Environment and Planning B: Planning and Design 2007, volume 34, pages 687 – 707

DOI:10.1068/b32150

Building a city *in vitro*: the experiment and the simulation model

Erez Hatna, Itzhak Benenson
Environmental Simulation Laboratory, Porter School of Environmental Studies and Department of Geography and Human Environment, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel; e-mail: erezh51@post.tau.ac.il, bennya@post.tau.ac.il
Received 11 November 2005; in revised form 21 April 2006

Abstract. All current urban models accept the 'first-order recursion' view, namely, that the state of an urban system at time t is sufficient for predicting its state at $t+1$. This assumption is not at all evident in the case of urban development, where the behavior of developers and planners is defined by the complex interaction between long-term and short-term plan guidelines, local spatial and temporal conditions, and individual entrepreneurial activity and cognition. In this paper we validate the first-order recursion approach in an artificial game environment: thirty geography students were asked to construct a 'city' on the floor of a large room, with each student using the same set of fifty-two building mock-ups. Based on the analysis of game outcomes, the first-order recursive set of behavioral rules shared by all the participants is estimated and further employed for computer generation of artificial cities. Comparison between the human-built and model patterns reveals that the constructed set of rules is sufficient for representing the dynamics of the majority of experimental patterns; however, the behavior of some participants differs and we analyze these differences. We consider this experiment as a preliminary yet important step towards the adequate modeling of decision-making behavior among real developers and planners.

1 Introduction

During the last decade we have witnessed a boom in urban modeling. Different from the regional models of 1980s, the new wave of high-resolution models focuses on behavior and transformations of urban objects (Benenson and Torrens, 2004). Numerous models of land-use dynamics, car and pedestrian traffic, residential migration, and service relocation, among others, have already been published. When the modelers aim at an explicit implementation of the behavior of individual urban agents—pedestrians, car drivers, householders, urban developers, or entrepreneurs—they apply multiagent systems (MAS); when they limit themselves to the infrastructure outcomes of human actions, they turn to cellular automata (CA) models.

Whether implementing an MAS or CA perspective, the modeler must formulate the rules of the objects' 'behavior' in space – time: the rules of agent behavior in MAS, or cell transition rules in CA. Analysis of existing models reveals a definite fashion in the formulation of these rules. If we limit ourselves to the dominant discrete-time and probabilistic view of urban systems, we can readily observe that the model rules describe the state and location of the objects that have been changed, added, or removed during the recent time step as being exclusively dependent on the state of the system at the previous time step. Possible dependence on the longer history is ignored, usually for a very practical reason—the experimental data necessary for testing this assumption are barely available. The evasion, however, is not at all evident—urban development is the outcome of efforts by developers and planners, the latter defined by the complex interaction between the long-term and short-term plan guidelines and the associated implementation policy, local spatial and temporal conditions, as well as individual entrepreneurial activity and cognition. In this paper we take the first step towards an experimental validation of the first-order recursion approach as an urban modeling perspective. We limit our study to an artificial environment and perform a series of thirty experiments in which a game participant constructs an artificial 'city' by locating 'buildings' on the floor of a large room. We then determine whether the first-order recursion is sufficient for representing the dynamics of the urban patterns that emerged in these thirty games.

2 Views of urban dynamics and human developers' behavior in space

2.1 Urban CA and MAS models

CA models of land-use dynamics consider urban areas as consisting of many parcels, each of which is in one of several discrete, easily recognized land-use states $\{s_1, s_2, s_3, ...\}$—dwelling, industry, open land, and so forth. The description of urban dynamics in CA land-use models is based on the Markov field model and its extensions (Benenson and Torrens, 2004). The Markov field model is a typical first-order recursion; it assumes that the transition probability p_{ij} that the state of a cell at $t+1$ will be s_j depends on the cell's state s_i and the states of the cell's neighbors at t only. Investigation of CA land use focuses on the equilibrium distributions of land uses and the convergence toward that equilibrium (Boerner et al, 1996; Gobin et al, 2002; Jahan, 1986).

Markov field models consider p_{ij} to be independent of cell location and time. CA land-use models go far beyond this limitation and account for p_{ij} as being dependent on location, say the distance to the nearest cell of a specific type (eg 'road'; see Clarke et al, 1997), or a (externally defined) demand for cells in a given state (White and Engelen, 1997). Other factors, such as population density (Yeh and Li, 2002) and major landscape characteristics, including elevation, slope, and suitability of the soil for agriculture (Brown et al, 2002; Hathout, 2002; LaGro and DeGloria, 1992; Lopez et al, 2001) are often included.

In the majority of CA models, the validity of the first-order recursion is taken for granted. However, this postulate has been indirectly confirmed by the model's high explanatory ability—in some cases up to 95% of future land uses could be predicted in this way (Gobin et al, 2002). Other cases exist, however, in which knowledge of previous land use is not helpful for predicting future changes (Bell, 1974), or in which the transition probabilities p_{ij} change in time (Weng, 2002).

Successful urban MAS applications of which we are aware include those of Benenson et al (2002), who captured residential dynamics in an ethnically mixed area of Tel Aviv, and Pentland and Liu (1999), who predicted drivers' actions in a controlled virtual reality environment with 95% accuracy. Both follow the first-order recursion approach.

2.2 Urban spatial games

One promising approach to studying the dynamics of an urban system as an outcome of developers' efforts is to imitate development by means of a dynamic game—either virtually implemented with a computer (Cecchini and Rizzi, 2001; Semboloni et al, 2003) or actually performed with physical objects. Until now, all the game-based studies of human spatial behavior that we are aware of have considered economic behavior alone. Most popular is the experimental reproduction of the famous Hotelling (1929) result, which states that the only stable outcome of competition between two firms for customers distributed over a 1D interval is the location of both sellers in the center of the interval. The games that imitate the spatial location and relocation of buyers and sellers have been investigated in several papers, yet these have not always confirmed the theoretical results (Camacho-Cuena et al, 2005; Collins and Sherstyuk, 2000; Huck et al, 2002).

Mock-up modeling is one basic element of architectural work, where setups of mock-up buildings and groups of buildings in urban environments are standard tools of project presentation (Radford, 2000). However, the shift from the use of static mock-ups to dynamic experiments in which participants *construct* a 'city' was proposed by Portugali only recently (1996). We are aware of two implementations of games of this kind (Mayer et al, 2004; Portugali, 2000), both of which involve *many players*, with each adding infrastructure units to the emerging pattern. The respective players can employ different types of construction behavior, thus introducing additional variance into the experimental results. We view our study as innovative, given that participant behavior is analyzed quantitatively; to establish the methodology for detailed analysis of human developers' behavior, we focus on a simpler situation, in which *only one person makes all the location decisions.*

3 Description of the experiment

Thirty undergraduate geography students were asked to 'construct a city' using fifty-two mock-ups to be situated on the floor of a large space (a hall). The game's area took the form of a trapezoid(1) of 28.2 m^2 in size. The set of mock-ups was developed by Portugali (1996) and has been used for various experiments (Portugali, 2002). The mock-ups represented real buildings at a 1:100 scale; they also represented different urban functions—small-height and medium-height dwellings, commercial sites, office buildings, and so on. The average size of a mock-up was 20 cm $\times$ 20 cm, although it could range from 10 cm $\times$ 10 cm to 40 cm $\times$ 40 cm (figure 1).

