffuse to the empty cell. That is

$$\text{if } \Phi_{\min} \leq \sum_{j \in \Omega_i} \frac{P_j(t)}{d_{ij}^\alpha} \leq \Phi_{\max} \begin{cases} \text{then } P_i(t+1) = 1 \\ \text{otherwise } P_i(t+1) = 0 \end{cases} , \quad (10)$$

where Φ_{\min} is the minimum and Φ_{\max} the maximum access thresholds of the neighbourhood Ω_i that need to be met if the cell is to be developed. In fact, this access threshold is entirely local being simply a count of developed cells. If we define $\Omega_i = \{N, S, E, W\}$, cells which are north, south, west and east of the cell i – the so-called von Neumann neighbourhood (Toffoli and Margolis, 1987) – then each distance in equation (10) is the same, that is $d_{iN} = d_{iS} = d_{iE} = d_{iW} = 1$, and (10) reduces to a count of cells in the neighbourhood.

There are many different rules based on neighbourhood counting. As Wolfram (1994) shows, there is a combinatorially explosive number of rules for even the simplest of two-dimensional CA and it is not usually possible to classify these. There are other rules too, based on voting, for example, as in counting cells associated with preferences in the Schelling (1978) model where states are changed subject to thresholds defined in terms of majorities or minorities. There are different kinds of simple neighbourhood to consider such as the 8 cell Moore neighbourhood in contrast to the 4 cell von Neumann. All these give rise to many possibilities but to illustrate the essential feature of CA, we will show what happens when we plant a seed in the centre of a square cellular space and then grow the structure – city if you like – around its central business district (CBD). There is one case worth noting before we continue. We have assumed that a reaction only takes place if the cell is empty as in equation (9) but if this rule abandoned, then only the diffusion takes place through equation (10). If we set $\Phi_{\min} = 2$ and $\Phi_{\max} = 3$, this generates Conway's Game of Life (Gardner, 1970). When the number of cells is less than 2, a cell which is already

developed ‘dies’ through isolation. When the number is greater than 3, the cell ‘dies’ through overcrowding. When between 2 and 3, an already developed cell remains developed but an empty cell is also developed, giving rise to a ‘birth’. These simple rules lead to structures which are self-perpetuating as is ‘life itself’ (Poundstone, 1985).

In Figures 1(a) and (b), we show two different structures where $\Phi_{\min} = 0$ and $\Phi_{\max} = 1$, and where $\Phi_{\min} = 0$ and $\Phi_{\max} = 8$. The grey tones indicate the order in which the cells are developed. It is quite clear in Figure 1(a) how simple rules lead to global patterns; how local rules applied over and over again repeat themselves at different spatial scales generating fractal-like structures whose form is difficult to predict simply from a knowledge of these local rules. The completely filled pattern in Figure 1(b) can in fact be generated when every value of $\Phi_{\max} > 1$, due to the way the von Neumann neighbourhood restricts the counting rule. In Figures 1(c) to (d), we have added some noise, that is, we have made all the reaction and diffusion rules in equations (9) and (10) subject to a meta rule which we specify as

$$\text{if } P_i(t+1) = 1 \text{ and } \varepsilon_i^P(t+1) < \Lambda \quad \begin{cases} \text{then} & P_i(t+1) = 1 \\ \text{otherwise} & P_i(t+1) = 0 \end{cases}. \quad (11)$$

Λ is a threshold above which if the random event $\varepsilon_i^P(t+1)$ occurs, the development is sustained. This is accomplished through drawing random numbers with appropriate scaling and using these values to modify the diffusion threshold. In these instances, it is clear that the patterns generated are much more irregular but still with some semblance of the global patterning of Figures 1(a) and (b).

Introducing agents into this mixture adds an entirely new dimension to such dynamics. An agent defined as $w^k = 1$ where we now assume there are K agents in total, $K = \sum_k w^k$, always exists with reference to a location i as $w_i^k(t)$, and the number of agents in any cell is $w_i(t) = \sum_k w_i^k(t)$. There is a special class of such agent-based models called ‘active-walker’ models (Kayser, Aberle, Pochy, and Lam, 1992). Here there is a strict separation between the landscape of cells which we define as before as

$\{P_i(t)\}$ and the walkers who exist on the landscape as $\{w_i^k(t)\}$. In active-walker models, agents change the landscape on which they walk and the landscape changes the agents in that it directs them where to walk. This can be specified in the set of coupled equations

$$\left. \begin{aligned} P_i(t+1) &= f[w_i^k(t), P_i(t)] \\ w_i^k(t+1) &= g[P_i(t), w_i^k(t)] \end{aligned} \right\} \quad (12)$$

The two functions f and g represent the landscaping function and the walker movement function respectively. Agents affect the future landscape by having been there, by walking upon it, while the landscape affects the walkers in terms of where the walkers are able to walk (and possibly other of their attributes).

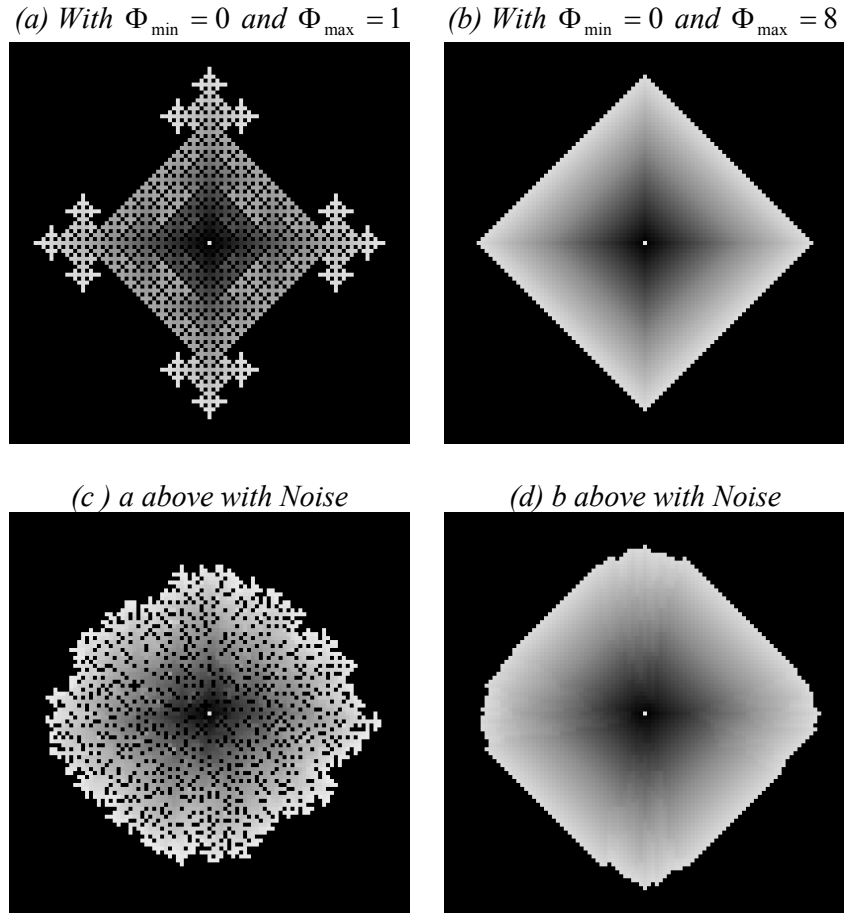


Figure 1: Local Rules Leading to Global Pattern: Generating Concentric Structures in von Neumann Neighbourhoods with and without Noise

One of the features about ‘walking’ is that the activity is entirely local. Walkers only move to adjacent cells and like CA, the same kinds of local rules and neighbourhoods apply. Let us now write the joint reaction-diffusion movement function as

$$P_{ij}(t) = \frac{P_i(t)P_j(t)}{d_{ij}^\alpha}, \quad (13)$$

where this interaction is only possible if i and j are part of the same local neighbourhood Ω . In this case, then $d_{ij} = 1$ and $P_{ij}(t) = P_i(t)P_j(t)$. We now define the landscape function as an encoding of the geometry of the system: if $P_i(t+1) = 0$, then this means it is possible to walk on the space, the space is empty and not occupied whereas if $P_i(t+1) = 1$, this means it is occupied, with a building or is illegal for walking upon. This implies the only possible moves in the system are given by the matrix elements defined by the conjunction $\rho_{ij}(t) = P_i(t) \wedge P_j(t) = 0$. It is now very easy to fashion a simple walking model. An agent at i , $w_i^k(t)$ can walk to j as $w_j^k(t+1)$ if and only if $\rho_{ij}(t) = 0$. However this would result in mindless walking because there is nothing else on the landscape to direct motion. A feature of the problems that we will illustrate here is that there is an objective in terms of walking and this can be encoded into the landscape as some form of location specific resource $R_i(t)$. This may vary through time by being consumed or replenished but it serves to direct the walker into available cells in which the resource is optimal in some local way. Thus our walking model might be written as

$$\text{if } \rho_{ij}(t) = 0 \quad \text{and} \quad R_j(t) = \max_{\ell \in \Omega_i} \{R_\ell(t)\} \quad \begin{cases} \text{then } w_i^k(t) \rightarrow w_j^k(t+1) \\ \text{otherwise } w_i^k(t) \rightarrow w_i^k(t+1) \end{cases}. \quad (14)$$

The reaction-diffusion structure is complete when we add some noise. As with our other models, randomness is essential when we have many walkers so that we can simulate slight deviations from intended direction and other elements of real world uncertainty (Helbing, Molnar, Farkas, and Bolay, 2001).

