The Probabilistic Generation of Characteristic Urban Structure

Stephen Marshall
s.marshall@ucl.ac.uk

Bartlett School of Planning,
University College London

12th June 2005

Synopsis

The paper suggests how certain typical urban layout patterns may arise in the absence of conscious overall ‘design’. The ‘characteristic’ urban structure is first introduced, and then programs are used to demonstrate the generation of forms, and the resultant probability of different structures arising. These are then compared with the structural properties of actual street patterns.

Background

Many urban layout patterns – especially certain kinds of traditional ‘unplanned’ patterns – are often interpreted as being ‘amorphous’, ‘chaotic’ or ‘formless’ (Marshall, 2005). Such patterns are often seen not only to be intrinsically without structure, or of an unfathomably complex structure, but are also seen as not requiring explanation, since they may be written off as ‘formless’ or ‘random’. This lack of attention did not matter in the era when Modernist town planning was in the ascendancy, because the focus was on creating new, rational layouts with definite prescribed structures.

However, with the upsurge of interest in neo-traditional urbanism, many contemporary urban designers and planners are once more interested in the qualities of traditional urban layout patterns, that may be considered desirably ‘functional’ in the sense of providing ‘urbanism that works’ (as opposed to the would-be functional*ism* of Modernism that often turned out dysfunctional in practice). If neo-traditional urban layouts are now to be considered as possible design ‘solutions’ then there is a need for a better understanding of what the traditional structures actually are, how those patterns were created historically (in the absence of deliberate town planning) and how those patterns might be replicated through contemporary design.

In addition to the 'design-led' prerogative for examining traditional urban structures, the emergence of the disciplines such as complexity theory, which have tackled previously unfathomable phenomena, would also suggest that the attempt to deconstruct the complex structure of traditional urban layouts might be worthwhile.[1]

Previously published research by the author has suggested that traditional urban layouts do have a typical or 'characteristic' structure which can be quantified according to route-structural parameters such as relative connectivity (X) and complexity (Ω) (Marshall, 2005). This paper now investigates possible generative explanations for this 'characteristic structure' – this combination of connectivity and complexity found in traditional planned urban layouts.

The paper first briefly introduces the concept of 'characteristic structure'. The paper then goes on to demonstrate how different 'programs' or rule systems for generating structure can give rise, probabilistically, to the observed characteristic structures.

Characteristic Structure

Research into the structure of street patterns has suggested the identification of a kind of urban structure – termed characteristic urban structure – which is typically found in traditional urban layouts (Marshall, 2005).

The term 'characteristic' embodies two meanings. Firstly, it implies the possession of distinctive *character*. Here this means structures that have a quintessential 'street pattern shape', that includes the following characteristics:

- a mixture of short and long routes, and more and less connective routes;
- some differentiation of routes by depth, but overall not too great a depth;

[1] The sciences of complexity have been linked to urbanism, for example, with regard to the generation of complex emergent patterns from simple rules (Batty, 1995, 2000; Batty and Xie, 1997); social networks and complex adaptive systems (Green, 1999); small world networks (Watts and Strogatz, 1998); interpretations of the complexity of architectural form (Salingaros, 1997); 'new science and new architecture' (Jencks, ed., 1997); emergence (Johnson, 2001); and network theory (Barabási, 2002).

- three-way junctions are typically in the majority, but there is a likelihood of at least some crossroads and culs-de-sac;
- a medium or ‘semi-griddy’ level of connectivity, with a relative connectivity (Χ) of around 0.35-0.45;
- a relatively high degree of irregularity and complexity, with complexity (Ω) typically in the range 0.35 to 0.6.

These characteristics has been explored in a previously published work (ibid., 154).

The second connotation of ‘characteristic’ is that of *likelihood*; and refers to the sense that a certain kind of characteristic structure is somehow ‘likely’ to exist as a street pattern. This second meaning of ‘characteristic’ is what will be demonstrated in this paper, and referred back to as an explanation for the first.

Programs for generating structures

Urban structure may be defined in a variety of different ways, and a variety of different processes of formation may be observed. Marshall (2001) has suggested three significant processes which may be identified with operations on different topological elements. The three elements considered are building footprints, plots and routes; and the associated processes are respectively termed footprint formation, plot subdivision and route propagation. There is no space here to discuss the significance of these, which will be reported more fully in another work (Marshall, forthcoming).

Here, the intention is to demonstrate how different ‘programs’ for generating structure can give rise, probabilistically, to the observed characteristic structures. This is demonstrated using one of the processes of formation – route propagation – by means of two simple programs: firstly a ‘T-tree’ program and secondly an ‘X-cell’ program.

By generating a finite set of ‘all possible patterns’, from a given program, we can see where the ‘likely’ patterns lie in relation to all patterns, and hence quantify their probability. This will illustrate how some urban patterns are more likely to arise than others, in the absence of overall planning.

The T-tree program

The creation of a set of 'all possible patterns' will depend on what rules are used to create the individual patterns. To illustrate this we will first use a very simple structure-generating program that generates a set of patterns, each of which is a tree structure formed by T-junctions. This shall be referred to as the 'T-tree program'. The simple nature of the program merely puts a limit on the absolute number and complexity of structures to manageable proportions, for the purposes of explanation; nevertheless, the underlying logic can be applied to any complexity of structure at any scale.