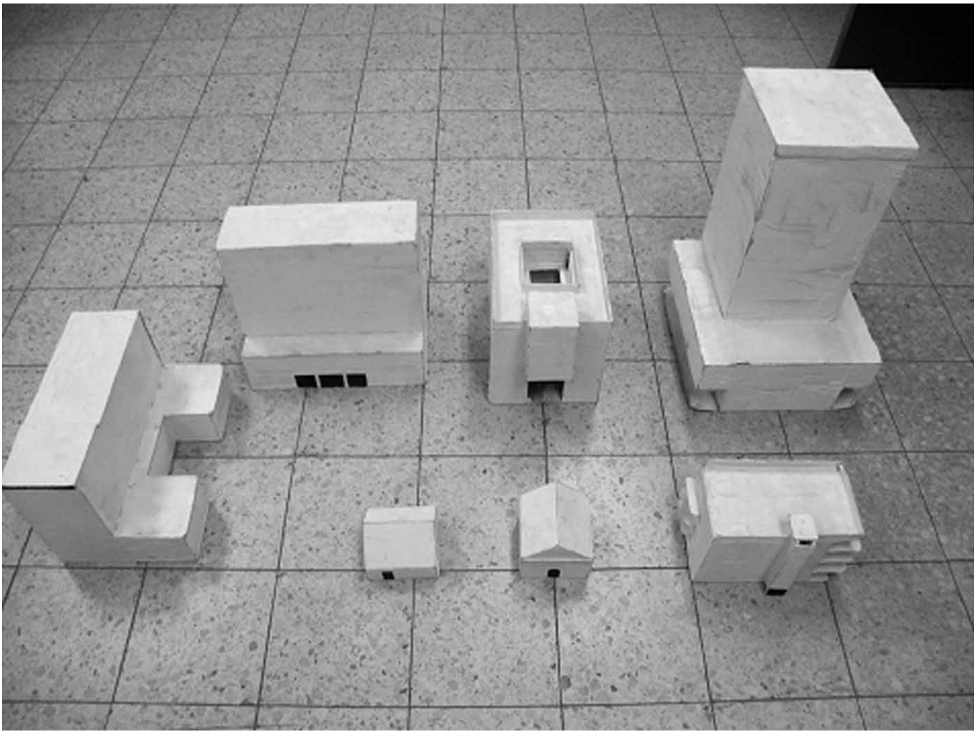
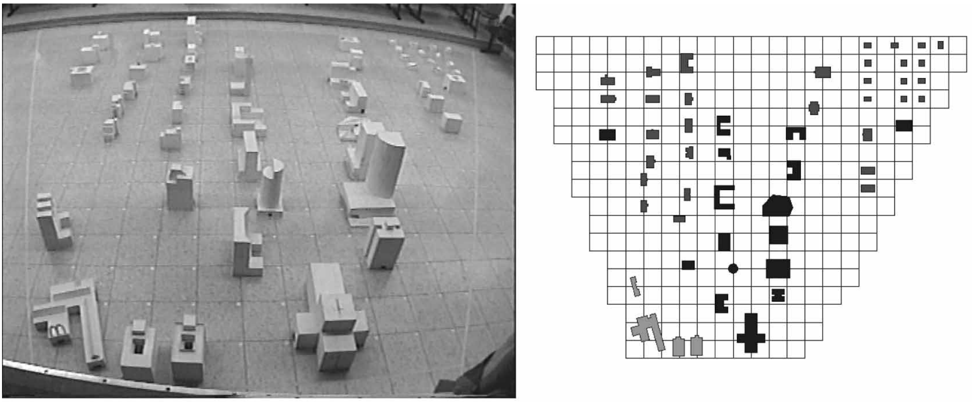


Figure 1. Seven of the fifty-two mock-ups used in the experiments (the floor tiles are of 20 cm $\times$ 20 cm size).

(1) The experiments were captured by video camera in fixed position, with a trapezoidal view.

Each participant constructed the 'city' once, by positioning mock-ups one at a time, in fifty-two *steps* of time *t*. During a time step, the participant chose a building from the remaining stock, declared its urban function, and then positioned it. The participant was not permitted to relocate or return a mock-up to the original stock once it was positioned. The experiment was concluded when all fifty-two mock-ups had been placed on the floor. While participating in the experiment, the participant could declare one of seven urban functions for a building—dwelling, commerce, entertainment, culture,



(a) (b)

Figure 2. (a) A snapshot of a game outcome; (b) its GIS presentation by means of mock-up foundation polygons.

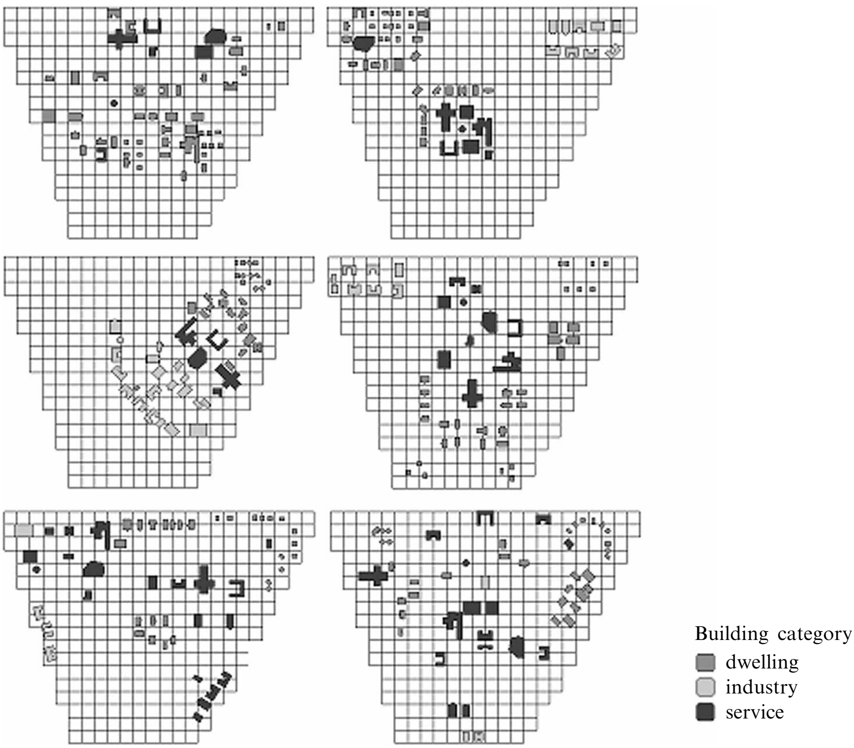


Figure 3. Six typical cities constructed in the experiments.

religion, industry, or office. After analyzing the experiments, we combined five infrequently employed building functions—commerce, entertainment, culture, religion, and office—into one, thereafter called service.

At each step, the mock-up's urban function, identifier, position, and orientation were recorded. Buildings were further represented as features of a GIS layer (figure 2), thus enabling spatial analysis of the results.

Figure 3 presents the final patterns of six cities, randomly selected from the thirty game outcomes produced.

4 A constant set of shared first-order recursive rules as the null hypothesis

The full history of the city game built by participant b from the beginning of the game until time t (inclusively) is given by a $(t+1)$-dimensional vector of the *city history* C_t^{b}, representing all 'construction' actions: $C_t^{\mathrm{b}} = X_0^{\mathrm{b}} \rightarrow X_1^{\mathrm{b}} \rightarrow, ..., \rightarrow X_t^{\mathrm{b}}$, where arrows denote ordered sequences. According to the experimental framework, each X_t^{b} is a pair $(f_t^{\mathrm{b}}, l_t^{\mathrm{b}})$ where f_t^{b} is the chosen building's function and l_t^{b} its location. It is important to distinguish between system history, given by C_t^{b} and *city pattern*, $P_t^{\mathrm{b}} = \{X_0^{\mathrm{b}}, X_1^{\mathrm{b}}, ..., X_t^{\mathrm{b}}\}$, with curled brackets denoting unordered sequences. Let us denote the rules that b employs at t as u_t^{b}; the game participant establishes X_t^{b} by applying u_{t-1}^{b}, given C_{t-1}^{b}:

$$X_t^{\mathrm{b}} = u_{t-1}^{\mathrm{b}}(C_{t-1}^{\mathrm{b}}), \qquad t = 1, 2, ..., 51. \tag{1}$$

Formula (1) expresses the general view of the behavior of the game participant. For real-world cities, the meaning of (1) is mainly conceptual; as discussed above, the typical modeling approach is to simplify (1), assuming that only the immediate past, given by P_{t-1}^{b}, matters. In terms of the game, this can be put as follows:

- The *game participant b* determines f_t^{b} and l_t^{b} on the basis of P_{t-1}^{b} only,
- The rules u_t^{b} that b applies do not change in time: $\forall t$, $u_t^{\mathrm{b}} = u^{\mathrm{b}}$,
- All the participants share the same set of rules: $\forall b$, $u^{\mathrm{b}} = u$.

Formally, *time-independent shared first-order recursion* can be expressed as follows:

$$X_t^{\mathrm{b}} = u(P_{t-1}^{\mathrm{b}}). \tag{2}$$

The range of views that fall between (1) and (2) is very broad. As a first step beyond (2) and towards (1), one can consider each participant as applying different sets of rules, with each set accounting for the previous pattern only: $X_t^{\mathrm{b}} = u^{\mathrm{b}}(P_{t-1}^{\mathrm{b}})$, or the shared set of rules which evolves with the evolution of the city: $X_t^{\mathrm{b}} = u_{t-1}(P_{t-1}^{\mathrm{b}})$. Furthermore, one can assume that u_t^{b} depends, among other things, on b's actions during several previous time steps, on some of the previous actions and so forth.

Accepting conceptual presentation (2), we have to specify the set of rules u; this can be done in numerous ways. We see our study as an innovation step, with formulation of such a set as one of our goals. The set we propose below is sufficient for describing the dynamics of the majority of patterns. Several of the patterns that deviate from this set are recognized and investigated.

5 Analysis of the experiment

As defined above, the behavioral actions u_{t-1}^{b}, $t = 1, 2, ..., 52$ represent two elementary acts performed by b during a time interval $(t-1, t)$—choice of the mock-up's function $f_t^{\mathrm{b}} \in$ {dwelling, service, industry} and choice of its position l_t^{b}. To begin the analysis, let us assume that P_{t-1}^{b} is sufficient for determining X_t^{b}, and u_{t-1}^{b} does not depend on b and t: $\forall b \forall t$, $u_t^{\mathrm{b}} = u$; if so, the experimental data can be considered as thirty repetitions of the same behavioral mode. To estimate u, let us proceed with analysis of the f_t^{b} and l_t^{b} sequences separately.