This kind of algorithm can be easily generalised and embedded into the generic structures introduced previously. The local optimality which this implies is easily visualised if we imagine an accessibility surface focused on the centre of a city. A walker arrives at the edge with a view to finding the most accessible location. The walking algorithm in equation (13) not only enables the walker to climb this surface to the optimal point but also to circumnavigate any obstacles. If the access surface is perfectly symmetrical and convex-up centred on a single point, the CBD say, then the algorithm will find the *optimum optimorum*. If there are instabilities in the surface, the algorithm will detect local optima and thus in real problems additional mechanisms are likely to be required to reduce the impact of any sub-optimality. Many examples of these types of model exist in the simulation of pedestrian behaviour and these are being generalised to other kinds of human motion using active particle techniques (Schweitzer, 2003). However the most impressive attempt to date for socio-economic system is the Sugarscape model (Epstein and Axtell, 1996). Here agents optimise their economic performance in climbing a Sugarscape, a resource surface, which acts as a metaphor for search through hill climbing to achieve an optimal consumption of resources.

Exemplars: Static Patterns, Cellular Growth, and Agent-Based Diffusion

We have now sketched the rudiments of a general structure for urban simulation although we will finish our elaboration far short of a well worked out theory. At this point, we will shift tack to illustrate our theories in more practical terms as it is worth emphasising just how close the different modelling strategies and paradigms of the last 50 years are in terms of the way activities and land uses in cities can be simulated. In this section, we will introduce three hypothetical applications – exemplars of the way cells and agents can be used to represent both static and dynamic models which build on the structures in the last section. Our focus will be upon explaining residential location in terms of cells becoming occupied by households who seek to optimise their accessibility.

At the heart of urban economic theory, lies a trade-off between a consumer's demand to minimise distance travelled to various activities and a desire to capture as much living space as possible. This is the theory first formally articulated by Alonso (1964) where it was assumed that in the monocentric industrial cities, residents arranged their locations around the CBD according to this trade-off between distance (travel cost) and space. The structure of preferences and the market for various land uses appears to have led to wealthy groups being able to capture more space at the edge of the city than poorer groups who have been confined to the inner areas around the CBD.

Our first exemplar operationalises Hansen's (1959) residential model where accessibility in equation (2) is measured on a cellular landscape around a CBD. Travel accessibility is measured lineally using 'negative' distance $D - d_{i0}$ while space available is measured as a nonlinear 'positive' function of distance ξd_{i0}^η . D is some limiting distance at the edge of the city, ξ and η are parameters, and the CBD is defined where $j = 0$. In fact, it makes more sense presentationally to describe locations i, j in terms of coordinates x, y . Thus we define travel accessibility d_{xy} and space available s_{xy} respectively as

$$\left. \begin{aligned} d_{xy} &= D - \{|x - x_o| + |y - y_o|\} \\ s_{xy} &= \xi \left[\{(x - x_o)^2 + (y - y_o)^2\}^{1/2} \right]^\eta \end{aligned} \right\} , \quad (15)$$

where it is assumed that the values associated with cell x, y are with respect to the CBD at x_0, y_0 . Note that to differentiate the two equations, we define a Manhattan-like distance where travel to the CBD is always at right angles for the travel access/cost and a direct crow-fly distance to the CBD with respect to the space available.

Accessibility T_{xy} is now defined as a simple sum of these two components in equation (14) as

$$T_{xy} = d_{xy} + s_{xy} . \quad (16)$$

The model is very easy to state. We define one unit of population (a household, say) associated with a cell as $P_{xy} = 1$, the total units of population to be allocated as $P = \sum_{xy} P_{xy}$, and we allocate population to cells so that

$$\text{if } P_{xy} = 1 \text{ then } T_{xy} > T_{x'y'} \text{ where } P_{x'y'} = 0 \quad . \quad (17)$$

This is a simple assignment that ensures that every household is in a cell which has higher accessibility than the empty cells and that this allocation is exhaustive. In fact the discreteness of the formulation leads to a slightly different allocation from that associated with equation (2) where the number of units is scaled according to the value of the accessibility with all units having some value, no matter however small. However with careful definition, this discrete CA representation can be made to replicate its continuous equivalent in equation (2).

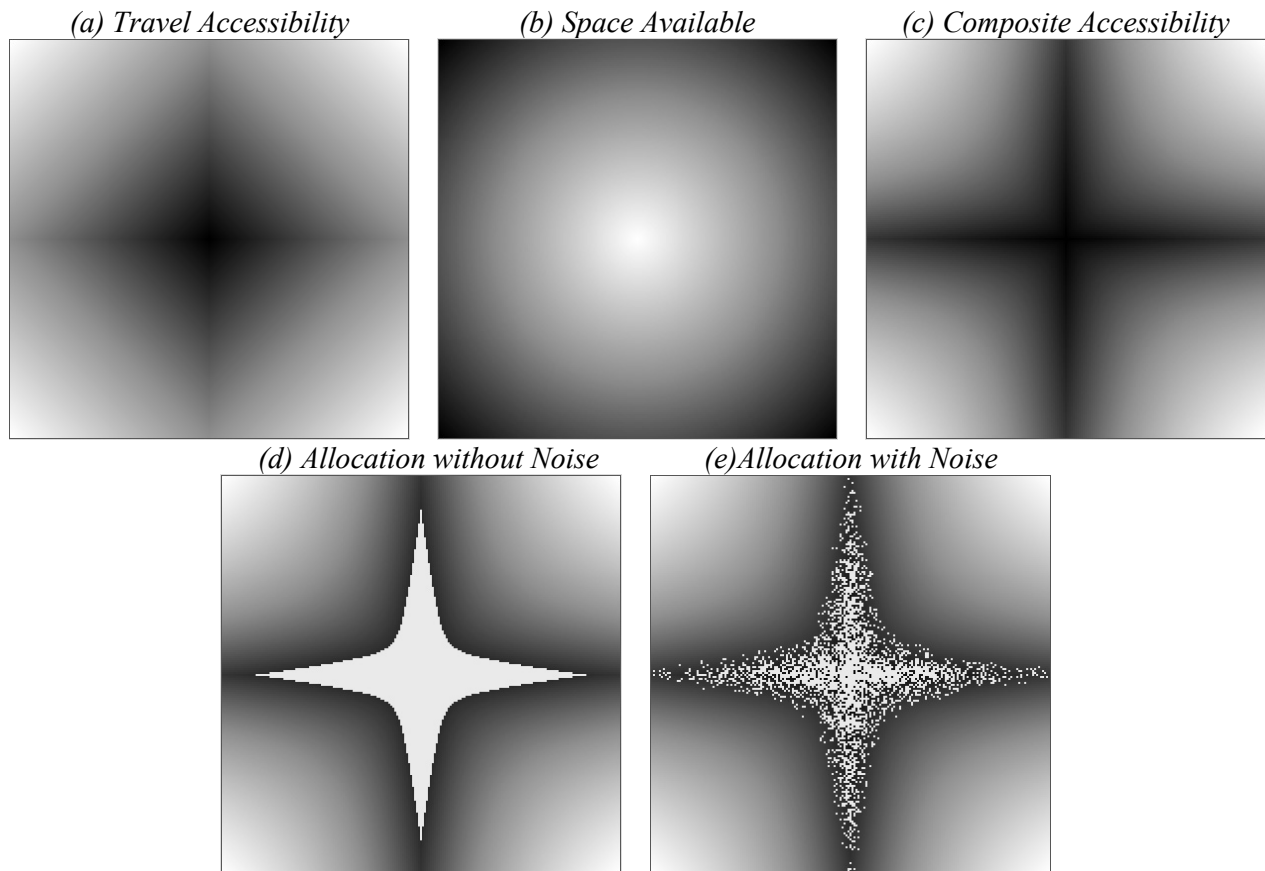


Figure 2: Hansen's (1959) Accessibility Model Implemented as a Cellular Automata with and without Noise

To illustrate its application, we use a cellular space of 201 x 201 pixels where the central cell 101, 101 is defined as the CBD. In Figure 2(a), we show the linear travel accessibility, in Figure 2(b) the space available, and in Figure 2(c), the aggregate accessibility, all defined using equations (15) and (16) where $D = 101$, $\xi = 0.4$, and $\eta = 0.1$. In Figure 2(d), we allocate 4000 households units where it is clear that the symmetric pattern generated simply reflects the form of the aggregate accessibility in Figure 2(c). This model is entirely based on an implicit diffusion process with no growth or decline (reaction) and no error or noise. It is easy enough to add noise as ε_{xy} and using an appropriate scaling, allocating 2000 households leads to the pattern shown in Figure 2(e). It is clear from these images that this kind of model is simply a mapping of one pattern onto another. Its morphology, even with some randomness, is entirely symmetric and hence predictable and there is no irregularity in terms of the morphology which is generated. It has all the hallmarks of a system generated from the top down without the kinds of dynamics and bottom-up processes which are so necessary to understand how cities change and evolve.

