Studying cities to learn about minds: some possible implications of space syntax for spatial cognition.

Bill Hillier

Bartlett School of Graduate Studies
University College London
b.hillier@ucl.ac.uk
http://www.bartlett.ucl.ac.uk/
http://www.spacesyntax.org/

Originally given as the keynote paper to the Workshop on Space syntax and spatial cognition at the Conference on Spatial Cognition 2006 at the University of Bremen. Revised 27.06.08 to appear in Environment and Planning B, Planning and Design 2009

‘Some, but not all, of the spatial maps identified by neurobiological and behavioural research impose a structure that goes beyond, and in consequence alters, our interpretation of the information available in the input alone. For example, the hippocampus appears to impose a Euclidean framework onto non-Euclidean inputs (O'Keefe and Nadel, 1978)…. who see in this process the instantiation of a Kantian a priori notion of absolute space. We propose that in ‘distorting’ the sensory inputs these spatial maps may impose an order and a structure that our spatial conceptual representations require…… We point out the importance of a careful analysis of the intrinsic ‘organising factors’ that interact with spatial information to structure our knowledge of the spatial world. These organising factors act like a kind of ‘syntax in accord with which inputs to spatial systems are ordered, and in doing so they contribute meaning to the spatial representations themselves. (Petersen 1996, 569)

ABSTRACT

What can we learn of the human mind by examining its products? The city is a case in point. Since the beginning of cities human ideas about them have been dominated by geometric ideas, and the real history of cities has always oscillated between the geometric and the ‘organic’. Set in the context of the suggestion from cognitive neuroscience that we impose more geometric order on the world that it actually possesses, an intriguing question arises: what is the role of geometric intuition in how we understand cities and how we create them? Here we argue, drawing on space syntax research which has sought to link the detailed spatial morphology of cities to observable functional regularities, that all cities, the organic as well as the geometric, are pervasively ordered by geometric intuition, so that neither the forms of the cities nor their functioning can be understood without insight into their distinctive and pervasive emergent geometrical forms. The city is often said to be the creation of economic and social processes, but here it is argued that these processes operate within an envelope of geometric possibility defined by human minds in its interaction with spatial laws that govern the relations between objects and spaces in the ambient world.

Introduction: spatial cognition and the city

What can we learn of the human mind by examining its products? This paper examines the potential of space syntax to throw light on spatial cognition as an instance of the question. At the urban level, space syntax is a theory of description for urban space that is testable by observation of morphological and functional patterns. To the extent that its descriptions correlate with observable regularities in the spatial morphology and functioning of cities, the theory will be successful. However, both the morphological and functional patterns that we find in cities involve the agency of human minds, so the question arises: to what extent can we learn about human minds by examining these patterns? Here we examine key aspects of space syntax research about the city from a cognitive point of view, and propose two theoretical conclusions. First that through an examination of the geometric form of the city we can see evidence of the role played by geometric intuitions in its construction. Second, that the relation between human minds and the physical and spatial city is mediated by spatial laws which are intuitively known to people in the same sense that we intuit physical laws in everyday behaviour. It is argued than that the city is a prime example of the imposition of a mental geometric order on space and a general mechanism by which this occurs is proposed.
The ideal and the organic

We begin with some observations about the city and its geometry. The most basic distinction we make about the form of cities is between the ideal and the organic (for example Morris 1972, Kossof 1991, Batty & Longley 1994). The ideal are geometric, the organic are not – or seem not to be. The geometric we define in terms of straight lines and 90 or 45 degree angles, the organic in terms of the lack of either. Figure 1 The ideal seem to be top-down impositions of the human mind, the outcome of reason, often in association with power. We easily grasp their patterns when seen ‘all at once’. The organic we take to be the outcome of unplanned bottom up processes reflecting the practicalities of everyday living rather than the ordering of human minds. We cannot easily grasp their patterns when we see them ‘all at once’ from above. But, curiously, when we walk about in them, and so see them a bit at a time, the very differentiation of their parts can make them easier to navigate than patterns where parts tend to be similar and in similar relations. From inside, we often find the organic easier to understand, from outside the geometric.