5.1 Choice of a building's urban function f_t^{b}

To verify the rules of building function choice, let us analyze $f_{t-1}^{\mathrm{b}} \to f_t^{\mathrm{b}}$ pairs of choices in thirty games (table 1):

The data in table 1 are strongly in favor of the dependence of f_t on f_{t-1} ($\chi^2 = 556.6$, $p < 0.001$, contingency coefficient $C = 0.52$). We thus continue the analysis, assuming that participants' choices of the next building's function depend on the function of the previously chosen building.

Table 1. Observed and expected, in cases of independent choice (in brackets), frequencies of $f_{t-1} \to f_t$ pairs.

f_{t-1}	f_t		
	dwelling	industry	service
dwelling	743 (570.8)	54 (115.1)	135 (246.1)
industry	52 (115.1)	93 (23.2)	43 (49.6)
service	142 (251.1)	42 (50.6)	226 (108.3)

To investigate the *sufficiency* of the first-order shared time-independent recursion regarding choice of building function, we tested whether the dependence of f_t on f_{t-1} can be considered as a first-order Markov chain. In the case of a Markov chain, the probability that f_t is chosen, given certain f_{t-s}, $s > 1$, is a power of s of the matrix $\mathbf{P} = \| p_{ij} \|$ of the first-order transition probabilities (table 2) (Isaacson and Madsen, 1976).

Table 2. Probability p_{ij} of choosing a building of function j, given building function i chosen at the previous time step (calculations are based on table 1 results).

$i(f_{t-1})$	$j(f_t)$		
	dwelling	industry	service
dwelling	0.797	0.058	0.145
industry	0.277	0.495	0.229
service	0.346	0.102	0.551

We thus compared the experimental frequencies of $f_{t-s} \to f_t$ transitions with the predictions according to the frequencies of the first-order $f_{t-1} \to f_t$ transitions (table 2). Table 3 presents the results of this comparison for $s = 1$, 2, and 3:

Table 3. Comparison of experimental and theoretical frequencies of $f_{t-s} \to f_t$ transitions for $s = 2$, 3, 4.

Time delay s	χ^2	Degrees of freedom	Significance	Contingency coefficient C
2	6.169	8	$p \sim 0.628$	0.064
3	23.743	8	$p \sim 0.003$	0.126
4	23.459	8	$p \sim 0.003$	0.127

Table 3 confirms the first-order Markov chain as a model of participants' choice regarding $f_{t-2} \to f_t$ transitions, although the result does not hold to the same degree for longer histories, that is, $f_{t-3} \to f_t$ and $f_{t-4} \to f_t$ transitions. A review of the χ^2 components demonstrates that industry $\to$ industry transitions are the main contributors to the discrepancy (table 4).

Table 4. Components of χ^2 when comparing experimental and theoretical frequencies of transition for $s = 2$, 3, and 4.

Transition		(observed – expected)^2/expected		
f_{t-s}	f_t	$s = 2$	$s = 3$	$s = 4$
dwelling	dwelling	0.890	3.711	2.354
industry	dwelling	0.494	3.516	1.864
service	dwelling	0.678	1.212	0.063
dwelling	industry	1.243	2.546	4.292
industry	industry	1.134	**6.429**	**12.445**
service	industry	0.314	0.330	0.826
dwelling	service	0.998	4.566	1.081
industry	service	0.038	0.202	0.472
service	service	0.382	1.232	0.062
Total χ^2		6.169	23.743	23.459

Looking closer at the industry$_{t-s}$ $\rightarrow$ industry$_t$ transitions, the observed and expected numbers are 60 versus 52.3, respectively, for $s = 2$; 50 versus 35.0 for $s = 3$; and 46 versus 27.5 for $s = 4$. That is, within the confines of the experiment, participants returned to industry mock-ups a few steps later at *more frequent* rates than might have been expected assuming that choice of a building's urban function is indeed part of a first-order Markov chain. We did not proceed with this analysis because, given the data, we could conclude that although the participants' choice of building function was not fully first-order Markov chain-related, it quite closely resembles it.

5.2 Choice of building location l_t^{b}

Just as we rejected the hypothesis that a building's function at t does not depend on the building's function selected at $t-1$, we can reject the hypothesis that l_t^{b} does not depend on P_{t-1}^{b}. To prove this formally, we applied the nearest-neighbor index r(1) (Clark and Evans, 1954) to the final game patterns, ignoring the building functions. Namely, we simulated the location patterns of fifty-two buildings provided that the next building was randomly located on the unoccupied area, and then compared the mean value of $\mathrm{r_e}(1)$ over 30 experimental patterns, with the $\mathrm{r_m}(1)$ calculated over 500 patterns generated by the model. The results were $\mathrm{r_e}(1) = 35.4$ cm ($\mathrm{SD_e} = 5.1$ cm), $\mathrm{r_m}(1) = 38.7$ cm ($\mathrm{SD_m} = 3.0$ cm); applying t-criteria, one obtains that the difference $\mathrm{r_m}(1) - \mathrm{r_e}(1) = 3.3$ cm is significant at $p < 0.005$. We can thus reject the hypothesis that l_t does not depend on the participant's previous choice of buildings location.

Let us propose the set of rules that describe *how* information from P_{t-1}^{b} is utilized when establishing position l_t^{b} of the building with a given function f_t^{b}.

5.2.1 *Which characteristics of P_{t-1}^{b} are taken into account?*

The assumption that P_{t-1}^{b} defines l_t^{b} demands specification—for example, one might assume that l_t^{b} is determined by the location l_{t-1}^{b} and the function f_{t-1}^{b} of the previously established building. We investigated this hypothesis and found that the rules based on unitary buildings did not engender satisfactory results. What did work was the dependence of l_t^{b} on the properties displayed by the *entire pattern* P_{t-1}^{b}. In the following discussion, we consider the following set of rules: the probability of choosing a certain location l for a building of function f_t^{b} is defined by (a) three distances between l and *the nearest buildings* of each of the three functions in P_{t-1}^{b}, and (b) the direction of the vector connecting the nearest building of function type f_t^{b} in P_{t-1}^{b} and l.

5.2.2 *Dependence of l_t^{b} on the distance to the nearest building in P_{t-1}^{b}*

Let us denote the locations of the dwelling, service, and industry buildings in P_{t-1}^{b} that are nearest to a certain location l, as $n\mathrm{dwelling}_{t-1,l}^{\mathrm{b}}$, $n\mathrm{service}_{t-1,l}^{\mathrm{b}}$, and $n\mathrm{industry}_{t-1,l}^{\mathrm{b}}$. The distributions of the distances between l and $n\mathrm{dwelling}_{t-1,l}^{\mathrm{b}}$, $n\mathrm{service}_{t-1,l}^{\mathrm{b}}$, and $n\mathrm{industry}_{t-1,l}^{\mathrm{b}}$ were estimated with experimental GIS maps [figure 2(b)] and are presented in figure 4.

The distributions of the distances in figure 4 reflect the participants' tendency to cluster: according to figures 4(a), 4(e), and 4(i), participants indeed prefer to locate a new building near a standing building of the same functional type, while staying far away from buildings of other types. The effect was strongest in the case of dwellings, where 80% of the buildings were located at a distance less than 60 cm from an existing dwelling [figure 4(a)]. The distances between industry and dwelling buildings [figures 4(b) and 4(d)] exhibited the greatest aversion between urban functions, while the