Our second example is even simpler in that our conception of travel accessibility and space available is entirely local. This is a model where we seek to locate households around a CBD which is the first active location initiating the development process. We will now mix coordinate with index notation. For development to occur at time $t + 1$ in cell i , the cell must be linked to the growing city – that is, it must be adjacent to some already developed cell $P_j(t) = 1$, any $j \in \Omega_i$, and the amount of space around this cell i must be a maximum for the system. We can easily implement these rules using the following conditional:

$$\left. \begin{array}{l} \text{if } P_i(t) = 0 \text{ and } \sum_{j \in \Omega_i} P_j(t) > 0 \text{ and } N_i(t) = \min_i \sum_{j \in \Omega_i} P_j(t) \\ \text{then } P_i(t+1) = 1, \text{ otherwise } P_i(t+1) = 0 \end{array} \right\} . \quad (18)$$

$N_i(t)$ is the number of neighbours around i which needs to be a minimum for the number of empty spaces to be a maximum and the cell to be developed.

We begin this process with the central cell as the seed, that is $P_{i=x_0, y_0} = 1$ and this leads to the structure shown in Figure 3(a) which is shown where 4750 cells have been occupied. It is clear that as the structure grows, it becomes compact as the average number of neighbours for the entire space begins to converge. The pattern of these numbers of neighbours is shown in Figure 3(b), and for the overall system at $t = 1600$ when the simulation is stopped, this average is 2.78. Although this pattern is clearly something which can only be grown from the bottom up and thus reflects all the principles of local cellular action, there is still a uniformity about its morphology which is unrealistic. In fact we could introduce some noise and this would distort the pattern in terms of aging but it would not distort the morphology. To generate more realistic structures, we need to turn to our third exemplar which adds the notion of active agency to cellular action.

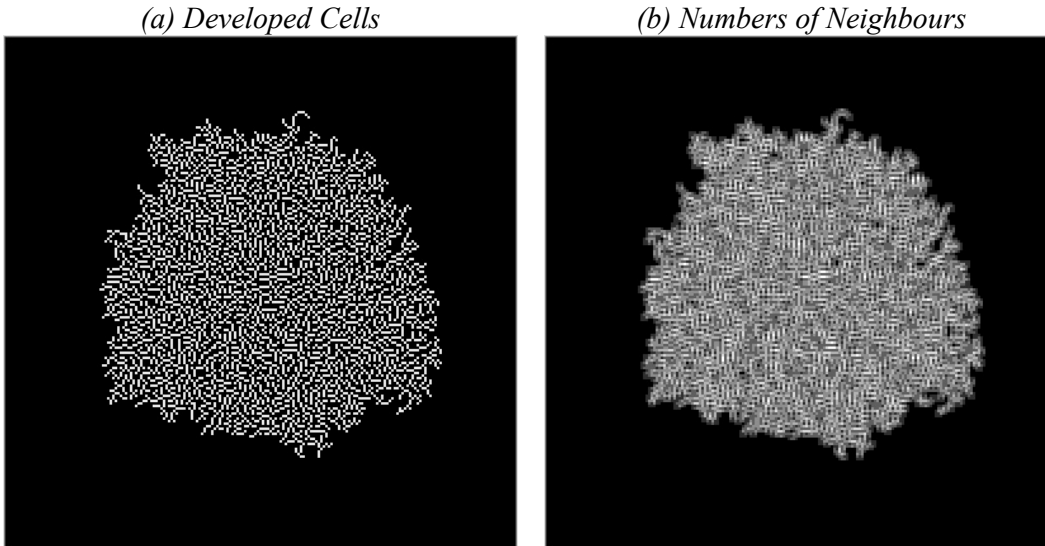


Figure 3: Connected Cellular Growth Around a Central Seed with Local Maximisation of Residential Space

This third model adds a new layer to cellular automata which in one sense might be thought of as an active, purposive dimension. Essentially our model is the same as the previous two in that all the action begins in or around the CBD. Agents w^k are launched at the edge of the city space and then engage in a random walk. If they walk outside the city space, they are moved back to the edge but if they walk to a cell adjacent to a cell that has already been developed, given by $P_j(t) = 1$, any $j \in \Omega_i$, then they will decide to locate there, the cell i in question being developed, that is

$P_i(t+1) = 1$. The simulation is begun with the central cell – the CBD cell – being developed, and with all the walkers being located randomly at the edge of the city space, that is, $w_i^k(0), \forall k$ where $d_{i=xy} \geq D$. Random walking is simply a method of exploring the space, and in this model, walkers move randomly to adjacent cells in the von Neumann neighbourhood $j \in \Omega = N, S, E, W$. Essentially the structure grows when walkers make first contact with developed cells and as in our previous model, the fact that cells are developed one by one ensures that the structure remains connected. In terms of maximizing space available, walkers have a greater probability of locating at the edge of the structure than towards its centre due to the way they are initiated.

Formally we can pose this model as follows. At each time t , we execute the test

$$\left. \begin{array}{l} \text{if } w_i^k(t) = 1 \text{ and } \sum_{j \in \Omega_i} P_j(t) > 0 \text{ and } P_i(t) > 0 \\ \text{then } P_i(t+1) = 1 \text{ and } w_i^k(t) \rightarrow w_z^k(t+1) \\ \text{otherwise } P_i(t+1) = 0 \text{ and } w_i^k(t) \rightarrow w_j^k(t+1) \end{array} \right\}, \quad (19)$$

In the first line of equation (19), we test to see if the cell location in which the walker is located is empty and if there is development in its neighbourhood. In the second line if this is so, the cell is developed and the walker returns to its initiating point on the edge of the city space where the index $z = \text{random } x, y \text{ with } d_{xy} \geq D$. The third line of the test is associated with failing the test in the first line and then the walker simply continues walking to a cell in its neighbourhood, chosen randomly as $\text{random } j \in \Omega_i = N, S, E, W$.

This is the very well known model of diffusion-limited aggregation (DLA) which is central to the physics of far-from-equilibrium systems. It was first introduced in this form by Witten and Sander (1981) and has been used to grow many kinds of structures which have a dendritic structure. Essentially what is generated is surprising. Unlike the previous model where the mass is compact, this structure is much more tree-like in form with branches reaching out to capture as much space as possible, not

unlike the way cities grow into their surrounding hinterland. It is essentially a fractal structure and its morphology can be tuned to produce dendrites of differing compaction with varying fractal dimensions. The di-electric breakdown model represents an equivalent form. Rather than being conceived as a bombardment of a growing mass with particles from the edge, growth takes place from the centre where tips of the evolving structure have the greatest probability of growth (Niemeyer, Pietronero, and Wiesmann, 1984; Stanley, and Ostrowsky, 1986).

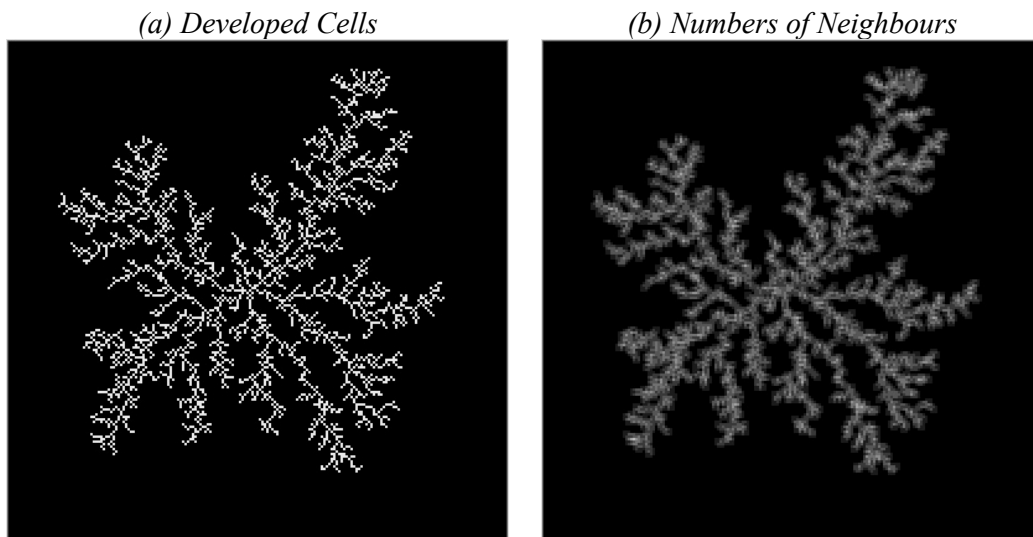


Figure 4: Diffusion-Limited Aggregation: Cellular Growth from Agent-Based Random Walks

We illustrate the structure on our 201 x 201 cellular space in Figure 4(a) where we show its morphology and in Figure 4(b) where we plot the number of developed neighbours associated with each of its cells. In contrast to our previous model, it is clear that as the structure grows, the average number of neighbours declines inexorably as the development reaches out into greater and greater regions of empty space, although the average number of those cells which have at least 1 developed neighbour is about 2.3, only a little less than in the previous model. This kind of irregularity can only be generated from the bottom up. It is a product of randomness with locational principles based on keeping the structure connected, agglomerated, combined with the search for greater and greater space in which to grow. It is as good an example of a reactive-diffusive structure as we have and in a sense it provides the baseline exemplar for the various applications that we will now elaborate.