The first assumption is that we are dealing with structures composed of *elemental parts* – intrinsically similar components, which are only subsequently differentiated by the ways they are put together.[2] In other words, we can imagine creating structures from a 'kit of parts' comprising a set of identical members (Figure 1). It is only when assembled together that these parts assume a structural identity, for example, as 'trunks' or branches'.

The second assumption relates to the fact that we are modelling the *cumulative*, *incremental* growth of a network. We shall assume that we start off with a primitive, single-route 'network' (imagine a single road along which a settlement starts to form). Additional routes are then added progressively, one at a time, with each additional route representing an iteration, or a new generation. Hence, at any stage, the number of generations equals the number of routes (Figure 1, i-iii).

Thirdly, we assume that the position at which a new element is added is selected *at random*, with all allowable positions being equally probable candidates for selection. This simulates the 'random' growth of unplanned patterns, as a result of a series of independent actions at the local scale.

[2] The assumption that each element is in principle identical means that we can regard all structures that are topologically equivalent to be identical. If each individual element were different in itself, no two structures would be the same, and all structures would be equally (im)probable.

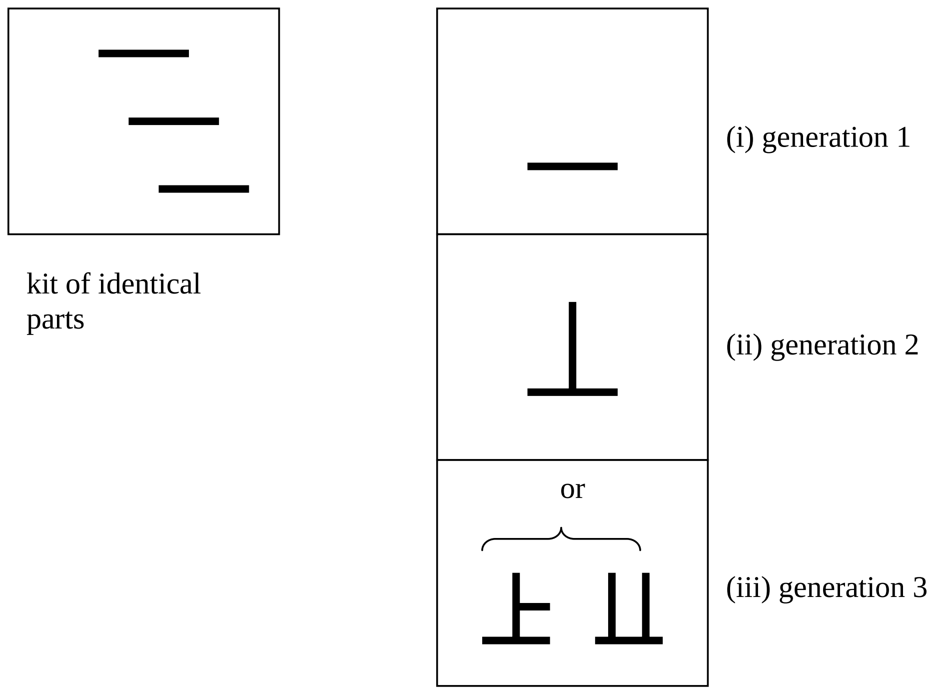


Figure 1. Assembly of elemental parts. We start with a kit of three identical parts (left hand inset). The first generation (i) creates a one-element figure; the second generation (ii) creates a two-element figure; the third generation (iii) creates two possible three-element figures – consistent with the 'T-tree' program. The elements become differentiated in character according to how they are connected: for example, 'trunk' and 'branch' elements become distinguishable.

The first three assumptions effectively represent the ground rules for the cumulative formation of incremental, random structure. The fourth and final assumption – which gives this program its particular structural character – is that that we will only allow each route to be added to the network in such a way that it forms a 'T-tree'. Basically, each new element is added to 'cantilever out' from an existing element[3]. The overall program is summarised in Box 1.

Box 1. The T-tree program.
1) Each constituent element is identical.
2) Each element is added to form structure one at a time.
3) Each new element occupies a position on the structure that is chosen at random.
4) Each new element joins the existing structure at only one of its ends, to form a 3-way connection (T-junction), such that the overall structure is a 'T-tree':

For a given program, we can work out all possible outcomes, and hence quantify the probability of any individual outcome relative to the set of all possible outcomes.

The full range of permutations for the T-tree program can now be explicitly worked out. Recall that the purpose is to work out the theoretical probability of each pattern arising, which will later be related to observed prevalence of actual street patterns. We follow the four assumptions of the T-tree program introduced above. Figure 2 shows the generation of all possible variants of structure up to the fifth generation.

With every generation, the number of possible permutations increases. For the fourth generation, there are 6 possible variants, for the fifth generation, there are 24 (i.e., Va to Vi, Figure 2). (In fact, the total number of permutations increases factorially. At the fifth generation, the number of permutations is 1x2x3x4 = 24. At the sixth generation, there are 1x2x3x4x5 = 120 permutations.)