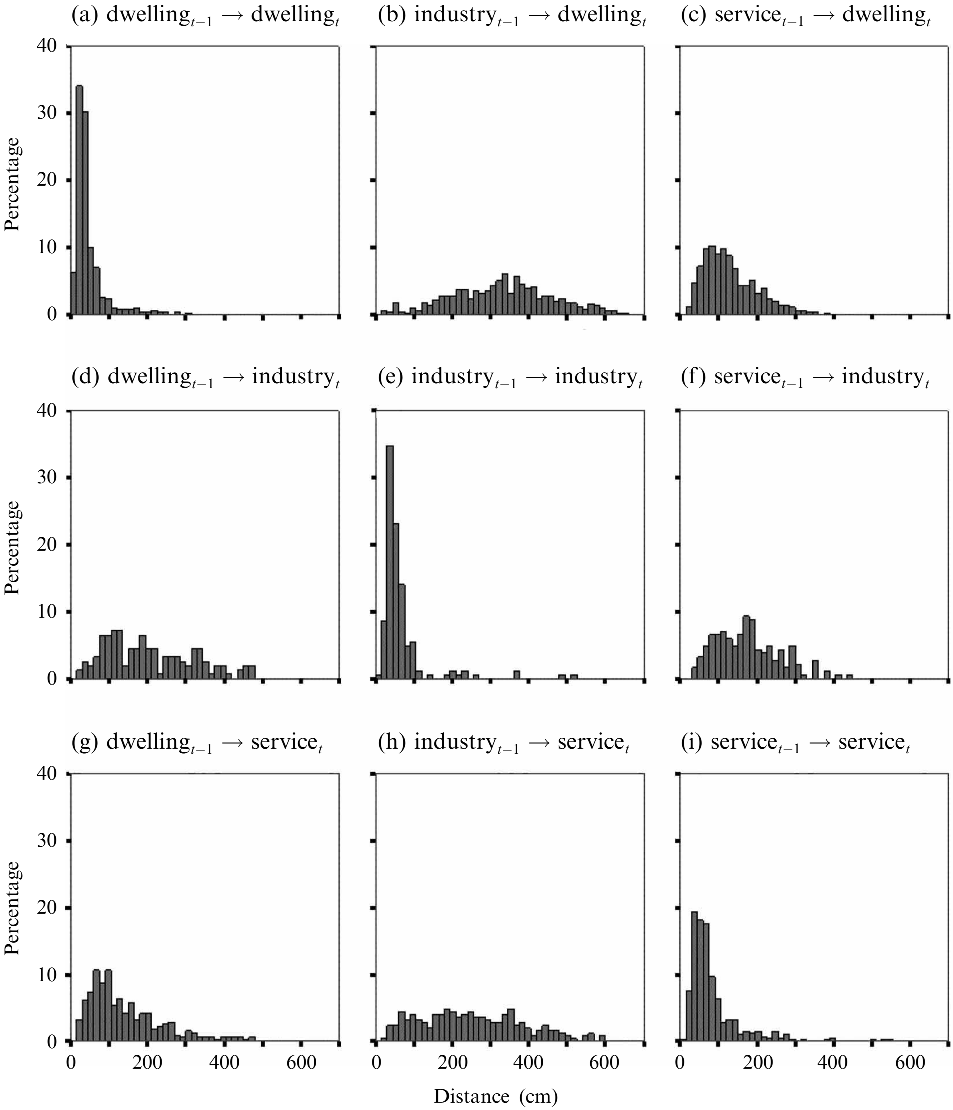


Figure 4. Distributions of distances between nearest neighbors for fifty-two steps.

distances between service and dwelling buildings represented an intermediate situation [figures 4(c) and 4(g)].

Location l_t^{b} also depends on the angle between the buildings. We consider this dependence as reflecting the participants' tendency to construct streets.

5.2.3 *The choice of location reflects the tendency to plan streets*
It is observable to the naked eye that the participants tended to locate buildings in street-like configurations (figure 3). This tendency can be quantified using the angle between consecutive buildings.

Let the building of a functional type f be located at l in P_{t-1}^{b}. Let us denote as nl_{t-1}^{b} the location of the f-type building nearest to l and as nnl_{t-1}^{b} the location of the f-type building nearest to nl_{t-1}^{b}. Let ϕ_{t-1}^{b} and $n\phi_{t-1}^{\mathrm{b}}$ be the angles between the vectors $(nl_{t-1}^{\mathrm{b}}, l_t^{\mathrm{b}})$ and $(nnl_{t-1}^{\mathrm{b}}, nl_{t-1}^{\mathrm{b}})$ and the x-axis, respectively (figure 5).

Analysis of the experiments' GIS maps reveals that the game participants prefer to orient houses either along the line of already established buildings or perpendicular to that line; formally, that means that ϕ_{t-1}^{b} depends on $n\phi_{t-1}^{\mathrm{b}}$. This tendency is strong when the distances between nnl_{t-1}^{b} and nl_{t-1}^{b}, as well as between nl_{t-1}^{b} and l_t^{b}, are below

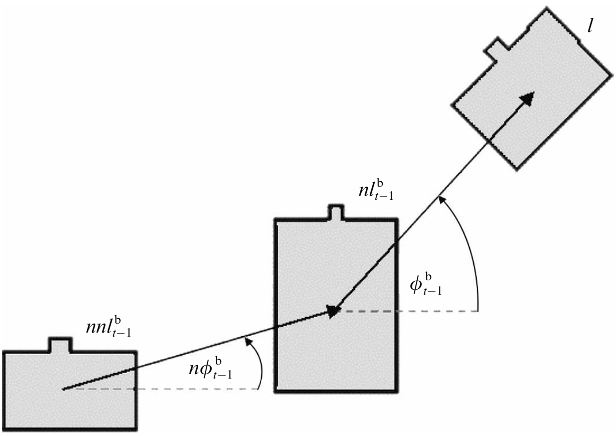


Figure 5. Vectors $(nnl_{t-1}^{\mathrm{b}}, nl_{t-1}^{\mathrm{b}})$ and $(nl_{t-1}^{\mathrm{b}}, l)$, and the difference in their direction.

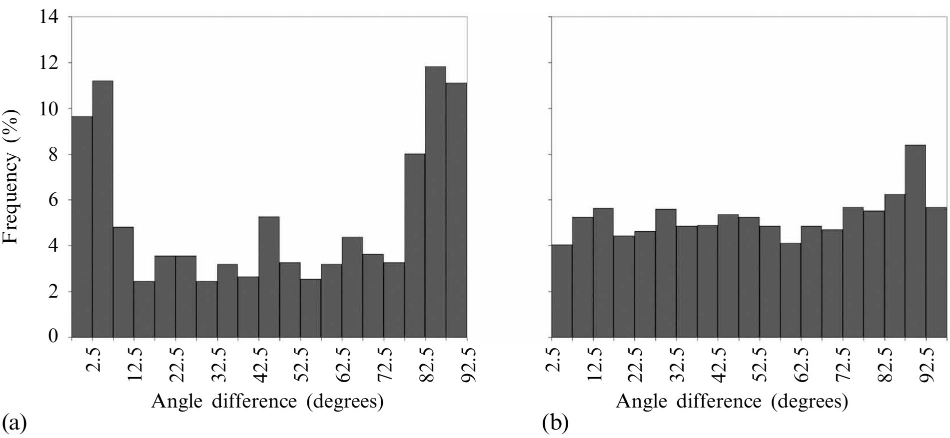


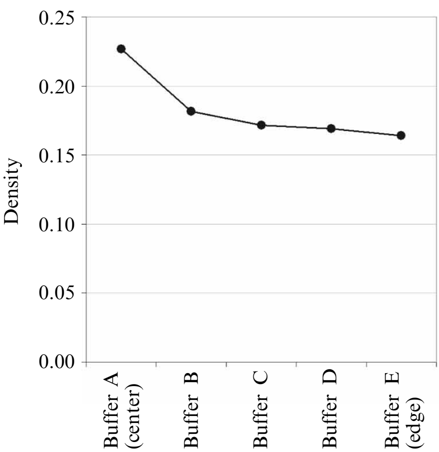
Figure 6. Histograms of the differences in direction of the vectors connecting the closest pairs of buildings (according to figure 5, angle $\phi_{t-1}^{\mathrm{b}} - n\phi_{t-1}^{\mathrm{b}}$): (a) each of two distances between nnl_{t-1}^{b} and nl_{t-1}^{b}, and between nl_{t-1}^{b} and l, is below 70 cm; (b) the remaining cases.

70 cm [figure 6(a)]; it quickly fades away and is not observed when one or both of these distances exceed 70 cm [figure 6(b)]. In what follows we employ this tendency for distances less than 70 cm.

5.2.4 *Lack of boundary effects in experimental patterns*

The above estimates of pattern development are based on the patterns' characteristics and do not account for the limited area available for construction. To distinguish between the view of the trapezoid borders as neutral, attractive (eg shorefront or recreation area) or, alternatively, unappealing (eg a highway) natural frontiers, we estimated the average density of the buildings over all thirty patterns as a function of the distance from the trapezoid's center[2] (figure 7).

According to figure 7 the overall density of the buildings is slightly higher in the absolute center and stabilizes at the higher distances; we thus accept the view that the experiment's participants regarded the trapezoid's boundaries as neutral physical borders. This assessment coincides with the participants' comments, made during the experiment. To illustrate, Amir stated: "I'm locating buildings like commerce and cultural buildings, which serve the entire city and should be highly accessible, in the center. Residential areas are located on the center fringe; this allows us to maintain good, quiet living areas, easily accessible to the center. This way people will visit the center more often."



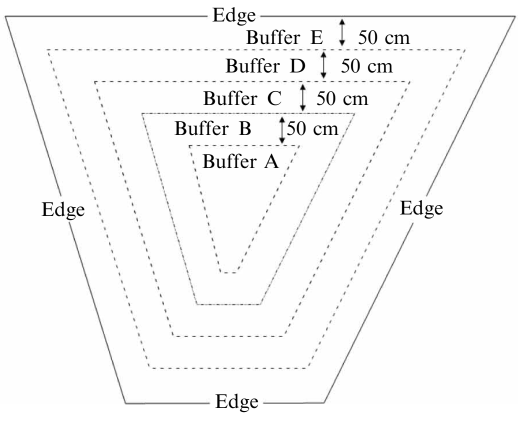


Figure 7. (a) Mean densities (per 30 cm × 30 cm area unit) of buildings per experiment as a function of the distance from the center of an area, according to the buffers of 50 cm width, constructed as in (b).