Simulation at the Very Smallest Scales: Pedestrian Movement in Buildings and Streets

Our three applications represent a classification of problems at different scales which also reflect different varieties of dynamics and different assumptions about the extent to which cells and agents engage in goal-seeking activities. At the smallest scale in built environments, routine, repetitive movement based on ‘fast’ dynamics is the focus where the frequency of interaction is measured in terms of seconds and minutes, sometimes hours and days but never any longer. Such activities usually respond to the environment through agents ‘using’ what has been already created rather than recreating it which takes place over much longer time periods. In contrast, at the very largest scales where we are dealing with systems of cities in Berry’s (1964) terms, the dynamics is ‘slow’. These are based on decisions which take place much more infrequently through migration, decisions to establish new settlements that evolve over years, probably decades, if not centuries and eons. Somewhere in between, we will deal with the city, the meso-scale –the focus of our exemplars so far – where the temporal scale is over years and decades. In fact associating slow with the large scale and fast with the small is always an oversimplification for in some applications, slow and fast exist together.

In terms of purposive activities, spatial and temporal scale tend to determine the degree to which agents and cells react to each other in terms of change to their environment. We will begin with models that simulate small scale movements over short time intervals in buildings and streets where the focus is on visiting places rather than on making decisions which change the composition of activities associated with those places. We have already mapped out a general structure for such agent models in which movement – walking – was articulated as the intersection of geometry defining where one might walk with responses to location attractions based on resources defining the landscape on which movement takes place. We will retain our previous notation: walking takes place on and between cells i and j adjacent to one another in appropriately defined neighbourhoods $j \in \Omega_i$, which meet the requirement that both are empty of any other activity, that is $\rho_{ij} = 0$. For movement in streets and buildings, we assume that the matrix \mathbf{p} does not vary through simulation time, and

thus it defines the ‘container’ within which walkers are able to respond to the resource landscape $\{R_i\}$ which is also unchanging in time. In fact, our models at this scale are not active-walker models at all, but passive-walker models. The landscape never changes although walkers do respond to each other which is an additional feature of models at this scale.

There are many variants ranging from those in which geometry is all important to those where the accessibility of resources takes precedence. Where geometry is important, these models deal with very fine scales at the level of corridors and rooms and tend to be used to predict panic situations and evacuation events in hazardous environments. Very detailed issues involving the physics of acceleration characterise these models (Helbing, Molnar, Farkas and Bolay, 2001). At the other extreme where geometry is unimportant but patronage of different locations is, as in shopping activities for example, the attraction surface is all important and this is reflected in the way such models are specified and implemented (Borgers and Timmermans, 1986). In our example here, both are important as we will simulate movement in a complex building where geometry does dictate where people go but the attraction of different exhibits is the prime reason why people move within the building in the first place. We will state the model by modifying equation (14) as follows. A walker will move

$$\text{if } \rho_{ij} = 0 \begin{cases} \text{then } w_i^k(t) \rightarrow w_j^k(t+1) \text{ where } j \leftarrow \max_j \{\nabla R_{ij} + \varepsilon_j^k(t)\} \\ \text{otherwise the walker engages in obstacle avoidance} \end{cases} . \quad (20)$$

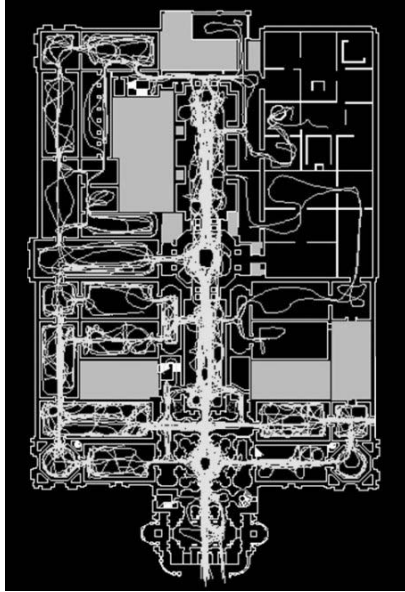
Note that we define ∇R_{ij} as the local gradient in the resource surface in the direction from i to j which is a maximum but add to this some random noise $\varepsilon_j^k(t)$. All this does is push the walker in the direction of greater resources. In fact obstacle avoidance is probably the more frequent occurrence in complex geometries and routines to effect this consist of moving walkers in different directions, dependent upon the previous history of how each walker has reacted to obstacles, how far they are able to see, and so on. The other feature of this model involves interactions between different walkers. There are limits on congestion which involve dispersing walkers if too many attempt to reach the same location. This is simply a matter of

ensuring that $\sum_k w_i^k(t)$ is within a certain threshold and initiating local movement if it is not. Panic can set in if congestion occurs across a wider region of cells and dispersion is not possible. In contrast, flocking or herding based on walkers ‘following the crowd’ is considered by assessing how the number of walkers $\sum_{j \in \Omega} \sum_k w_j^k(t)$ in a wider area Ω attracts even more walkers.

To illustrate this model, we have applied it to the movement of visitors in the Tate Gallery on London’s Millbank. We have good data on the circulation patterns of 550 visitors observed over a short time period in August, 1995 which we show in Figure 5(a) (UAS, 1996). At that time, paintings were on display in 49 rooms of the building and we simplified the problem to consider only those visitors – some 97% of those visiting in fact – who entered the Gallery through the main entrance. We did not deal with movement in the Clore Annex which is excluded from Figure 5. What we were interested in is the pattern of visitation to the various rooms. To measure this, we introduced walkers into the Gallery over a short period of time and then examined the pattern of room occupancy in the steady state which emerges when the model is run through many time periods. What we assess is the average number of walkers visiting each room Ω_r over T time periods which we compute from $\sum_{\tau=1}^T \sum_{j \in \Omega_r} \sum_k w_j^k(\tau) / T$. Once this quantity stabilises as it clearly does, we are then in a position to assess the ‘fit’ of the model to the actual pattern of room visitation.

One of the greatest advantages of agent-based models is that as we run the model with different numbers of agents, we can derive different kinds of information about the problem. In our Tate model, when we launch just one agent, we can consider this as a probe used to explore a complex space and in doing so, assess how well it is dealing with problems of obstacle avoidance which we illustrate in Figure 5(b). As we run the model with more, then many agents, we can also assess the role of randomness on the pattern of visitation and the steady state, enabling the ‘right’ level of variation in overall spatial behaviour to be defined. In Figure 5(c), we show a typical snapshot of agents within the gallery. It is immediately clear that although this does tend to show those rooms more frequently patronised as well as the position of different agents at

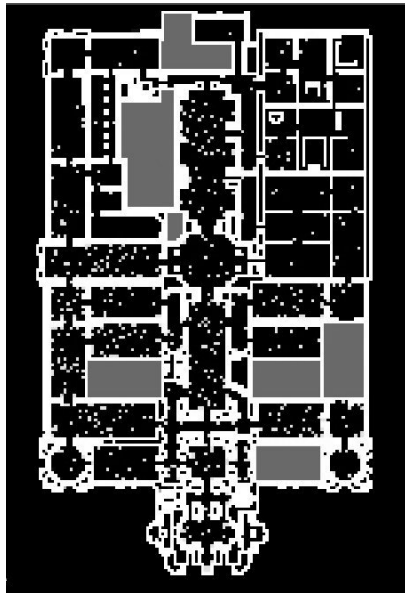
*(a) Walker Paths in the Gallery,
August 1995: 12noon to 12-15pm*



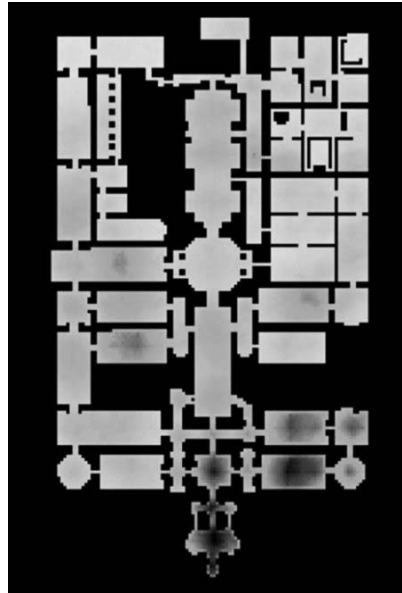
*(b) The Progress of a Single
Agent Exploring the Gallery*



*(c) A Snapshot of Agents Moving
in the Gallery*



*(d) The Steady State Pattern:
Visitors to Cells and Rooms*



*Figure 5: Movements of Real and Simulated Walkers
in the Tate Gallery, Millbank, London*

every point in the space, it is not possible to say anything about the long term visitation pattern. In short, this is as good an example as we have where cross-sectional patterns mean very little in terms of the longer term dynamics. In Figure 5(d), we show the average patronage in the steady state in each cell, not in each room although it is possible to get a sense of the frequency of room visitation from this. The use of this kind of model for ‘what if’ analysis is fairly obvious. In this case, closing or opening rooms or changing their configuration for various types of exhibition as well as showing how different kinds of attractions in rooms affect movement is what the application is all about.