In fact, the relations between ‘organic’ and ‘geometrical’ are much more complex than any simple typology. Most cities combine the organic and the geometric in different phases of growth. Rome, for example, like many cities including London, began organic and grew more geometric as it expanded. Tokyo, again like many cities, began geometric and grew organic. We also find cities with bewildering juxtapositions of differently orientated and differently shaped grids, a kind of organic mix of geometric grids. Cities often enclose patches of organic settlement as they grow, and of course we also find geometric interventions in organic grids.

Cities then seem hugely different in the way in which their grids are put together. It is something of a surprise then to discover that in spite of these differences pretty well all cities share certain common geometric properties. To understand this, we first need to represent urban grids in a consistent way. What we have been looking at so far are the least line maps of cities developed by space syntax (Hillier & Hanson 1984) which are probably the simplest consistent representations of urban grids. These can in small scale cases be created algorithmically by using the UCL DepthMap software (Turner 2002, Turner, Penn & Hillier 2005) but for large scale urban systems this is computationally prohibitive, so least line maps are commonly digitised using the rules for creating and checking maps set out in (Hillier & Penn 2004)

Some consistencies

With least line maps for real cities we find some remarkable consistencies. Most strikingly, at all scales, from the local area to the whole city, and for both ‘organic’ and ‘geometric’ cities, we find least line maps are made up of a very small number of long lines and a very large number of short lines, so much so that in terms of the line length distributions in their least line maps cities have scale-free properties [Hillier 2002, Carvalho & Penn 2004]. Practically speaking, we also find that wherever we are, we are not far from a line much longer than the one we are on. Formally, it means that these often haphazard urban growths have acquired
some mathematical structure. This poses a puzzle. How can mathematically well-formed networks emerge from decades or centuries of activity by innumerable uncoordinated agents acting in very different social, economic and cultural situations and working with very different, and highly variable, geometries?

But this is not all we find. If we look carefully at organic grids, we begin to see geometrical consistencies. For example, if we look at the section of Tokyo in Figure 1 the eye intuitively picks out line continuities, formed by series of nearly straight end-line to end-line connections. If we move along one of these lines, another is very likely to be found at the end of the line, and then another. This happens at all scales, but at each scale the lines are locally longer than lines which lack this kind of angular connection. It can be said that, probabilistically, the longer the line, the more likely it is to end in a nearly straight connection to another line.

The eye will also note a large number of shorter lines, often with near right angle connections, forming more grid-like local patterns. Again if one such line is found, then it is likely that there will be several others in the immediate neighbourhood. We can also say the shorter the line, the more likely it is to end or intersect in a right angle or near right angle. These are the opposite properties to those we find in highly formal cities, like Brasilia or pre-Columbian Teotihuacan, where the longest lines end at right angles on the most important buildings.

So organic grids tend to have a kind of informal geometry. They are more regular than they appear at first sight. There is, in effect, a hidden geometry in organic cities: they are quite grid like, in spite of seeming irregular (Hillier 1999). We can call them deformed grids. At the same time, geometric grids are not so regular. Lines are of very different lengths and connectivities, because many are interrupted, either by buildings or other artefacts. We can perhaps speak of two kinds of grid: deformed grids and interrupted grids, as in the sections of Chicago and Tokyo shown in Figure 1 (Hillier 1996).

Where then do the surprising regularities of both types of grid, and especially organic grids, come from? Three types of process identified in space syntax studies seem relevant. The first refers to how spatial patterns emerge from the simple fact of the aggregating buildings in space (Hillier & Hanson 1984) The second refers to the effects on the geometry of the emergent space of how buildings are shaped, scaled and oriented (Hillier 2002). The third refers to the effects of the emergent spatial geometry on movement patterns (Hillier et al 1987, Hillier et al 1993, Hillier & Iida 2005). Each process, it is argued in space syntax, is subject to simple spatial laws which can be formally expressed as well as intuitively understood at an every day level. We can term these the aggregative law, the spatial law, and the movement law. We now take each of these in turn.