[3] The T-branching tree is a network that is mathematically a tree (no circuits, zero redundancy) whose nodes are all 3-way, or T-junctions. In terms of our rule system, we will therefore only allow a new route to join the existing network by a connection at one of its ends, and this connection must be made at a mid-link position on an existing route, in other words, not at an existing node.

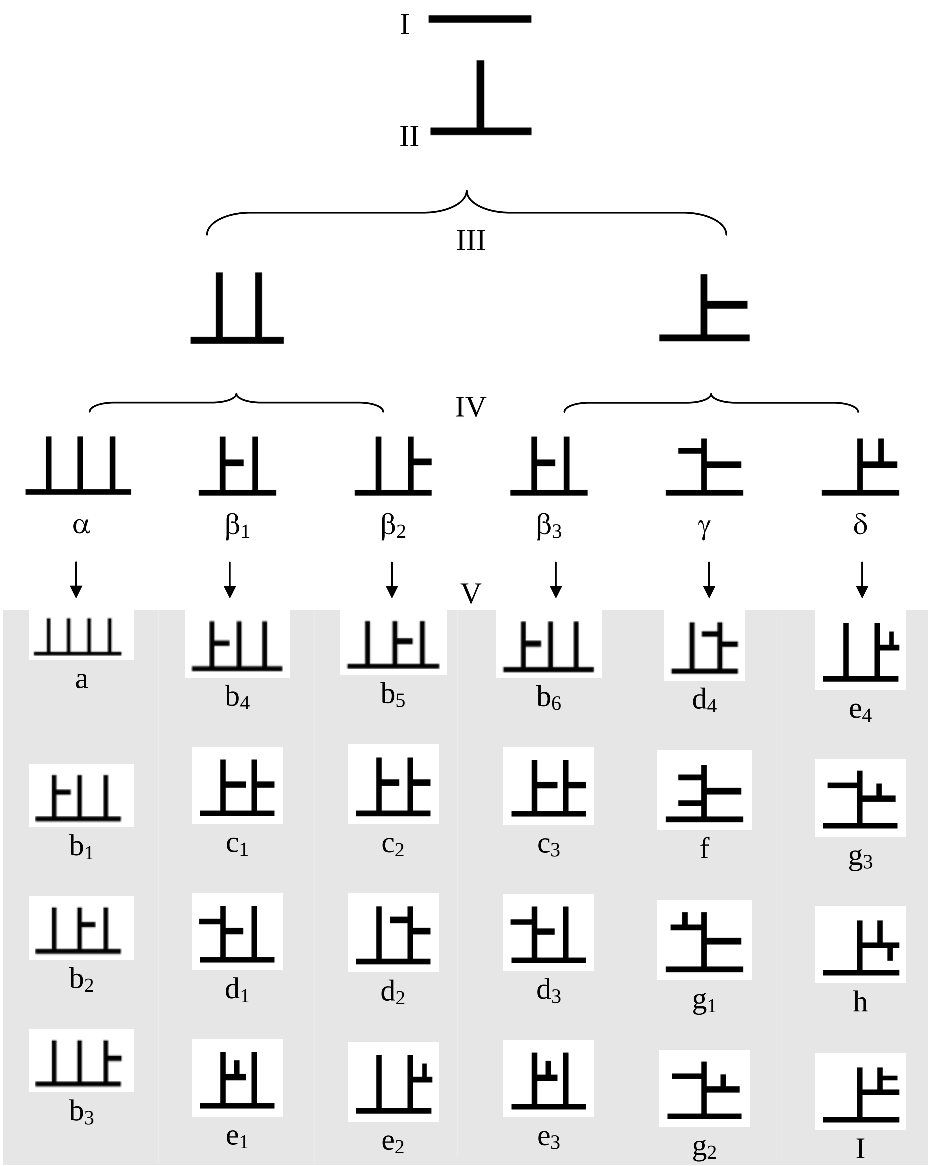
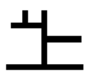
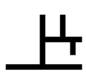


Figure 2. Demonstration of all possible patterns generated by the 'T-tree' program. The rule set for the program is given in Box 1. The program is shown here to the fifth generation. At the fifth generation, there are 24 resulting outcomes, of which there are nine unique topologies (a-i).

With each generation, there is an increasing divergence of types. We can clearly see the 'family resemblance' between the structures – they are often similar; in some cases the variants are topologically identical: indeed there are only 9 distinct structural types at the 5th generation. Hence some topologies will be more numerous, that is to say *more probable,* than others. For example, the type (b) – represented by variants

$b_1, \ldots b_6$ – is more likely than (in this case, six times as likely as) the type represented by (a), which is a singular case. For a given program, then, the theoretical probabilities of all possible patterns can be explicitly calculated for each generation. Table 1 does this for the T-tree program's fifth generation.