We have but touched the surface of this kind of modelling as the field is currently exploding. In terms of using this particular model, our main applications have been to simulating shopping patronage in town centres (Haklay, Thurstain-Goodwin, O’Sullivan and Schelhorn, 2001), predicting shortest routes in pedestrian networks (Batty and Jiang, 2000), and modelling street parades, in particular the Notting Hill Carnival where public safety was the main focus (Batty, Desyllas, and Duxbury, 2003a, 2003b). Models at much finer scales involving panic and evacuation possibilities tend to include much more basic physics and there are strong links to CA models of traffic movement (Helbing, Farkas and Vicsek, 2000). Useful summaries are provided by Schreckenberg and Sharma (2002) and Vicsek (2002). There are also active-walker versions of these models where the landscape is altered by the act of walking. In particular, Helbing, Schweitzer, Keltsch and Molnar (1997) show how paths become established where none exist before as walkers move across a space, interacting with the tracks established so far and the paths of other walkers. These models are being extended to other kinds of economic system by Schweitzer (2003) and there are many applications which mirror animal movements in the biological sciences (see Camazine, Deneubourg, Franks, Sneyd, Theraulaz and Bonabeau, 2001). In all these cases, the models treat agents quite literally as distinct objects such as people but in their extension to larger scales and different temporal dynamics, the idea of the agent changes, becoming more abstract and instrumental. It is to these kinds of problem that we now turn.

Simulation at the Very Largest Scales: The Emergence of Systems of Cities

We will evolve a system of cities where walkers are ‘migrants’ and resources are their ‘jobs’. These are active-walker models where walkers respond to jobs which define an economic landscape which in turn directs where migrants search. Our landscapes have no geometry, being featureless plains in the grand sense, although it is entirely possible that geometric obstacles could be introduced and thus the mechanisms of the very small scale (in the previous section) might feature in directing walkers. The ultimate model is in fact composed of two such landscapes, the first linking people to jobs through physical networks based on origins and destinations, the second being defined in terms of ‘resource’ potential which enables new walkers – new migrants to the system – to be located to meet an appropriate growth rate. Positive feedback and diffusion are reflected in the way both landscapes are formed. It is quite possible to specify models where these landscapes are compatible but separate, networks simply being the consequence of where jobs and people are located. However in the model we will present, these landscapes interact with one another through time, thus adding a meta-level coupling, defining not simply ‘active-walker’ models but ‘active-landscape’ models. In this sense, we consider this to be an extension to the state-of-the-art in agent-based modelling.

We will first introduce a model with fixed origins for walkers and fixed destinations for resources, and this will enable us to predict the paths that walkers take between them. We will then show how the capacity of the network channels which emerge can be used to define the potential for locating new walkers. In this way, we link our network model to a location model, thus tying together walkers and landscapes in several different ways, each reflecting various positive feedbacks and diffusive effects that drive the system’s evolution. We will index each origin cell for a walker as I and each destination cell for a resource as J where $w_I^k(0)$ and $R_J(0)$ are the initial distributions of these quantities. In fact in our first model, these will not change in time as our focus is entirely on generating the networks that connect these two distributions.

The way the model works is by letting walkers move randomly through space, starting from their origins in search of resources at the given destinations. When a walker discovers a resource, essentially it tells other walkers of its discovery. But most of these will not be in the vicinity if the walk is truly random and thus the walker has to have some means of communicating this information. The walker does this by returning to its origin with the resource – back home to consume the food, if you like – and in making this trip, it lays down a path that other walkers can observe. This path will mark the straight-line distance from the origin to the destination, subject to any noise that interferes with the process. If this process is operated continually, then more and more walkers will discover resources, more and more walkers who have not discovered resources will detect the paths that lead to these resources, and ultimately everybody will be travelling on a route that takes them directly from their origin to a resource destination.

This is quite similar to Helbing, Schweitzer, Keltsch, and Molnar's (1997) model of trail formation and it figures widely in the way insect populations such as ants forage for food (Camazine, Deneubourg, Franks, Sneyd, Theraulaz, and Bonabeau, 2001). The model is sometimes called a swarm algorithm because when all movement is random, this is akin to a swarm moving out from some source. It is used to predict shortest paths in the Notting Hill model where such paths within the street network were unknown (Batty, Desyllas, and Duxbury, 2003a) although we will use it here to predict straight-line distances in our featureless plain. In essence if there are enough walkers swarming out from known origins $\{w_i^k(0)\}$ to known destinations $\{R_j(0)\}$, then once such a destination is discovered, the agent in question heads back directly from J to I impressing a track $s_{ij}(t+1)$ on every i, j cell pair which defines this line. This is added to the existing track, if there is one, as $S_{ij}(t+1) = S_{ij}(t) + s_{ij}(t+1)$ and in this way, the track gains in capacity. Walkers who are still in search of resources then react to the gradient formed by this track $\nabla S_{ij}(t)$ following the route from i to j which is the maximum of this gradient. Ultimately as tracks or network channels emerge, this is a reflection of the relative nearness of the origins to the destinations and the numbers of walkers involved.

The formal mechanism is little different from what has been stated already but for any walker in search of a resource

$$w_i^k(t) \rightarrow w_j^k(t+1) \text{ where } j \leftarrow \max_j \{ \nabla S_{ij}(t) + \varepsilon_j^k(t) \} , \quad (21)$$

where we have added the usual term for noise. This procedure works in a trackless landscape where movement is entirely dictated by random noise. Eventually all the walkers discover all the resources and the network landscape begins to stabilise in its morphology. When a track is formed as walkers who have discovered resources head back to base, it is usual to simply set $s_{ij}(t)$ to a constant which reflects a simple addition to the capacity due to the actions of one agent. A more detailed elaboration of the swarming that occurs with and without geometric constraints is presented in Batty, Desyllas and Duxbury (2003b).

An application of this network formation is illustrated in Figure 6 where we have simulated the tracks formed between 10 fixed walker and 13 resource locations where 1000 walkers have been randomly assigned to the 10 origins. In Figure 6(a) we show the origins and destinations, and then in 6(b) to (d), the distribution of the agents in the 201 x 201 cellular space, all paths taken in the landscape, and the tracks formed which is a subset of all paths taken: these patterns are illustrated for $t = 50$, $t = 500$, and $t = 5000$. The convergence from random walks to nodal structure in the landscape is impressive. There is more information, however, contained in this simulation. We have assumed that the numbers of agents visiting each resource location is unknown even though we may know the amount of resources there. It is however possible to compute the numbers of walkers visiting these locations in a cumulative manner which would give some indication of their size as

$$\bar{R}_j(T) \sim \sum_{\tau=1}^T \sum_k w_j^k(\tau) \quad . \quad (22)$$

As the network stabilises, so will the numbers attracted to each resource destination and to express these in terms of the total number of walkers, it is a simple matter to