An aggregative law

The basic form of all cities is one of discrete groups of contiguous buildings, or ‘blocks’, usually outward facing, defining a network of linear spaces linking the buildings. How can this arise?

![Figure 2: Aggregating dyads of open and closed cells by a restricted random process](image)

If we take cell dyads (Figure 2, top left), representing buildings linked by entrances to a bit of open space, and aggregate them randomly apart from a rule that each dyad joins its bit of
open space cell to one already in the system (forbidding vertex joins for the buildings, since no one joins buildings corner to corner), a pattern of buildings and spaces emerges with the topology of a city - outward facing blocks defining a linking network of linear space - but nothing like its geometry, in spite of being constructed on a regular grid (Hillier & Hanson 1984). The ‘blocks’, and so the spaces, are the wrong shape. Where then does the characteristic urban geometry come from?

**A spatial law**

This brings us to the second process identified in space syntax. To understand this we need first to think a little about the network of space in cities and how we interact with it, and the role that different notions of distance might play. Space in cities is about seeing and moving. We interact with space in cities both through our bodies and our minds. Our bodies interact with the space network through moving about in it, and bodily the city exists for us as a system of metric distances. Our minds interact with the city through seeing. By seeing the city we learn to understand it. This is not just a matter of seeing buildings. We also see space, and the city comes to exist for us also as a visually more or less complex object, with more or less visual steps required to see all parts from all others, and so as a system of visual distances. This warns us that distance in cities might mean more than one thing.

But we also need to reflect on the fact that cities are also collective artefacts which bring together and relate very large collections of people. Their critical spatial properties of cities are not then just about the relation of one part to another, but of all parts to all others. We need a concept of distance which reflects this. We propose that if specific distance means the common notion of distance as the distance, visual or metric, from a to b, that is from an origin to a destination, universal distance means the distance from each origin to all possible destinations in the system, and so from all origins to all destinations (Hillier 1996).

Why does this matter? Because universal distance behaves quite differently from the normal metric and geometric concepts of distance that we use habitually. For example, if, as in Figure 3 we have to place a cell to block direct movement between two cells, the closer we place it to one of the outer cells the less the total distance from each cell to all others will be, because more cell-to-cell trips are direct and do not require deviations around the blocking object.

![Figure 3](image)

*Figure 3 Moving an object between two others from edge to centre increases the sum of distances from all cells to all others*

The same applies to intervisibility from all points to all others (Figure 4). As we move a partition in a line of cells from centre to edge, the total inter-visibility from each cell to all others increases, though of course the total area remains constant.
Both metric and visual effects arise from the simple fact that to measure inter-visibility or inter-accessibility we need to square the numbers of points on either side of the blockage. So all we need to know is that twice the square of a number, $n$, will be a smaller number than $(n - 1)^2 + (n + 1)^2$ and that in general:

$$2n^2 < (n - x)^2 + (n + x)^2$$ (1)

We can call this the ‘squaring law’ for space. It applies when, instead of being interested in, say, the distance from $a$ to $b$, we are interested in the distance, metric or visual, from each point in the system to all others. In space syntax these ‘all to all’ properties are called configurational to distinguish them from simple relational or geometric properties.

So why does this matter? Because how we place and shape physical objects, such as urban blocks, in space, determines the emergent configurational properties of that space. For example, one consequence of the squaring law is that as we move objects from corner to edge and then to central locations in bounded spaces, total inter-visibility in the system decreases, as does visual integration (or universal visual distance) defined as how few visual steps we need to link all points to all others Figure 5 (left) The same applies to metric integration (or metric universal distance) defined as the sum of shortest paths between all pairs of points in the ambient space, which increases as we move the obstacle from corner to centre (right).

Figure 5 Moving an object from corner to centre decreases intervisibility (left – light means less visual distance to all other points, and dark more) and increases the mean length of trips (right – light is less metric distance, and dark more).