Table 1. Frequency and probability for patterns generated by the T-tree program, at the 5th generation (see Figure 2)

Case		Frequency	Probability
a		1	4%
$b_{1\text{-}6}$		6	25%
$c_{1\text{-}3}$		3	12.5%
$d_{1\text{-}4}$		4	17%
$e_{1\text{-}4}$		4	17%
f		1	4%
$g_{1\text{-}3}$		3	12.5%
h		1	4%
i		1	4%
Total		24	100%

The comb structure (a) and the fractal structure (i) are both unlikely because, in each case, only one possible generational path leads to that pattern (for this given program). Hence, a random selection of a structure from Figure 2 is more likely to return a 'characteristic' or 'irregular' pattern like structure (b) or (c) than a comb (a) or fractal (i). The comb and fractal can be regarded as extreme cases. It turns out that the most likely structures are those that are neither extremely regular nor extremely irregular.

This demonstrates the basic mechanism whereby urban elements can grow, probabilistically, and, as it were, 'entropically'. That is, it shows how the cumulative effect of random generation can yet give rise to characteristic or typical looking urban patterns. And it demonstrates for urban patterns the difference between the field of *possible* types and the landscape of *probability* – in other words, although here 9 types (a-i) are possible, some are more probable than others.[4]

In fact we can recognise an emerging relationship between apparent type and typicality. If we look back at the fourth generation (Figure 2), there are six outcomes, representing four distinct topologies or types. We could choose to group these into three classes of type – two pure classes and a mixed class – let's say 'comb', 'fractal' and a third 'irregular' class whose members have a mixture of the characteristics of both. Figure 3 demonstrates the start of an emerging pattern whereby some types of urban pattern are more likely than others, partly on the basis of how those patterns unfold over time, and partly on the basis of what types are recognisable in the first place.[5]

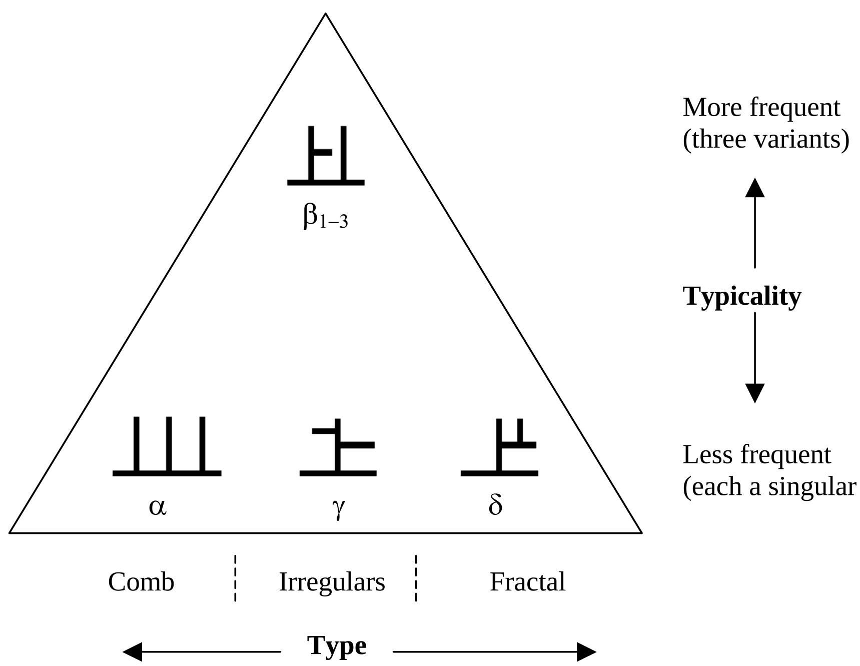


Figure 3. Emerging impression of type and typicality. The horizontal dimension represents the number of distinguishable classes of type. In this case three classes are recognized (where β and γ are grouped in the same 'irregular' class.). The vertical dimension represents frequency and hence probability (here interpreted as typicality). The apparent probability will depend on what types are recognised as distinct in the first place.

---

[4] The underlying 'landscape of probability' is discussed by Marshall (forthcoming).

[5] Here, we class all types b_{1-6} as equivalent, since all have three branches and one sub-branch – though the exact location of the sub-branch differs between cases. Some generalisation is necessary, otherwise all urban patterns will be geographically and geometrically unique, and all equally (im)probable.

It is not just route structures that will unfold with this kind of probabilistic logic. Although the demonstration here has been limited to a program building up structures of routes, the same mathematical mechanism could apply to other forms of urban generation, such as footprint formation or plot subdivision. These are not pursued here, but are being reported elsewhere (Marshall, forthcoming).

The basic message here is that some patterns are more likely than others to arise by chance, in the absence of 'anthropic' direction at the scale of the pattern as a whole; or, put another way, in an incremental pattern generating process, some patterns are more likely to be offered for selection than others.

We have quantified the likelihood of different theoretical patterns occurring, out of all possible patterns. The question then becomes: to what extent do the 'likely' patterns reflect the actual patterns seen on the ground?

Matching the Actual and the Theoretical

Street patterns typically arise from a combination of 'anthropic' influences – conscious design – and 'entropic' outcomes based on the cumulative effect of random growth. We want to see how our theoretical entropic programs might lead towards the kind of patterns found in actual urban layouts.