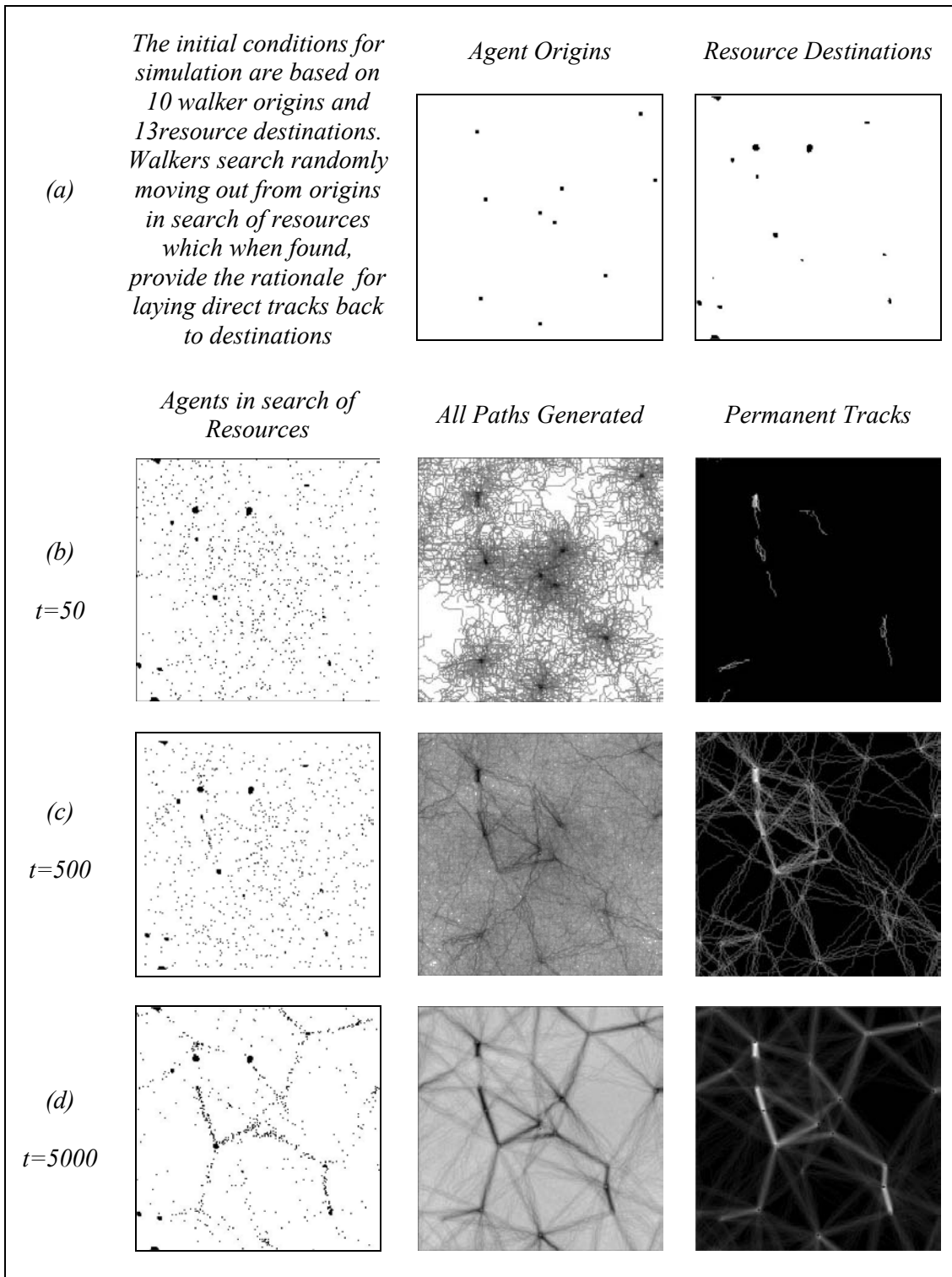


Figure 6: Network Formation Between Walker Origins and Resource Destinations

scale these totals as

$$\hat{R}_j(T) = \varphi \bar{R}_j(T) = \bar{R}_j(T) \frac{\sum_k w_i^k(0)}{\sum_j \bar{R}_j(T)} \quad . \quad (23)$$

$\hat{R}_j(T)$ can in fact be regarded as a measure of potential – network potential – of the resource node which can then be used to condition an extension of the model to incorporate growth in different locations.

Imagine we now wish to grow the number of agents from the initial base. One way of locating them would be to form a measure of potential and then to seek locations for new walkers where this potential is maximised. We form a generic potential using a reaction-diffusion equation

$$P_i(t+1) \sim P_i(t) + \omega \nabla^2 P_i(t) + \hat{w}_i(t+1) + \varepsilon_i(t) \quad . \quad (24)$$

ω is a weight on the diffusion term and $\hat{w}_i(t+1)$ is the location of one new agent in each time period, reflecting uniform growth through time whose location is chosen so that $\hat{w}_i(t+1)$ is determined from the cell given by $\max_j P_j(t)$. If we start with one walker, then what happens is that the first walker is located randomly as the potential surface is uniform. Reaction and diffusion ensures that this initial location survives and a path dependence then sets in which can only be broken if the noise in the system $\varepsilon_i(t)$ is large enough. In such applications, it is likely that the initial cluster will be reinforced. However if we make the connection between generic potential and the potential interaction at the node J as $\hat{R}_j(T) = P_j(t)$, we set in motion a process in which population becomes a function of resource potential which in turn is a function of the way populations discover resources through the emergence of their networks. To make the structure more elaborate, we might introduce a second potential equation in which new resources are located as a function of population, developing a structure in which agents and landscapes interact in all possible ways.

In fact, we will illustrate the simpler structure in which resource potential is used for generic potential in equation (24) rather than the more elaborate model just sketched. We will begin with 20 fixed resource locations and 100 walkers randomly located. As walkers begin their random walk using equation (21), new walkers are introduced one in each time period using equation (24). We show the tracks in space, the potential, and the location of the populations at an early stage ($t = 100$) in Figures 7(a) to (c), and then at a much later stage ($t = 2000$) in Figures 7(d) to (f). Although there is considerable persistence in the spatial structures generated over quite long time periods, over thousands of time periods, clusters of activity can change quite radically. Much depends on the level of noise introduced into the simulation but the patterns produced do mirror real systems of cities in terms of their social physics (Batty, 2001). Similar although more simplified models have been developed by Manrubia and Zanette (1998), and similar arguments are presented by Schweitzer (2003).

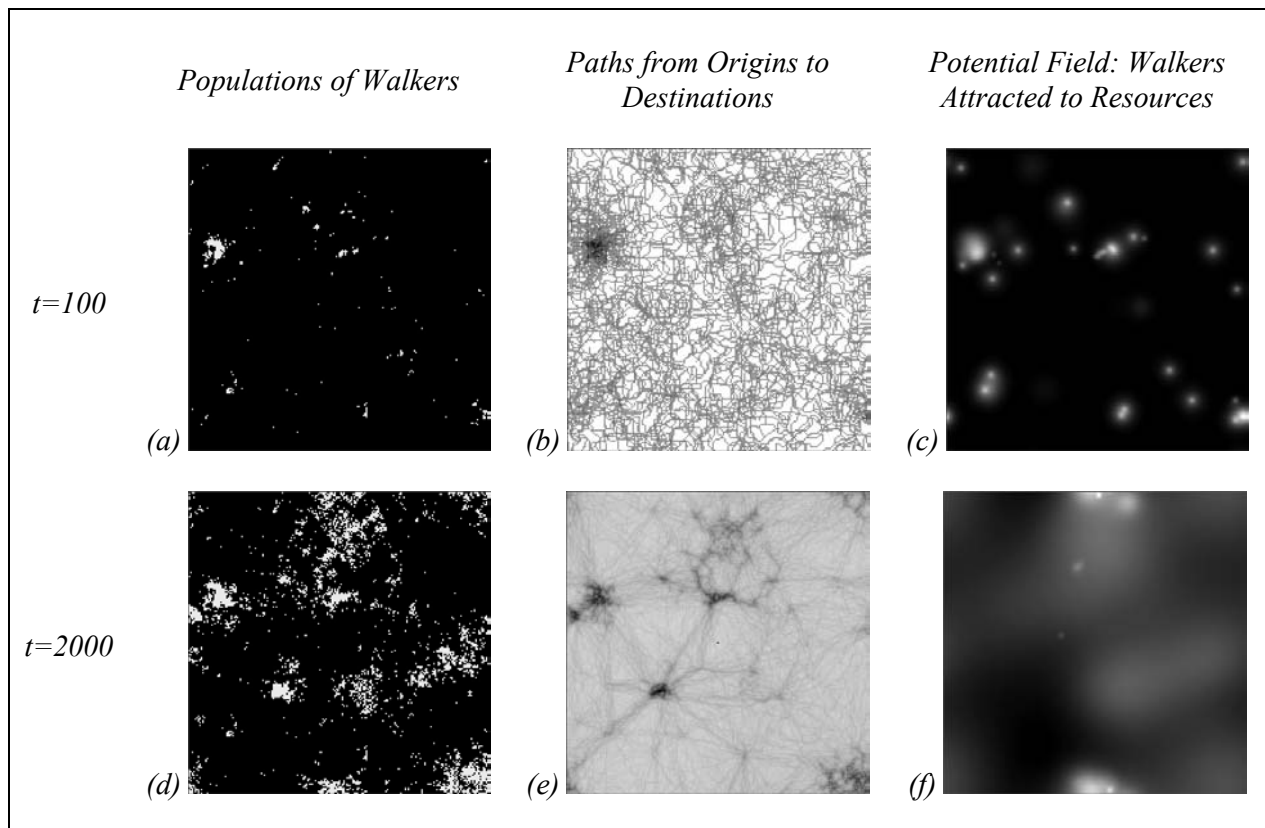


Figure 7: The Generation of a Coupled Active Urban Landscape

Cities at the Meso Scale: Metropolitan Dynamics and Urban Sprawl

Between the very small and the very large scale lie cities. New urban models at this meso scale have developed very rapidly but these have been largely fashioned around aggregating agents into cells, building on the traditional idea that land use and related urban activities take place in zones. These developments have been driven by GIS, treating space as a raster or pixel grid inspired by remotely sensed digital data. Moreover, urban growth, particularly in its current manifestation as sprawl, has spurred on these developments with a major focus on land consumption and conversion.

CA methods provide the workhorse for these new models although very few of these are built around strict CA and thus they are better thought of as cell-space (CS) automata. Generally cells are small enough to be associated with one and only one state, a land use, with several such states being represented, and the model dynamics being based on transitions between these land uses. Typically, a land use k in cell i at time t , $P_i^k(t)$, has the potential to spawn a new land use ℓ in cell j at time $t+1$, $P_j^\ell(t+1)$, where the location depends on how the neighbourhood is configured. In short, these CA/CS models enable land uses to grow or decline as a function of space, time and type but not as an explicit function of spatial interaction which is their Achilles' heel.