6 From specific rules to an integral description of participants' behavior

To combine the revealed dependencies, let us consider the distributions in figures 4 and 6 as potentials—three *distance potentials* regarding the nearest neighbor of each of the three functional types and a *direction potential* regarding buildings of the same functional type for expanding the street network. Assuming that these potentials determine the participants' location decisions, we can investigate their sufficiency for reproducing the patterns that emerged in the experiments (figure 3). Before attempting to do so, we have to establish how game participants might react to these potentials and how we can formally model that behavior.

[2] Thanks to an anonymous referee for raising the question of the importance of boundaries.

6.1 Possible formalization of participants' behavior: bounded rationality versus optimization

The two views of human ability to consider several factors that potentially influence behavior are generally referred to as perfectly rational and bounded rationality (Rubinstein, 1998). As a rule, if the participant's rationality is bounded, the utility of a location and the consequent location decisions are determined by 'sufficiently powerful' factors, but not necessarily the most powerful or all of the factors potentially influencing those decisions (Gigerenzer and Goldstein, 1996). In our case, participants desirous of being perfectly rational should be able to combine all four of the potentials previously listed.

Data on thirty artificial cities are sufficient to verify how participants combined *pairs* of potentials: the tests confirmed that the dependence of l_t^{b} on the distance to the *two* nearest neighbors in p_{t-1}^{b} can be presented as a *product* of the marginal dependences of l_t^{b} on the distances to each member of the pair. To illustrate, table 5 presents the experimental potential of establishing industry depending on the distances to the nearest industry *and* service buildings in p_{t-1}^{b}, versus the product of the marginal potentials for establishing industry when depending on the distance to the nearest industry and service buildings *separately* in the same pattern.

Table 5. Observed and expected (in brackets) frequencies of industry buildings, depending on distances from previously located nearest industry and service buildings ($\chi^2 = 7.4$, degrees of freedom $= 12$, $p \sim 0.83$).

Distance from industry (cm)	Distance from service (cm)				
	0–60	60–120	120–180	180–240	>240
0–60	4 (4.5)	30 (29.2)	27 (26.0)	23 (24.0)	18 (18.2)
60–120	3 (1.9)	10 (12.0)	9 (10.7)	11 (9.9)	9 (7.5)
>120	0 (0.6)	5 (3.7)	4 (3.3)	3 (3.1)	1 (2.3)

Table 6 presents the χ^2 comparisons for several pairs of potentials, estimated according to the distance to the two nearest neighbors in p_{t-1}^{b}, and as a product of two marginal potentials in six cases arbitrarily chosen among the twenty-four possible cases.

Table 6. The results of the χ^2 comparison between directly estimated two-dimensional potentials and the product of two one-dimensional potentials.

f_t^{b}-type	Nearest-neighbor type	Nearest-neighbor type	χ^2	Degrees of freedom	Significance
industry	industry	service	3.802	8	$p \sim 0.87$
industry	industry	dwelling	11.831	10	$p \sim 0.30$
service	industry	service	7.391	12	$p \sim 0.83$
service	industry	dwelling	33.122	15	$p \sim 0.05$
dwelling	industry	service	53.228	18	$p < 0.001$
dwelling	industry	dwelling	27.436	18	$p \sim 0.07$

Supported by the two-dimensional tests, we assumed that participants choose l_t^{b} following the *product* of all four potentials; they are, therefore, rational within the set of the factors we examined. Our next step was to simulate the game cities by applying the above rules. In this way, we are able to verify the sufficiency of the entire set.

7 The simulation model of participants' behavior

The model 'participants' built the city in fifty-two steps on a grid identical to the experimental grid. At each step t they applied the same set u of rules:

(1) Assign urban function f_t based on the function of the previously selected building according to the Markov transition matrix **P**, given in table 2.

(2) For selected f_t and each unoccupied cell l, calculate three distance potentials according to the distributions presented in figure 4, and a direction potential according to the distribution presented in figure 6, as well as their product. Use this product as the overall potential $a_{t,l}$ for locating an f_t building at l.

(3) Normalize the values $a_{t,l}$

$$p_{t,l} = \frac{a_{t,l}}{\sum_i a_{t,l}}. \tag{3}$$

(4) Consider $p_{t,l}$ as the probability to locate an f_t building at l and choose one of the locations according to the $p_{t,l}$ distribution.

The details of the simulation are as given in the following discussion.

7.1 Urban space

We represent the city space as a trapezoid (just as in the experiment) of cells, each 20 cm × 20 cm in size, the same as the average size of a mock-up foundation (figure 8). Each cell can contain one building only.

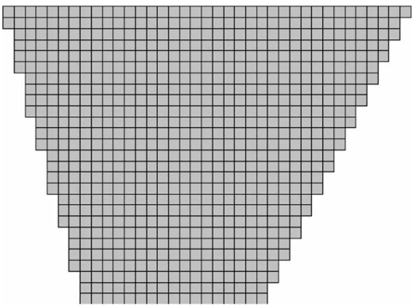


Figure 8. Representation of the 20 cm × 20 cm cell space used in the model.

7.2 Modeling a building's function f_t

7.2.1 *Choice of the initial building's function*

We initiate each model run according to the experimental estimates of the start frequencies for buildings of each of the three functional types (table 7):

Table 7. Probability to start the simulation with a building of a given functional type.

Building function	Probability of assigning the function at $t = 0$
dwelling	0.300
industry	0.067
service	0.633

7.2.2 *Type of f_t as determined by f_{t-1}*

For $t > 0$ building type is assigned according to the first-order Markov chain, with the matrix of transition probabilities **P** given in table 2.

7.3 Modeling building location l_t

7.3.1 *Locating the first building*

The mean position of the thirty mock-ups situated at the first step is only about 34 cm from the trapezoid's center. Thus, we use the center as a reference point for locating the first building in the city. Depending on its functional type, which is already chosen according to table 7, the first building is located according to the distance to the trapezoid's center. The distance to the center is estimated on the basis of thirty locations of the first building (table 8), while the angle of vector connecting the trapezoid's center and the building's location is randomly set according to a uniform distribution.

Table 8. Distributions of distances between buildings of a given functional type, with trapezoid center at $t = 0$.

Distance from the trapezoid center (cm)	Probability to locate first building at the given distance		
	dwelling	industry	service
0–25	0.00	0.00	0.50
25–75	0.22	0.00	0.30
75–175	0.22	0.00	0.05
175–275	0.33	0.50	0.05
>275	0.22	0.50	0.10

Note that, according to tables 7 and 8, more than half of the buildings in the first step are services that, at $t = 0$, are located close to the trapezoid's center.

7.3.2 *Locating* f_t, $t > 0$

The model evaluates three distance potentials, as given in figure 4, and one direction potential, as given in figure 6, for all unoccupied buildings at $t - 1$ locations. l_t is then calculated according to their product, as given in (3). Note that the potentials in figures 4 and 6 could be biased because of the high density of buildings in the latter time steps. In order to reduce that influence, we based the model on the distributions of distances obtained with the first twenty steps of each experiment; these distributions were very close to those presented in figures 4 and 6.

When estimating the direction potential, if the distances between nnl_{t-1}^{b} and nl_{t-1}^{b}, and between nl_{t-1}^{b} and l were below 70 cm, the direction potential of the l was calculated according to the distribution presented in figure 6(a); otherwise, the direction potential of l was set to uniform.

8 Evaluation of model results

To evaluate the model we generated 500 patterns of fifty-two buildings and constructed distributions of the distances between each building and its nearest neighbor of each functional type. Figure 9 presents typical examples of the simulated urban patterns.

As could be expected, the patterns expressed strong aggregation—buildings of the same functional type tended to cluster together, buildings with different functions were positioned further apart, industrial buildings tended to be located at the periphery, dwelling and service clusters were close together, and so forth.

The means and standard deviations of the nearest-neighbor index for the model and the experiments are presented in table 9. For purposes of comparison, table 9 likewise contains the nearest-neighbor indices estimated for 500 patterns produced by a 'reduced model' in which the type of building was chosen according to the Markov transition matrix **P**, given in table 2, while the location was randomly assigned.