The X-cell program

The previous T-tree program was useful for a simple demonstration, but it gives rise to labyrinthine branching networks joined only by T-junctions. In order to generate a more realistic diversity of network structures, we can generate a new program by (1) allowing routes to be added at existing junctions, thereby allowing crossroads and other multi-way junctions to form, and (2) allowing routes to join (terminate at) two different routes, thereby allowing circuits to form. We can all this the 'X-cell' program, since it gives rise to a network with enclosed cells or circuits, where the joints may be four-way (X) or multi-way connections.[6] The rule set is summarised in Box 2.

---

[6] We can readily imagine two other related programs (not illustrated): a 'T-cell' program, allowing circuits (cells) but joined only by T junctions; and an 'X-tree' program, allowing multi-way junctions but in a tree pattern only.

Box 2. The 'X-cell' program.

1) Each constituent element is identical.

2) Each element is added to form structure one at a time.

3) Each new element occupies a position on the structure that is chosen at random.

4) Each new element joins the existing structure at one or both of its ends (but not along its middle); the resulting structure can have multi-spoked nodes and form 'circuits'.

The 'X-cell' program generates a total of 39 permutations to the 4th generation (Figure 4). Having synthesised a set of theoretical structures, it is possible to compare these with the structures of actual street patterns.

Growth of a single layout over time

For a start, it is possible to compare the actual historic sequence of urban growth with a trajectory through the theoretical solution space of Figure 2. Figure 5 (a to e) shows the early historical growth of Perth, Scotland, cross-referenced to the corresponding abstract structures of the X-cell program (Figure 4)[7]. Note that the IVe structure happens to be one of the equally most likely forms at the fourth generation (3 out of 33 cases).[8]

[7] Strictly speaking, the last iteration departs from the X-cell program by extending a route to connect with another, rather than generating a new route.
[8] Note also, however, that any correspondence between theoretical and actual is based on the assumed program in the first place. The actual is being compared with a theoretical that contains built-in assumptions.