The model we will illustrate here was first developed by Xie (1994) under the acronym **DUEM** (Dynamic Urban Evolutionary Model) and a detailed presentation of the form to be presented is given by Batty, Xie and Sun (1999). In essence, the model integrates a life cycle for each land use with the possibility that any land use in its early life will spawn new land uses in the manner just suggested. Uses are classified into initiating, mature, or declining (and then vacant – the null state) with a regular transition between each controlled by parameters specific to the average time each land use is in each cycle. Initiating land uses have the potential to grow new land uses of any type within a field which is wider than the neighbourhood within which they are located. Constraints posed by distance and direction first determine the probability of new land uses but these are subject to a series of land use type and density

constraints within the narrower neighbourhood. Overall regional constraints dictate whether or not a cell can be occupied for a particular land use.

The rules are too tricky and extensive to formalise in the simple manner illustrated previously in equations (9) to (11) but we can sketch the way they operate. In each time period, the probability of converting a land use k in cell i to ℓ in j is computed as $p_{ij}^{k\ell}(t+1) = F\{P_i^k(t), P_j^\ell(t+1)\}$ where the functional form is based on a series of ‘*if* → *then* rules’ of the conventional CA type. For any land use k in i , the probability is first determined by the distance from i to j in the field around i . Any distortions in direction are added, and then this probability is checked for legitimacy against a series of density constraints on the occurrence of different land uses in the neighbourhood around i . If the probability of new growth survives, this is tested against the presence of some street pattern in the neighbourhood for it is essential that any new use be ‘near’ some transport. Streets are land uses too and these are grown in a similar but slightly more restrictive manner. In this way, land development and transportation are coordinated in physical terms. Once this probability matrix is in its final form, a new land use is chosen from the probability $p_{ij}^{k\ell}(t+1)$ using some random mechanism, or if it is decided that a new land use must always be grown, the maximum probability determines what this is. There is no symmetric process to simulate decline for this is determined by the aging of land uses.

In ***DUEM***, there are currently five land uses – residential (population), manufacturing and primary uses, services and commerce, streets, and vacant land. In the current software, further uses can be defined although these five are the default, with plausible distance, aging, density, and related parameters defining their respective transitions. Another problem of CA/CS models is the inability to control the extent to which local growth meets global targets although this might also be regarded as a predictive ability. In fact at each time cycle, in operational versions which must meet some target, this is achieved by a crude scaling. The version we illustrate here enables both real and hypothetical examples to be developed and we will begin with the hypothetical. We have randomly planted a series of land uses of these five types in a 350 x 250 pixel space and using the default rules, grown the system through time to the point where the entire space is occupied. Some land use seeds never take off

because they do not satisfy constraints in their local neighbourhoods and fields but most do. As the space fills up, the total quantities of land use grow logistically to upper limits but as aging takes place, land uses become vacant and the trajectories of growth begin to oscillate. This is indicative of a simple capacitated system but the interest here is on the long term balance of land uses. As the rules are specified locally, we have no idea in advance as to how these will combine together to produce realistic structures and thus the experimentation which we show in Figure 8 is essential for tuning the transition probabilities governing the system's evolution.

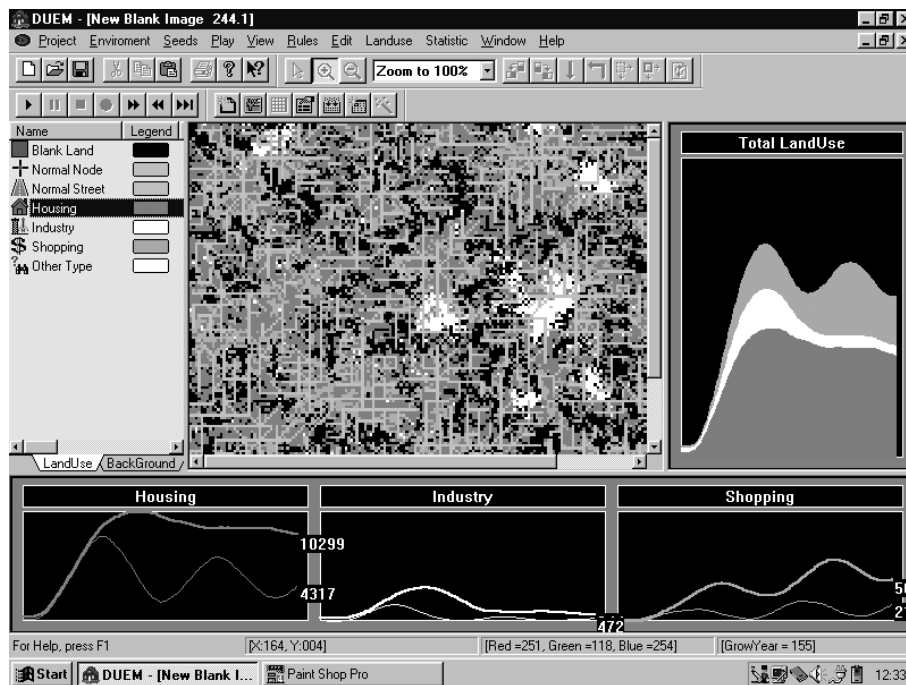


Figure 8: Logistic Growth of a Capacitated Urban System

Our second example simulates growth in a real system, in metro Detroit which although declining in total population like many cities in the US rustbelt, is growing rapidly on its edges. Detroit and towns in its suburbs like Ann Arbor which we have also modelled, represent the equivalent of jobless growth in the economy with rapid change taking place in terms of land consumption and abandonment but with total activity declining or, at least, not growing much. The software that we have allows us to import land use patterns from desktop GIS; we begin with the pattern of development in Detroit in 2000 and simulate its growth and spread forward to a long term steady state some 200 years hence when the region reaches capacity. We

illustrate a typical output in Figure 9 where the size of the space in which the simulation takes place is 514 x 613. In fact we can handle up to 9 million pixels/cells in the current software and thus we have no difficulties in ensuring that the cell size is appropriate to single land uses.

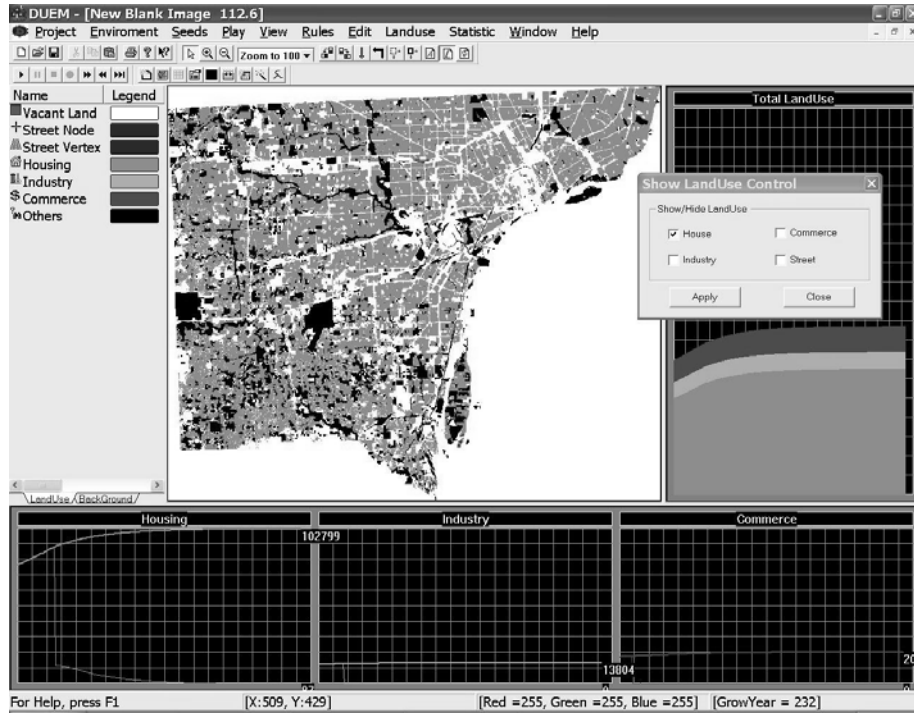


Figure 9: Residential Development in the Very Long Term in the Detroit Metro Area

There are now upwards of 50 applications of CA/CS models of urban growth with sustained effort taking place in some half dozen places, namely at RIKS (White and Engelen, 2000), Santa Barbara (Clarke, Hoppen and Gaydos, 1997), Southampton-Cardiff (Wu and Martin, 2002), Hong Kong (Li and Yeh, 2000), Tel Aviv (Portugali, 2000), and Brisbane (Ward, Murray, and Phinn, 2000) besides our own. Most but not all these models are inspired by applications involving urban sprawl but some such as the Tel-Aviv models are geared to segregation and polarisation while there have been various attempts to endow such structures with stronger urban economic content. These applications are all subject to the limitations that plague this area generally, the lack of explicit transportation, poor command over control totals, and rather stylised representations of land uses in cells. Nevertheless progress is being made and there are many new developments in the pipeline.

Conclusions: Next Steps

This long review warrants a short conclusion. Currently the extent to which CA/CS and agent-based models of urban systems can be fully implemented for policy applications is quite limited. As the scale gets finer and the agents and their cells become more like real objects, their operationality increases to the point where substantive policy applications are possible. Such is the case with the pedestrian models that we began with where the definitions of objects are quite unambiguous in comparison to larger scales. At the largest scale, these approaches stretch the field in terms of theory and such is their role. But the biggest problem with all these models is their lack of parsimony. The richness of the data required makes calibration and estimation difficult and predictive accuracy hard to assess in terms of past simulations. Thus many of these models are demonstrated in pedagogic fashion to show emergence and path dependence in terms of their processes. Use in forecasting is largely restricted to very long term futures where their role in scenario testing is clear.