The same same squaring law governs the effect of shape (Figure 6): the more we elongate shapes, keeping area constant, the more we decrease inter-visibility and increase trip length in the ambient space. The effect of a long and short boundary is to create greater blockage in the system through the squaring law.

Figure 6 Changing the shape of an object from square to rectangular decreases inter-visibility and increase mean trip length. Again, light means less visual distance (left) and metric distance (right)
Even at this stage, this spatial law has a critical implication for cities: in terms of configurational metrics a short line and a long line are, other things being equal, metrically and visually more efficient in linking the system together than two lines of equal length. (Figure 7), as would be a large space and a small space, compare to two equal spaces.

![Figure 7](image)

*Figure 7* Other things being equal, a short and long line integrate more than two lines of equal length. Again, dark means less visual distance.

Another consequence is for the mean length of trip (or metric integration) from all points to all others in different types of grid, holding ground coverage of blocks, and therefore total travellable distance in the space, constant.

![Figure 8](image)

*Figure 8* Changing the scaling of a grid changes mean trip length. In this case, for graphical clarity, dark means less metric distance from each point to all others. The mean distances for each system are: top left 2.53, top right 2.59, bottom right 2.71, bottom left 2.42

In the four grids in Figure 8, darker (for clarity) means shorter mean trip length to all other points. Compared with the regular orthogonal grid (top left), interference in linearity on the right slightly increases mean trip length. But more strikingly, if we reduce the size of central blocks and compensate by increasing the size of peripheral blocks, we reduce mean trip length compared to the regular grid. This of course is the ‘grid intensification’ that we often note in looking at centres and sub-centres in cities. As so often, we find a mathematical principle underlying an empirical phenomenon.

How we place and shape objects in space then determines the emergent configurational properties of that space. But what kind of block placing and shaping make space urban?

![Figure 9](image)

*Figure 9* Two slightly different arrangements of identical blocks, with strong linear relations between spaces on the left and weak on the right
On the left of Figure 9, we aggregate buildings in an approximately urban way, with linear relations between spaces, so we can see where we are going as well as where we are. On the right we retain the identical blocks but move them slightly to break linear connections between the spaces. If we then analyse metric and visual distances within the two complexes, we find that all to all metric distances (not shown) increases in the right hand case, so trips are on average longer, but the effect is slight compared to the effect on all to all visual distances, which changes dramatically (shown in Figure 10). Showing visual integration – dark mean less visual distance as before – we see that the left case identifies a kind of main street with side and back streets, so an urban type structure has emerged. But the right case has lost both structure and degree of inter-visibility. Even though the changes are minor, it feels like a labyrinth. We can see where we are but not where we might be.

Figure 10 Visual integration analysis (light is high, and so low visual distances from all points to all others) showing how non-urban layout on the loses both integration and structure through the slight block changes

The effect on computer agents moving around the system is striking, if obvious. In Figure 11 we move 10000 computer agents with forward vision in the space, again using the software by Alasdair Turner (Turner 2002). The agents randomly select a target within their field of vision, move 3 pixels in that direction, then stop and repeat the process. On the left, the traces of agent movement ‘find’ the structure of visual integration. On the right, they wander everywhere and tend to get trapped in fatter spaces. This is an effect purely of the configuration, since everything else is identical.

Figure 11 Traces of 10000 forward looking agents moving nearly randomly

But what about human beings? Human beings do not of course move randomly, but purposefully, and successful navigation in an unfamiliar environment would seem to depend on how good a picture of the whole pattern we can get from seeing it from a succession of points within it. One way we might plausibly measure this property is by correlating the size of the visual field we can see from each point with the visual integration value (its visual distance from all others), so in effect measuring the relation between a local property that we can see from each point, and a non-local one that we cannot see (Figure 12).