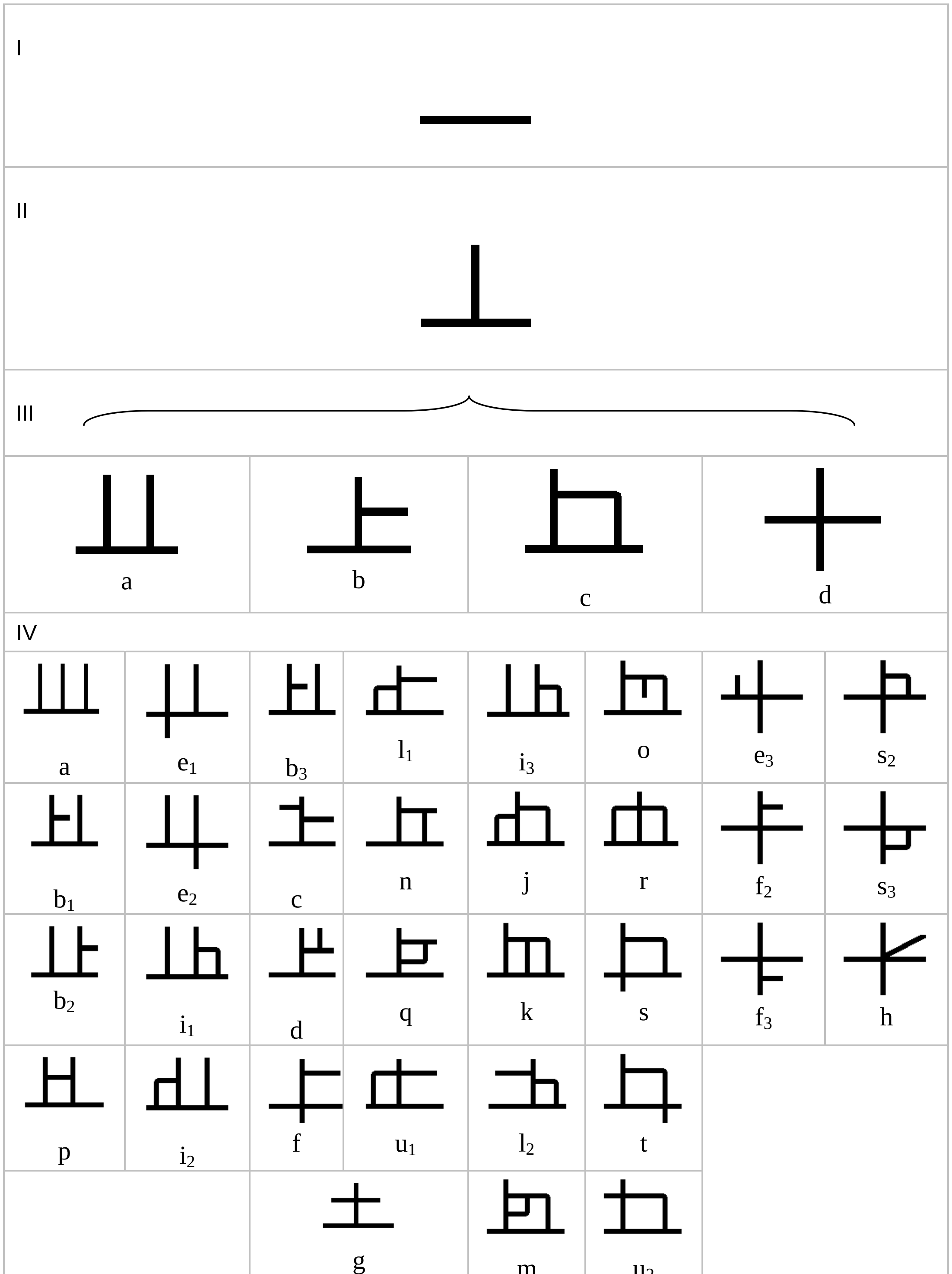


Figure 4. Demonstration of all possible patterns generated by the 'X-cell' program. The rules are given in Box 2. The program is shown run to the fourth generation. At the fourth generation, there are 39 resulting outcomes, of which there are 21 unique topologies (a-u).

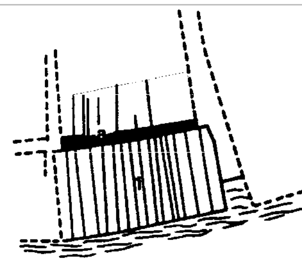
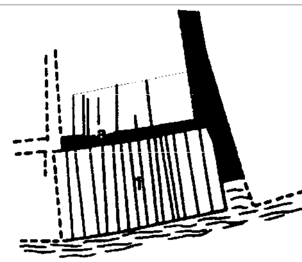
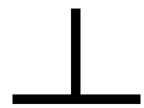
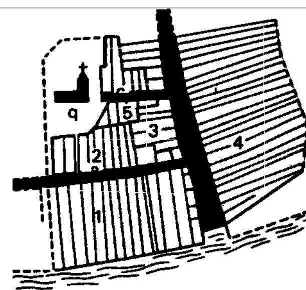
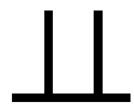
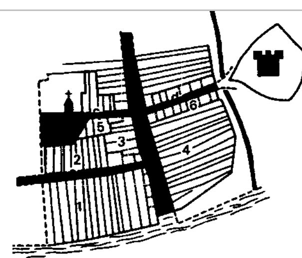
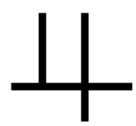
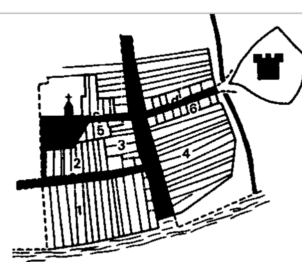
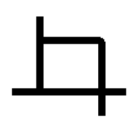
Stage of growth	Base plan	Structure and code (Figure 4)
(a) Earliest 1st generation (1 route)		 I
(b) Intermediate 2nd generation (2 routes)		 II
(c) Early 12th Century 3rd generation (3 routes)		 IIIa
(d) Mid 12th Century 4th generation (4 routes)		 IVe$_2$
(e) Mid to late 12th Century 4th generation (4 routes extended)		 IVt

Figure 5. The early growth of Perth, Scotland, cross-referenced to structures in Figure 4. Base plan after Spearman (1988).

Note also that with each generation, the probability of any particular type tends to decrease. Perth's third generation pattern (IIIa) had a 25% chance of occurring (according to the given program), while its fourth generation type (IVe) had a 9% chance. The smaller the network, the more likely a particular structure will occur, and moreover, given that there will be more smaller networks than larger ones, the likelihood of matching cases will be greatest in small networks. Clearly, the number of urban structures comprising a simple crossroads will be very large, while the topological structure of any city (or any town of any size) will be practically unique.

From this example, we can see in principle how growth of the actual town plans over time will correspond to trajectories through the solution space of Figure 4. What we need is an impression of where a whole range of town plans would lie in relation to each other. To do this, we will invoke properties of a set of actual town plans that has been studied with respect to the properties of their route structure, making use of a graphic device called the 'netgram'.

'Netgram' comparison

It is possible to characterise the structure of street patterns in terms of three route-structural variables known as relative continuity (λ), relative connectivity (χ) and relative depth (δ) (Marshall, 2005). These three properties are so defined that they sum to one (i.e. $\lambda + \chi + \delta = 1$). As a result, a triangular plot termed the 'netgram' may be used to locate the structure of any street pattern in relation to any other, in terms of these three variables. This bounded diagram effectively defines a solution space in which all theoretically possible structures would lie.

The Actual

Figure 6 shows a zone occupied by structures that characteristically appear as actual street patterns.[9] It is apparent that this zone fills out only a limited portion of the overall territory: in fact, the actual patterns lie within a zone occupying just over 6% of the whole plot.[10]

---

[9] This zone is an indication of the area in which 36 'actual' street patterns lie, as analysed in Marshall (2005).

[10] Marshall (2001). Estimated at resolution of a grid of triangles of 0.05 side length of the whole figure.