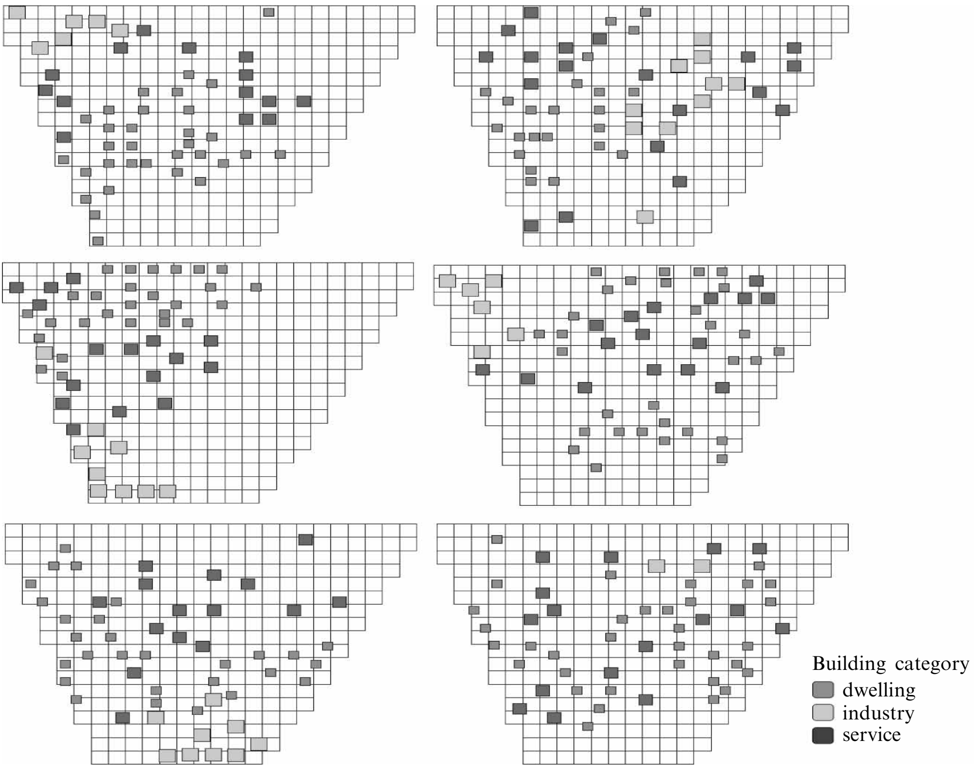


Figure 9. Examples of model patterns at step 52. The area of the squares of a given shade equals the average foundation area of the mock-ups of a given building function.

Table 9. Means and standard deviations of the distance between a building and its nearest neighbors in the experiment, model, and reduced model.

Pair	Nearest-neighbor distance constructed					
	experiment		model		reduced model	
	mean	SD	mean	SD	mean	SD
$\text{dwelling}_{t-1} \rightarrow \text{dwelling}_t$	30[a]	6[b]	43	4	51	8
$\text{dwelling}_{t-1} \rightarrow \text{industry}_t$	184[a]	97[b]	153	64	50	14
$\text{dwelling}_{t-1} \rightarrow \text{service}_t$	112[a]	56[b]	85	19	50	11
$\text{service}_{t-1} \rightarrow \text{dwelling}_t$	104	41[b]	106	24	80	19
$\text{service}_{t-1} \rightarrow \text{industry}_t$	131	63[b]	120	46	81	30
$\text{service}_{t-1} \rightarrow \text{service}_t$	61	14	62	15	84	25
$\text{industry}_{t-1} \rightarrow \text{dwelling}_t$	280[a]	97[b]	258	59	132	55
$\text{industry}_{t-1} \rightarrow \text{industry}_t$	58	32	65	37	140	74
$\text{industry}_{t-1} \rightarrow \text{service}_t$	212	63	196	58	131	55

[a] Model and experimental means differ significantly at $p = 0.05$.
[b] Model and experimental standard deviations differ significantly at $p = 0.05$. All means and a majority of the standard deviations for the reduced model differ from both experimental and model ones at $p < 0.05$.

As could be expected, the mean values of the nearest-neighbor indices obtained in the second case significantly differed from the experiment and from the model.

Regarding the experimental and model results, in about half of the cases, the differences are insignificant; in all cases excluding one, the difference between the experimental and model means is less than 20%. Nonetheless, despite its low level, the noncorrespondence demands further investigation.

We should stress that the experimental information included in the model does not automatically indicate successful simulation. Indeed, the model incorporates important assumptions that simply *cannot be tested experimentally*:

- participants' decisions at each time step are two-staged decisions, entailing choice of function and location;
- participants' location decisions regarding distance are determined by the distances to the buildings' nearest neighbors of three functional types;
- participants' location decisions regarding direction are determined by the relative location of buildings of the same functional type only;
- in the course of decision making, model participants adjust their behavior to the product of three distance potentials and one direction potential.

One or more of these assumptions may be wrong and cause the above-mentioned noncorrespondence between the experiment and the model results. Model studies make it possible to investigate a spectrum of 'what-if' questions regarding the importance of one or more of the associated assumptions. In what follows we limit ourselves to the investigation of only two of these assumptions.

Contrary to the assumption that location potential is the product of three distance potentials, and one direction potential, we assumed that the overall location potential is estimated according to the product of the *highest* of the three distance potentials and the direction potential. This 'bounded rationality' view was actually our theoretical starting point with respect to the ability of human participants to combine potentials. In brief, according to the nearest-neighbor analysis, the model patterns obtained under these assumptions corresponded much less to the experimental results than to those based on the product of all four potentials, as presented in table 9.

In the following section, we investigate in greater detail the consequences of recalling the assumption that all game participants follow the shared and time-independent set of rules.

9 Behavior of the game participants versus the average model results

Let us recall two main disagreements between the model and the game city patterns that do remain. First, the test of the hypothesis regarding the Markov chain as the model of the choice of the mock-up functional type resulted in significant χ^2 values for three-step and four-step histories (table 3). Second, disagreements remained between the mean and variance of the nearest-neighbor index, especially in the case of the dwellings' neighbors (table 9); in almost all cases, the model nearest-neighbor index varies less for the model patterns than for those obtained in the experiments.

We should also recall that the above results were obtained in a model that simulated participant behavior according to the rules formulated in section 7, with the parameters given in table 2 and in figures 4 and 6. Let us call them the set of model rules (SMR). In the following discussion we explain the differences between the model and the experiments by altering the assumption that the SMR is shared by all human participants and employed during the entire game. One can thus conceive of a participant who employs the rules with parameters different from those of the SMR, or whose rules change as the game develops, or possibly both.

To identify possible individual deviations, we compared the potentials of the buildings' locations which the participants actually produced, with the potentials of those locations as estimated according to the model. The idea was that if participant b acts according to the SMR at each time step, then the potentials of the actually selected locations l_t^{b}, $t = 0, 1, 2, ..., 51$, would be sufficiently close to the highest possible potential, calculated according to the SMR; participants' patterns P_t^{b} should therefore be close to that generated by the model.

We perform this investigation on the basis of two measures. First, we quantify the *participant's tendency to abide by the SMR* by calculating the fraction q_t^{b} of locations, free at t, where the estimated potential was higher than the potential of the actually selected location. To characterize the correspondence between the participant's behavior and the model from the beginning of the experiment until time step t, we calculate the time average, $\hat{q}_t^{\mathrm{b}}$, of the q_i^{b}-values for i between 0 and t [figures 10(c) and 10(d)]. If the participant followed the SMR, then his or her $\hat{q}_t^{\mathrm{b}}$ series remained close to zero for all t. The $\hat{q}_t^{\mathrm{b}}$ series revealed the participants who strictly followed the SMR [as in figures 10(a) and 10(c)] and those who systematically deviated from it [as in figures 10(b) and 10(d)].

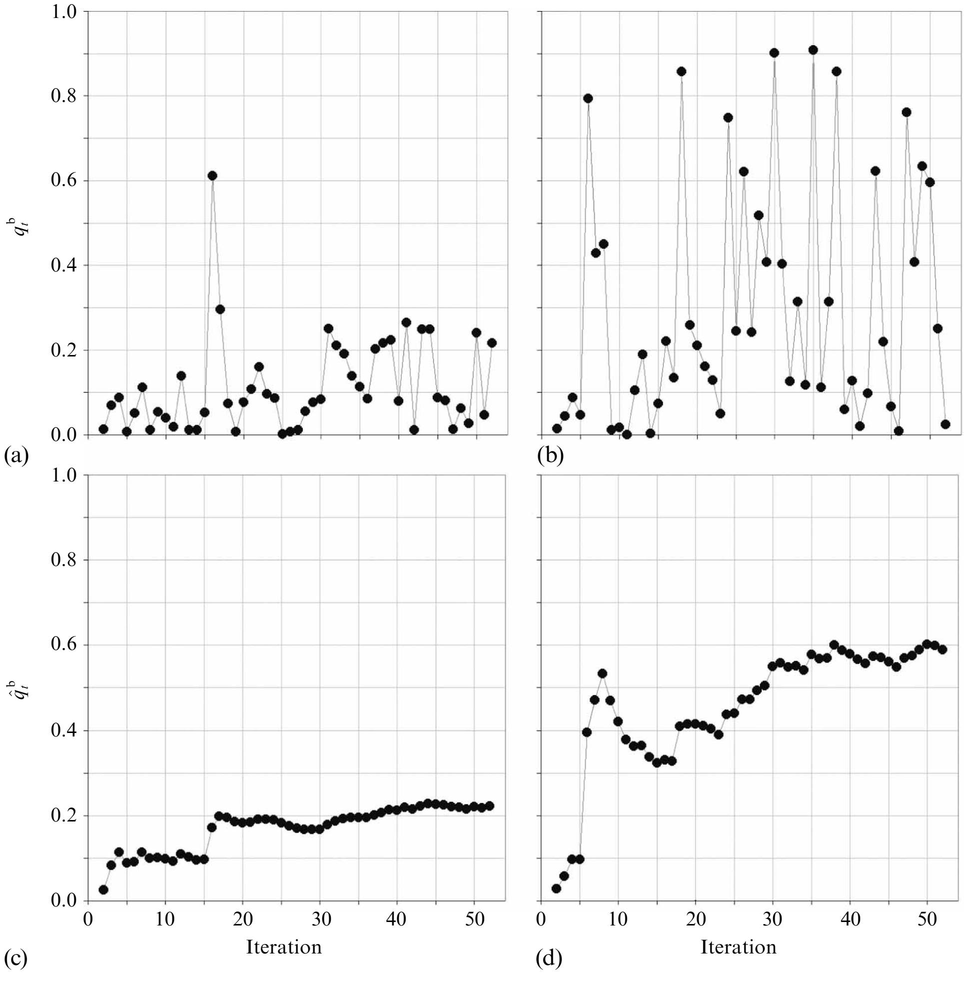


Figure 10. The difference between the player's behavior and the model: (a) model-like behavior; (b) behavior deviating from the model; (c) cumulative averge of (a); (d) cumulative average of (b).