Figure 12 Intelligibility scattergrams for the two layouts in Figure 9
In space syntax this is called this the *intelligibility* of the system. The $r^2$ for the ‘intelligible’ layout on the left is 0.714 while for the right case it is 0.267. Defined this way, the intelligibility of a spatial network depends almost entirely on its linear structure. Both field studies (Hillier et al 1987) and experiments (Conroy-Dalton 2001) suggest that this does work for humans. For example, Conroy Dalton took a linearised ‘urban’ type network (Figure 13 left below) and asked subjects to navigate in a 3D immersive world from left edge to ‘town square’ and back. As the traces show, they manage to find reasonable routes. But she then moved the (identical) blocks slightly to break the linear structure and reduce intelligibility (Figure 13 right below), and repeated the experiment. The subjects found the modified layout labyrinthine and many wandered all over the system trying to perform the same way-finding task.

![Figure 13 Trace of human agents in an intelligible (left) and unintelligible (right) layout](image)

So if, coming back to our aggregative process, we modify it by requiring those adding cells to the system to avoid blocking a longer local line if they can block a shorter one (Figure 14, left), we find a much more urban type layout emerges approximating the mix of long and short lines we find in real systems and emulating certain structural features (Hillier 2002). With the contrary rule — always block long lines (Figure 14, right) — we construct a labyrinth in which lines are of much more even length. So urban space networks seem to be shaped in some degree by a combination of spatial laws and human agency, with the human agents implementing, and so in a sense knowing, the spatial laws. The consistency we find in urban space patterns suggests that human beings ‘know’ the configurational laws of space in some sense — perhaps in the same sense that they ‘know’ simple ‘intuitive physics’ when they throw a ball of paper so that its parabola leads it to land in a waste paper basket.

![Figure 14 A layout generated by a ‘conserve longer lines’ rule (left) and one generated by the inverse rule](image)

**A movement law**

But this cannot be all. Cities are also shaped by economic and social processes. How do these fit into the picture? This is the third thing we need to know. First we can note that the ‘urban’ type pattern on the left of Figure 14 is *dual* in the sense that it is composed of a
dominant pattern of long lines against a background of areas made up of short lines, approximating what we saw cities to be like. But why this duality? For this we must understand the functional effects of the network, and to understand these we must first learn to analyse the network in terms of what the network is primarily for: that is movement.

Let us first reflect for a moment on human movement. Spatially speaking, every human trip is made up of two elements:

- an origin-destination pair—every trip is from an origin space to a destination space—we can call this the to-movement component;
- the spaces passed through on the way from origin to destination—we can call this the through-movement component.

In fact, both of these potentials can be measured:

- to-movement is about the closeness or accessibility of spaces from all others— in space syntax this is normalised and called the measure of integration: how close is a space to all other spaces (Hillier & Hanson 1984, Sabidussi 1966);
- through-movement is about the propensity of spaces to be passed through on the way from all origins to all destinations—this is the betweenness measure, called choice in space syntax, meaning how likely is a space to be chosen as part of a route between spaces (Hillier et al 1987, Freeman 1977).

Starting from the least line map, we divide each line into its segments (between intersections) and represent the result as a graph (Figure 15). We then assign integration (closeness in mathematical parlance) and choice (betweenness in maths) measures using shortest path (metric), least angle change (geometric), fewest turns (topological) weightings to relations between each segment and all others, and we apply them at different radii from each segment, also defining radii metrically, geometrically and topologically. This yields a matrix of configurational measures which we can use to see if we can find significant structure-function relations. So we can look at each segment in a system in terms of either its to- or through-movement potential, defining distance and radius metrically (shortest paths), geometrically (least angle change paths) or topologically (fewest turns paths) (Hillier & Iida 2005).