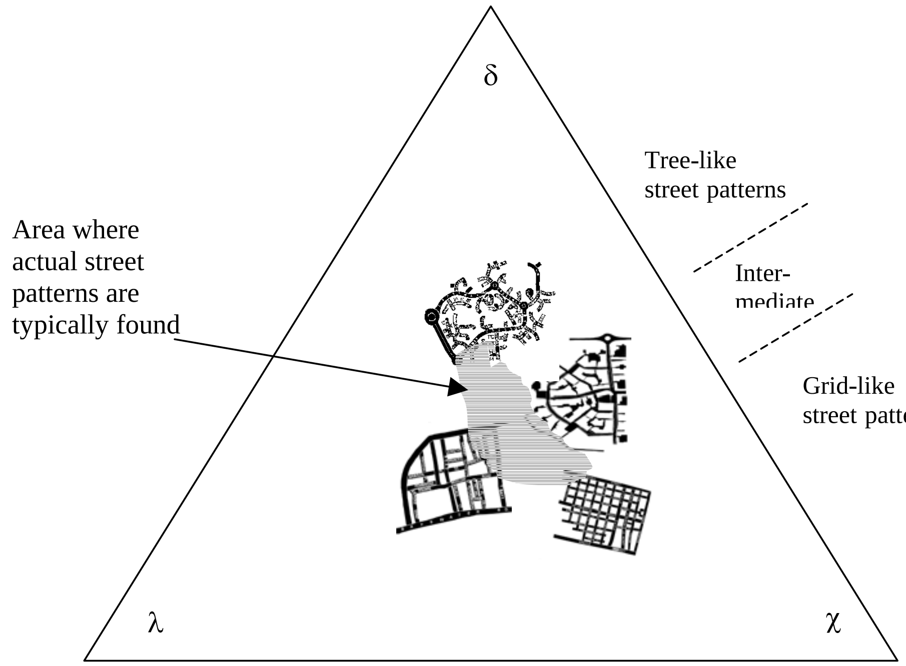


Figure 6. Location of actual street patterns on the 'netgram'. Actual street patterns tend occupy a region (shown shaded) in the middle of the netgram.

It turns out that the areas of the netgram in which no real street patterns are observed – towards the vertices and around the edges in general – are occupied by structures with many repeated or recursive elements: an esoteric collection of multi-spoked stars and fractals (Marshall, forthcoming). Overall, we can say that actual street patterns tend to be heterogeneous, relative to all possible patterns, and this heterogeneity places them where they are relative to all theoretically possible patterns.

The Actual and the Theoretically Probable

The theoretical structures generated by the X-cell program can now be compared with the set of actual street patterns (Figure 7). This would not be expected to be an exact match, conceptually or empirically, since the theoretical cases here are relatively small (four-route) networks, and network properties tend to vary with size. Nevertheless, it can be seen that they lie in an interior zone on the netgram close to, and overlapping with, the zone occupied by a set of actual street patterns evaluated. This represents a common area that is still under 7% of the total area of the netgram.

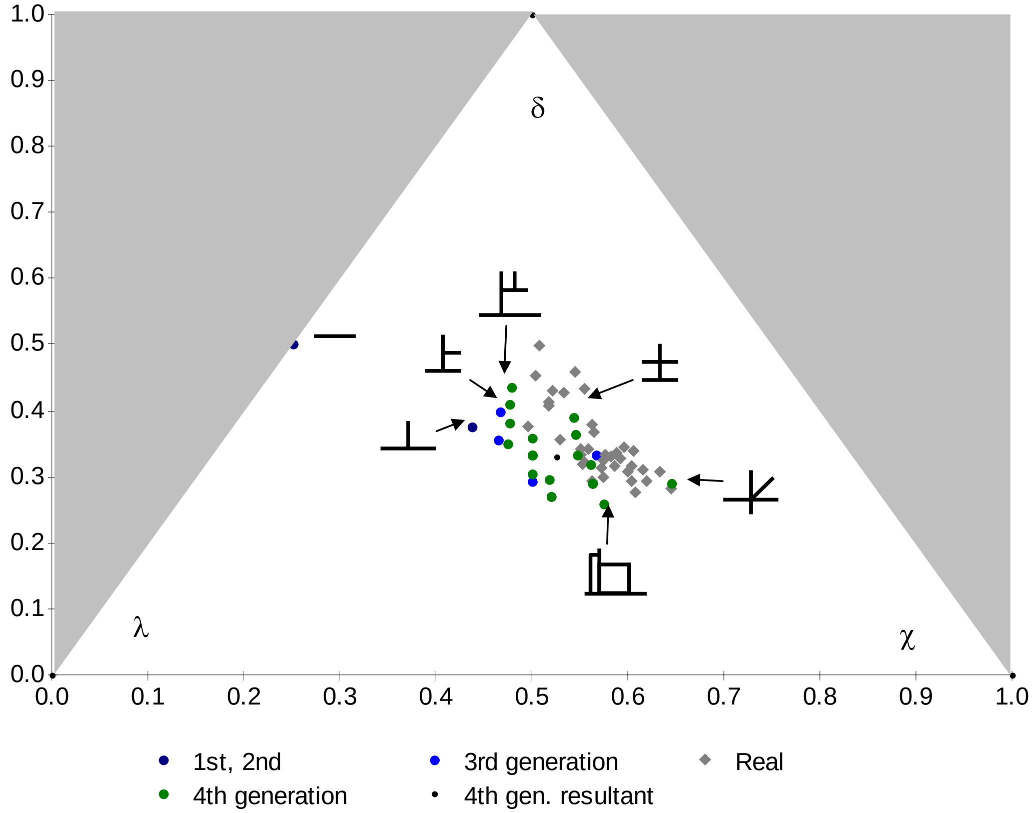


Figure 7. The position of 36 actual street patterns plotted against the theoretical 'X-cell' patterns in Figure 4. The location of the theoretical structures tends towards the zone in which the scatter of actual (street pattern) structures lies. The theoretical structures are much smaller than the real structures, and lie in a less deep position on the netgram.