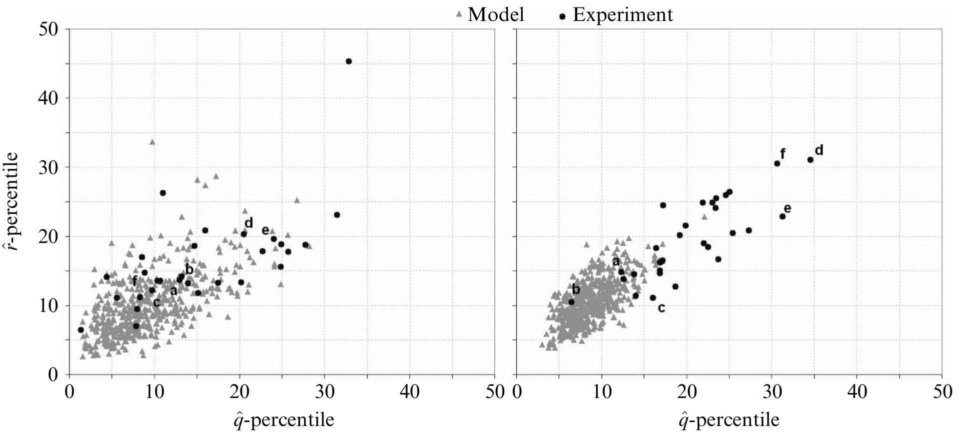
Second, to characterize whether the *game pattern resembled that produced by the model*, let us consider every building i in the experimental pattern P_t^{b} produced by b; let a building's functional type be f_i^{b} and let it be located at l_i^{b}. We then estimate the fraction $r_{i,t}^{\mathrm{b}}$ of locations at t, where the model potential for i is greater than that at its actual location l_i^{b}, assuming that the other buildings in P_t^{b} are already in place. One can readily note that the calculation procedure is equivalent to that employed for calculating q_t^{b} but that, this time, *the order of development is ignored*. The average of $r_{i,t}$ values over the pattern P_t^{b}, $\hat{r}_t^{\mathrm{b}}$, estimates the extent to which P_t^{b} *could be* an outcome of the model, although it is not at all necessary for those patterns to be constructed by a game participant who follows the SMR. To wit, one can consider the pattern P_{51} generated with the model, and rebuild it by placing the same mock-ups into the same locations but in a different order—first for all dwellings, then for all industry, and finally for all services. The P_{51} will be thus repeated by a process that has nothing in common with the SMR.

Comparison of the $(\hat{q}_t^{\mathrm{b}}, \hat{r}_t^{\mathrm{b}})$-trajectory of the individual participants b with the spectrum of the model $(\hat{q}_t, \hat{r}_t)$-trajectories makes it possible to identify game participants of three types.

First, if the trajectory of a certain game participant always remained within the model $(\hat{q}_t, \hat{r}_t)$-domain, then he or she made decisions that followed the SMR, with the preservation of the time correspondence between the individual and the model patterns.

Second, if an individual $(\hat{q}_t^{\mathrm{b}}, \hat{r}_t^{\mathrm{b}})$-trajectory extends beyond the $(\hat{q}_t, \hat{r}_t)$-domain generated by the model, two other options are possible. First, if both $\hat{q}_t^{\mathrm{b}}$ and $\hat{r}_t^{\mathrm{b}}$ differ from the model values, then the behavior of the game participant does not follow the SMR *and* the pattern P_t^{b} that he or she constructed differs from those generated by the model. Second, when the values of the q_t^{b} only fall beyond the model $(\hat{q}_t, \hat{r}_t)$-domain, then b does not follow SMR, although the patterns that the participant constructed can nonetheless be generated by the model.

To verify these three options, figure 11 displays the scatterplots of $(\hat{q}_{20}^{\mathrm{b}}, \hat{r}_{20}^{\mathrm{b}})$ and $(\hat{q}_{51}^{\mathrm{b}}, \hat{r}_{51}^{\mathrm{b}})$. As can be seen, up to $t = 20$, all participants, excluding three, behave in an SMR fashion and produce game cities that do not differ from model-generated cities [figure 11(a)]. However, with time, the situation changes [figure 11(b)], and some participants keep following the SMR until the very end, while others do not. Figure 12 presents three patterns that do not deviate from the model at $t = 51$ (marked a, b, and c



(a) (b)

Figure 11. (a) $(\hat{q}_t, \hat{r}_t)$ scatterplot at $t = 20$; (b) $(\hat{q}_t, \hat{r}_t)$ scatterplot at $t = 51$.

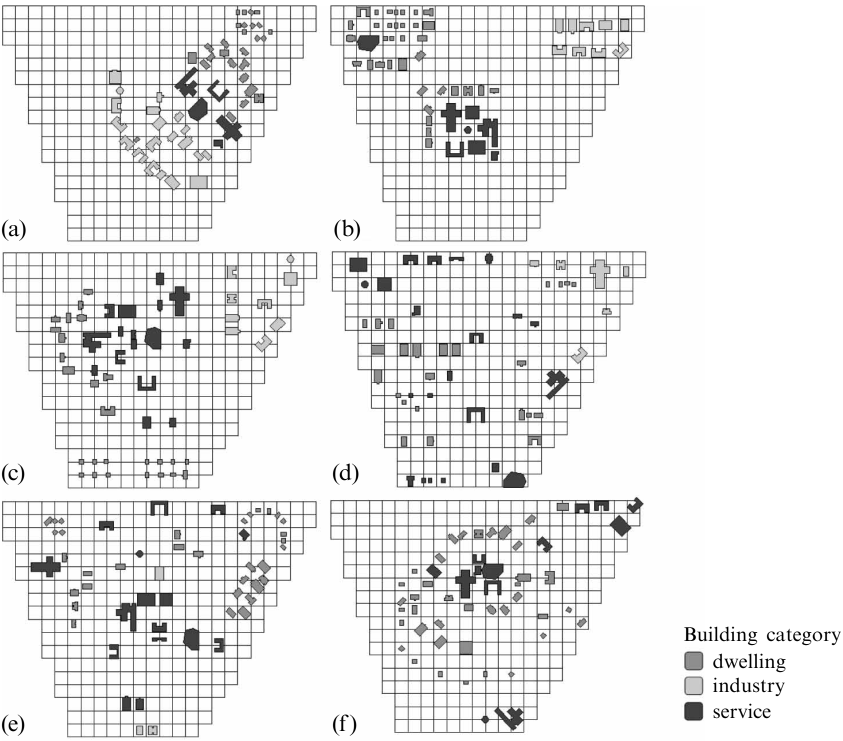


Figure 12. (a), (b), and (c) The final patterns produced by players who followed the model during all fifty-two time steps (points a, b, and c in figure 11); (d), (e), and (f) the patterns of the three players who deviated most from the model (points d, e, and f in figure 11).

on figure 11) and those for which the deviation at $t = 51$ is maximal (marked d, e, and f on figure 11). One can see that the SMR followers produced clustered cities of a definite sort [figures 12(a), (b), and (c)], while the participants who produced patterns deviating from the model [figures 12(d), (e), and (f)] likewise differed among themselves. The participant who produced the pattern presented in figure 12(d) seems to have failed to implement some definite program; the city appears quite disorganized by the end of the experiment. The cities presented in figures 12(e) and 12(f) are clearly organized: the city in figure 12(e) seems to be a mixture of the organized and disorganized parts, and that in figure 12(f) appears to be fully organized. At the same time, none of these three patterns can be produced by the SMR and they all look different from those in figures 12(a), (b), and 12(c). Did their game participants follow first-order rule sets that differ from the SMR or did they vary the rules during the game? Our data are insufficient for an answer.