In scientific terms, the biggest hurdle to surmount involves the role of transportation but in the agent-based modelling strategies we illustrated for the very small and very large scales, explicit interaction can be simulated. Much more work needs to be done in this area while more routine extensions need to be implemented with respect to cellular representation. It is interesting that apart from a few notable exceptions (Sanders, Pumain, Mathian, Guerin-Pace and Bura, 1997), hardly any agent-based model have yet been developed at the meso scale although there is active development for more environmentally-based land cover systems (Parker, Manson, Janssen, Hoffman, and Deadman, 2003). Applications will doubtless increase but at the end of this review, we are still left with the perennial question which dominates all discussion of science in public affairs: to what extent can formal models be built which will provide robust enough forecasts for real policy analysis? These new approaches only provide a part of this answer. Although promising in that new forms of representation clearly get to grips somewhat more effectively with the way contemporary problems are articulated, this perspective raises a new set of questions which limit their applicability in rather different ways from traditional urban models. This is a recurrent feature of this field which suggests that not one but many different approaches will always be required.

References

- Albin, P. (1975) *The Analysis of Complex Socio-Economic Systems*, Lexington Books, Lexington, MA.
- Alonso, W. (1964) *Location and Land Use: Toward a General Theory of Land Rent*, Harvard University Press, Cambridge, MA.
- Batty, M. (1971) Modelling Cities as Dynamic Systems, *Nature*, 231, 425-428.
- Batty, M. (1976) *Urban Modelling: Algorithms, Calibration, Predictions*, Cambridge University Press, Cambridge, UK.
- Batty, M. (1984) *Pseudo Dynamic Urban Models*, PhD University of Wales, NCB85-19527, University Microfilms International, Ann Arbor, Michigan 48106.
- Batty, M. (2001) Polynucleated Urban Landscapes, *Urban Studies*, 38, 635-655.
- Batty, M. and Jiang, B. (2000) Multi-Agent Simulation: Computational Space-Time Dynamics in GIS, in P. Atkinson and D. Martin (Editors) *Innovations in GIS VII: GIS and Geocomputation*, Taylor and Francis, London, 55-71.
- Batty, M., Desyllas, J., and Duxbury, E. (2003a) Safety in Numbers? Modelling Crowds and Designing Control for the Notting Hill Carnival, *Urban Studies*, 40, 1573-1590.
- Batty, M., Desyllas, J., and Duxbury, E. (2003b) The Discrete Dynamics of Small-Scale Spatial Events: Agent-Based Models of Mobility in Carnivals and Street Parades, *International Journal of Geographic Information Science*, forthcoming.
- Batty, M., Xie, Y., and Sun, Z. (1999) Modeling Urban Dynamics Through GIS-Based Cellular Automata, *Computers, Environments and Urban Systems*, 23, 205-233.
- Berry, B. J. L. (1964) Cities as Systems within Systems of Cities, *Papers and Proceedings of the Regional Science Association*, 13, 147- 163.
- Borgers, A., and Timmermans, H. (1986) A Model of Pedestrian Route Choice and Demand for Retail Facilities within Inner-City Shopping Areas, *Geographical Analysis*, 18, 115-128.
- Camazine, S., Deneubourg, J-L., Franks, N. R., Sneyd, J., Theraulaz, G. and Bonabeau, E. (2001) *Self-Organization in Biological Systems*, Princeton University Press, Princeton, NJ.
- Clarke, K. C., Hoppen, S., and Gaydos, L. (1997) A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area, *Environment and Planning B*, 24, 247-261.
- Couclelis, H. (1985) Cellular Worlds: A Framework for Modeling Micro-Macro Dynamics, *Environment and Planning A*, 17, 585-596.

- Epstein, J. M., and Axtell, R. (1996) *Growing Artificial Societies: Social Science from the Bottom Up*, Brookings/MIT Press, Cambridge, MA.
- Forrester, J. W. (1969) *Urban Dynamics*, MIT Press, Cambridge, MA.
- Gardner, M. (1970) The Fantastic Combinations of John Conway's New Solitary Game of 'Life', *Scientific American*, 223, 120-23.
- Haklay, M. Thurstain-Goodwin, M. O'Sullivan, D. and Schelhorn, T. (2001) "So Go Downtown": Simulating Pedestrian Movement in Town Centres, *Environment and Planning B*, 28, 343-359.
- Hansen, W. G. (1959) How Accessibility Shapes Land Use, *Journal of the American Institute of Planners*, 25, 73-76.
- Helbing, D., Farkas, I., and Vicsek, T. (2000) Simulating Dynamical Features of Escape Panic, *Nature*, 407, 487-490.
- Helbing, D., Molnar, P. Farkas, I. J., and Bolay, K. (2001) Self-Organizing Pedestrian Movement, *Environment and Planning B*, 28, 361-383.
- Helbing, D., Schweitzer, F., Keltsch, J., and Molnar, P., 1997, Active Walker Model for the Formation of Animal and Trail Systems. *Physical Review E*, 56, 2527-2539.
- Kayser, D. R., Aberle, L. K., Pochy, R. D., and Lam. L. (1992) Active Walker Models: Tracks and Landscapes, *Physica A*, 191, 17-24.
- Li, X., and Yeh, A. G-O. (2000) Modelling Sustainable Urban Development by the Integration of Constrained Cellular Automata and GIS, *International Journal of Geographic Information Science*, 14, 131-152.
- Lowry, I. S. (1964) *Model of Metropolis*, Memorandum RM-4035-RC, Rand Corporation, Santa Monica, CA.
- Manrubia, S. C., and Zanette, D. H. (1998) Intermittency Model for Urban Development, *Physical Review E*, 58, 295-302.
- Niemeyer, L., Pietronero, L., and Wiesmann, H. J. (1984) Fractal Dimension of Dielectric Breakdown, *Physical Review Letters*, 52, 1033-1036.
- Parker, D. C., Manson, S. M., Janssen, M. A., Hoffman, M. J., and Deadman P. (2003) Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review, *Annals of the Association of American Geographers*, 93, 314-337.
- Portugali, J. (2000) *Self-Organization and the City*, Springer-Verlag, Berlin.
- Poundstone, W. (1985) *The Recursive Universe*, William Morrow, New York.

- Sanders, L., Pumain, D., Mathian, H., Guerin-Pace, F., and Bura, S. (1997) SIMPOP: A Multiagent System for the Study of Urbanism, *Environment and Planning B*, 24, 287-305.
- Schelling, T. C. (1978) *Micromotives and Macrobehavior*, W. W. Norton and Company, New York.
- Schreckenberg, M. and Sharma, S. D. (Editors) (2002) *Pedestrian and Evacuation Dynamics*, Springer-Verlag, Berlin.
- Schweitzer, F. (2003) *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*, Springer-Verlag, Berlin, Germany.
- Stanley, H. E. and Ostrowsky, N. (Editors)(1986) *On Growth and Form*, Martinus Nijhoff, Dordrecht, Holland.
- Stewart, J. Q. and Warntz, W. (1958) Physics of Population Distribution, *Journal of Regional Science*, 1, 99-123.
- Toffoli, T. and Margolis, N. (1987) *Cellular Automata Machines: A New Environment for Modeling*, MIT Press, Cambridge, MA.
- UAS (1996) *Tate Gallery, Millbank: A Study of the Existing Layout and New Masterplan Proposal*, Unit for Architectural Studies, Bartlett School, University College, London.
- Vicsek, T. (Editor) (2002) *Fluctuations and Scaling in Biology*, Oxford University Press, Oxford, UK.
- Ward, D. P., Murray, A. T., and Phinn, S. R. (2000) A Stochastically Constrained Cellular Model of Urban Growth, *Computers, Environment and Urban Systems*, 24, 539-558.
- Wegener, M. (1994) Operational Urban Models: State of the Art, *Journal of the American Planning Association*, 60, 17-30.
- White, R. W. and Engelen, G. (2000) High-Resolution Integrated Modelling of the Spatial Dynamics of Urban and Regional Systems, *Computers, Environments and Urban Systems*, 24, 383-400.
- Wilson, A. G. (1970) *Entropy in Urban and Regional Modelling*, Pion Press, London.
- Wilson, A. G. (2000) *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis*, Pearson Education, Harlow, UK.
- Witten, T. A. and Sander, L. M. (1981) Diffusion-Limited Aggregation: A Kinetic Critical Phenomena, *Physical Review Letters*, 47, 1400-1403.
- Wolfram, S. (1994) *Cellular Automata and Complexity: Collected Papers*, Addison Wesley, Reading, MA.

Wu, F., and Martin, D. (2002) Urban Expansion Simulation of Southeast England using Population Surface Modelling and Cellular Automata, *Environment and Planning A*, 34, 1855-1876.

Xie, Y. (1994) Analytical Models and Algorithms for Cellular Urban Dynamics, unpublished PhD dissertation, State University of New York at Buffalo, Buffalo, NY.