Discussion

This exercise provides a link between theoretical probability and the observed distribution of actual street pattern structures as appearing on the netgram. The actual street patterns, characterised by heterogeneity, probabilistically lie in a small zone close to the centre of the possible solution space. The patterns lying at the extrema of this zone tend to be planned structures of consistent regularity: tree-like networks lie towards the top left of the zone of actual street patterns, while repetitive grids lie towards the lower right. Overall, therefore, the distribution of actual street patterns can be explained to some extent from the distribution of theoretical patterns, and in turn by the type of program generating them.

Effectively, the extremes of connectivity correspond with the most ordered or patterns, requiring the greatest design intent to contrive them, whereas the 'unplanned' patterns form the mid range of connectivity. Put another way, it seems that it requires careful planning or ordering to achieve the balance required to avoid a natural differentiation of routes. This reflects the existence of 'likely' patterns.

Conclusion

This paper has demonstrated how, using simple programs (the so-called 'T-tree' and 'X-cell' programs) we can start to see the probabilistic emergence of characteristic semi-regular, semi-complex forms that we associate with the street pattern shape. This lends weight to the suggestion that there is a relationship between type and typicality, or the two connotations of the term 'characteristic', suggested earlier. Indeed, it may be said that characteristic structures are not 'likely to be designed' because they nicely fit an abstract archetype of the quintessential 'street pattern shape', but rather the quintessential 'street pattern shape' is identifiable precisely because it is 'likely' to emerge – and hence is pervasive – in the *absence* of design intent.

By reaching some sort of correspondence between the actual and theoretical, this is not to 'prove' that random, local-rule processes account for all or any observed patterns (not least because any pattern, however 'naturally' unlikely, may be deliberately contrived by the designer). The point here is simply to highlight the idea that probabilistic processes will account for *some* proportion; that the cumulative effect of simple random processes can on their own give rise to the typical patterns seen on the ground. This reminds us that street patterns are not solely the product of overall design will.

By implication, if urban planners and designers wish to create (or recreate) traditional style, irregular street patterns, then it is not necessary to 'cut and paste' the structure of model layouts or impose a rigid straitjacket of layout templates to achieve desired structures. Rather, it may be sufficient to use a program or rule system, in which local relationships between elements are specified (cf. the T-tree and X-cell programs), which are enacted incrementally over time. This means there is no overall master-planner or designer prescribing (or able to predict) every detail of the final form, but certain familiar-looking characteristic structures are likely to emerge. The design of urban layout could therefore use local programs such as urban codes to generate urban structure, rather than conventional top-down 'master-planning' or 'town planning'.

While this suggestion may seem natural enough in the context of a generative theoretical exercise, the idea that urban layout might be generated with no overall 'planning' is yet to be accepted or even fully aired within the planning and design disciplines themselves.

The key to success in this respect would seem to be the ability to demonstrate that the design of parts can lead to workable wholes. Here, the complex end-result may be not so much a product of design, as a *by-product* of design attention at a lower level.

References

Barabási, A-L. (2002) *Linked: The New Science of Networks.* Cambridge: Perseus.

Batty, M. (1995) New ways of looking at cities, in *Nature* **377**, p574.

Batty, M. (2000) Less is more, more is different: complexity, morphology, cities and emergence; editorial in *Environment and Planning B: Planning and Design*, **27**, 167–168.

Batty, M. and Xie, Y. (1997) Possible urban automata, in *Environment and Planning B: Planning and Design*, **24**, 175–192.

Green, N. (1999) Art and complexity in London's East End, in *Complexity*, **4** (6), 14–21.

Jencks, C. (ed.) (1997) *New Science = Architecture?* London: Academy Editions.

Johnson, S. (2001) *Emergence. The Connected Lives of Ants, Brains, Cities and Software.* London: Penguin.

Marshall, S. (2001) *Transport and the Design of Urban Structure*. Unpublished PhD thesis, Bartlett School of Planning, University College London.

Marshall, S. (2005) *Streets and Patterns*. London: Spon Press.

Marshall, S. (forthcoming) *Cities Design and Evolution*. London and New York: Taylor and Francis.

Salingaros, N. A. (1997) Life and complexity in architecture from a thermodynamic analogy, in *Physics Essays*, **10**, 165–173.

Spearman, R. M. (1988) The medieval townscape of Perth, in Lynch, M., Spearman, M. and Stell, G. (eds.) *The Scottish Medieval Town*. Edinburgh: John Donald Publishers Ltd.

Watts, D. J. and Strogatz, S. H. (1998) Collective dynamics of ‘small-world’ networks, in *Nature* **393**, 440–442.