To conclude the analysis of the patterns, the game experiments clearly favor the idea of a shared set of rules that can be formally considered as first-order time-independent recursion for representing the behavior of the majority of game participants. With the increase in city complexity, however, the behavior of some participants tends to deviate from the behavior generated in the model. For some of the participants this behavior results in disorganized patterns; for others the city patterns remain organized despite behavior that differs from that characterizing the majority.

Let us also note that the analysis of the experiment's spatial outcomes does not permit recognition of whether the students developed patterns that could potentially function as urban systems or whether they simply retrieved a city image from memory and then filled in the gaps in locations with the models available.[3] However, the students' comments provide informal arguments in favor of 'zoning-oriented' and 'developer-like' behavior. Consider the comments made by Edna: in the very beginning—"I'm putting the religious building, like a synagogue, in the center of the city—everyone would want to visit there." Soon after that—"And now I'm building a residential neighborhood around here", and in explaining the final pattern—"My intention was to locate the dwellings in a way that all kinds of services, such as culture and leisure, were close by." Eitan provides another example: "Here are the streets that bring people to the center of the city" and "I have located some of the buildings along a diagonal, to destroy the symmetry."

10 From games back to cities

By means of our 'development game', we investigated the individual behavior of thirty students. Although lacking any experience of real development, they had been taught what cities should look like. Each was asked to construct a city of fifty-two buildings, and was then left alone to complete the task; their behavior and the resulting game patterns were recorded in fine detail. Based on these records, the set of shared first-order recursive rules that described the choice of the urban function and the location of the buildings was formulated. A model that simulates participant behavior based on these revealed rules was likewise constructed. Comparisons between the urban patterns generated by the model and those constructed by the game participants clearly demonstrated that the majority of students acted according to the rules.

Despite the game being a remote semblance of a real city, we view our experiment as providing some important insights into the relationship between models of urban land-use dynamics and the human activities that, in effect, determine these dynamics. As summarized at the beginning of this paper, a first-order recursive set of land-change rules characterize contemporary urban modeling's default setting. Land uses, however, do not change by themselves—human beings decide where to build offices, dwellings, malls, or parks. To understand and appropriately model urban dynamics, we are required to relate between these two views.

This relationship between rules and human behavior was recently proposed by Portugali in his *interrepresentation network* (IRN) hypothesis: emerging urban patterns essentially determine ('enslave') the future decisions of urban developers and force them to further reproduce those patterns (Portugali, 1996; 2000). We consider our game as experimentally supporting the IRN hypothesis, yet extending it still further. Namely, on the basis of our experiment, we suggest that: (1) humans are able to extract the rules of a pattern's development, and (2) the majority of individuals extract similar and simple rules, rooted mainly in the immediate past. Because our participants had no experience of real development and planning activity, we can assume that the initial state of the majority of real developers and planners is similar to that of our participants—they may be enslaved from the very start. Numerous real-world constraints can only intensify this tendency to behave similarly.

Some participants did fail to produce an organized pattern, while others produced partially organized or well-organized patterns, acting in an organized but unusual fashion. The latter cases are evidently insufficient for generalizations but, nonetheless, inspire some parallels between behavioral and genetic variability. Just as genetic variability is a source for natural selection, individual variability in the formulation of development rules may be inherent in humans. Uncommon rules, just like recessive genes, can be preserved by some developers and planners. With changes in exogenous

(3) We wish to thank the anonymous referee who raised this issue.

or endogenous conditions (such as the density of the buildings in our experiments), this behavior can become advantageous. If repeated by the masses, the new path of urban development can be fixed, thus providing the bifurcation mechanism Portugali (2000) discusses but does not specify in his IRN theory.

To conclude, a real city is built by many experienced participants, all involved in numerous vertical and horizontal relationships, from cooperation to antagonism, with the majority abiding by a mass of planning directives and limitations and reacting to varying demands for land uses of different kinds. They are all, nonetheless, human beings. The results of our experiments support the view that, over extensive periods of time, the majority of developers obey a simple set of common rules that do not necessarily correspond to the written rules. We can also argue that not all participants tend to share the same rules; hence, the inherent variability of development behavior can be preserved.

We consider our experiment as a preliminary yet important step towards adequate modeling of decision-making behavior among real developers and planners.

Acknowledgements. We would like to acknowledge Dr Hernan Casakin for his valuable contributions in the experimental part of the research.

References

Bell E J, 1974, "Markov analysis of land use change: an application of stochastic processes to remotely sensed data" *Socio-Economic Planning Science* **8** 311 – 316

Benenson I, Omer I, Hatna E, 2002, "Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel-Aviv" *Environment and Planning B: Planning and Design* **29** 491 – 512

Benenson I, Torrens P M, 2004 *Geosimulation: Automata-based Modeling of Urban Phenomena* (John Wiley, New York)

Boerner R E J, DeMers M N, Simpson J W, Artigas F J, Silva A, Berns L A, 1996, "Markov models of inertia and dynamism on two continuous Ohio landscapes" *Geographical Analysis* **28** 56 – 66

Brown D G, Pijanowski B C, Duh J D, 2002, "Modeling the relationships between land use and land cover on private lands in the Upper Midwest, USA" *Journal of Environmental Management* **59** 247 – 263

Camacho-Cuena E, Garcia-Gallego A, Georgantzis N, Sabater-Grande G, 2005, "Buyer – seller interaction in experimental spatial markets" *Regional Science and Urban Economics* **35** 89 – 108

Cecchini A, Rizzi P, 2001, "Is urban gaming simulation useful?" *Simulation and Gaming* **32** 507 – 521

Clark P J, Evans F C, 1954, "Distance to nearest neighbor as a measure spatial relationships in populations" *Ecology* **35** 445 – 453

Clarke K C, Hoppen S, Gaydos L J, 1997, "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area" *Environment and Planning B: Planning and Design* **24** 247 – 261

Collins R, Sherstyuk K, 2000, "Spatial competition with three firms: an experimental study" *Economic Inquiry* **38** 73 – 94

Gigerenzer G, Goldstein D G, 1996, "Reasoning the fast and frugal way: models of bounded rationality" *Psychological Review* **103** 650 – 669

Gobin A, Campling P, Feyen J, 2002, "Logistic modelling to derive agricultural land use determinants: a case study from southeastern Nigeria" *Agriculture, Ecosystems and Environment* **89** 213 – 228

Hathout S, 2002, "The use of GIS for monitoring and predicting urban growth in East and West St Paul, Winnipeg, Manitoba, Canada" *Journal of Environmental Management* **66** 229 – 238

Hotelling H, 1929, "Stability in competition" *Economic Journal* **39** 41 – 57

Huck S, Muller W, Vriend N J, 2002, "The East end, the West end, and King's Cross: on clustering in the fourplayer Hotelling game" *Economic Inquiry* **40** 231 – 240

Isaacson D L, Madsen R W, 1976 *Markov Chains: Theory and Applications* (John Wiley, New York)

Jahan S, 1986, "The determination of stability and similarity of Markovian land use change processes: a theoretical and empirical analysis" *Socio-Economic Planning Sciences* **20** 243 – 251

LaGro J A, DeGloria S D, 1992, "Land use dynamics within an urbanizing non-metropolitan county in New York State (USA)" *Landscape Ecology* **7** 275 – 289

Lopez E, Bocco G, Mendoza M, Duhau E, 2001, "Predicting land-cover and land-use change in the urban fringe A case in Morelia city, Mexico" *Landscape and Urban Planning* **55** 271 – 285
Mayer I S, Carton L, de Jong M, Leijten M, Dammers E, 2004, "Gaming the future of an urban network" *Futures* **36** 311 – 333
Pentland A, Liu A, 1999, "Modeling and prediction of human behavior" *Neural Computation* **11** 229 – 242
Portugali J, 1996, "Inter-representation networks and cognitive maps", in *The Construction of Cognitive Maps* Ed. J Portugali (Kluwer, Dordrecht) pp 11 – 43
Portugali J, 2000 *Self-organization and the City* (Springer, Berlin)
Portugali J, 2002, "The seven basic propositions of SIRN (synergetic interrepresentation networks)" *Nonlinear Phenomena in Complex Systems* **5** 428 – 444
Radford A, 2000, "Games and learning about form in architecture" *Automation in Construction* **9** 379 – 385
Rubinstein A, 1998 *Modeling Bounded Rationality* (MIT Press, Cambridge, MA)
Semboloni F, Assfalg J, Armeni S, Gianassi R, Marsoni F, 2003, "CityDev, an interactive multi-agents urban model on the web" *Computers, Environment and Urban Systems* **28** 45 – 64
Weng Q, 2002, "Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling" *Journal of Environmental Management* **64** 273 – 284
White R, Engelen G, 1997, "Cellular automata as the basis of integrated dynamic regional modelling" *Environment and Planning B: Planning and Design* **24** 235 – 246
Yeh A G O, Li X, 2002, "A cellular automata model to simulate development density for urban planning" *Environment and Planning B: Planning and Design* **29** 431 – 450