Figure 15 How the angular segment map is derived from the least line map

Taking the least line map of the left generated system in Figure 14, we can visualise the pattern of values, which we call the structure of the system, by ‘colouring up’ the network in the usual way for integration (Figure 16 left) and choice (Figure 16 right) without radius restriction. Since integration measures the accessibility of nodes as destinations from origins, then from the principle of distance decay (and other things being equal), we must statistically expect more movement potential for nodes that are closer to all others at some radius. Likewise, since choice measures the sequence of segments we pass through so we must expect a similar bias in real movement. In effect integration measures the to-movement, and choice the through-movement, potential of spaces.
Natural movement

Is there a link to real movement? This will depend, among other things, on the question, much discussed in the cognitive literature, of how people make distance judgements in complex space (for example Winter 2002, Timpf et al 1992, Hochmair & Frank 2002, Conroy-Dalton 2003, Duckham & Kulik 2003, Golledge 1995, Montello 1992, 1997, Sadalla 1980, Duckham, Kulik & Worboys 2003, Kim & Penn 2004) So how? Shortest paths? Fewest turns? Least angle change? (Hillier & Iida 2005) suggest this can be resolved by examining real flows. They applied the three weightings to the two measures to make six different analyses of the same urban system, and correlated the resulting patterns of values for each segment with observed movement flows on that segment (Table 1, 2), arguing that if across cases there were consistently better correlations with one or other weighting, then the only logical explanation would be that this weighting reflects better how people are biasing spatial movement choices, since everything else about the system is identical. In fact, across four separate studies in areas of central London, we consistently found that geometric, or least angle weightings yields the strongest movement prediction, with an average of around 0.7 for vehicular movement and 0.6 for pedestrian, closely followed by the topological or fewest turns weighting. Metric shortest paths are markedly inferior in most cases, and in general, to-movement potentials are slightly stronger than through-movement potentials, though this varies from case to case. (Hillier & Iida 2005)

![Figure 16 Integration, or closeness, (left) and choice, or betweenness, (bottom right) analysis of the least line map of the left generated system of Figure 14](image)

### Tables 1 and 2 showing $r^2$ values for observed movement and spatial values

<table>
<thead>
<tr>
<th>Gales</th>
<th>Measure</th>
<th>Least Length</th>
<th>Least Angle</th>
<th>Fewest Turns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VEHICULAR MOVEMENT</strong> $r^2$ values for correlations between vehicular flows and shortest path, least angle and fewest turns analysis applied to accessibility and choice measures. Best correlations are marked *. Numbers in brackets indicate best radius of segments for accessibility measures.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BURNSBURY</td>
<td>accessibility choice</td>
<td>.31(60)</td>
<td>.579</td>
<td>.679(90)</td>
</tr>
<tr>
<td>CALTHORPE</td>
<td>accessibility choice</td>
<td>.20(65)</td>
<td>.585</td>
<td>.83* (90)</td>
</tr>
<tr>
<td>SOUTH KEN</td>
<td>accessibility choice</td>
<td>.75(93)</td>
<td>.645</td>
<td>.688(24)</td>
</tr>
<tr>
<td>BROMPTON</td>
<td>accessibility choice</td>
<td>.56(61)</td>
<td>.475</td>
<td>.692(33)</td>
</tr>
<tr>
<td><strong>PEDESTRIAN MOVEMENT</strong> $r^2$ values for correlations between pedestrian flows and shortest path, least angle and fewest turns analysis applied to accessibility and choice measures. Best correlations are marked <em>. $</em>$ or $**$ for combined multiple values indicates whether accessibility or choice is dominant. Numbers in brackets indicate best radius of segments for accessibility measures.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BURNSBURY</td>
<td>accessibility choice</td>
<td>.19(57)</td>
<td>.579</td>
<td>.719(18)</td>
</tr>
<tr>
<td>CALTHORPE</td>
<td>accessibility choice</td>
<td>.35(102)</td>
<td>.430</td>
<td>.857(39)</td>
</tr>
<tr>
<td>SOUTH KEN</td>
<td>accessibility choice</td>
<td>.15(87)</td>
<td>.514</td>
<td>.523(21)</td>
</tr>
<tr>
<td>BROMPTON</td>
<td>accessibility choice</td>
<td>.11(81)</td>
<td>.456</td>
<td>.632(53)</td>
</tr>
</tbody>
</table>
This shows the configuration of the network is responsible for a substantial part of movement flows in two senses: the objective to- and through-movement potentials of the network itself contributes what we might call network effects on shaping flows; and these are modified by how human minds contribute distance effects through how they read distance in complex spaces. So we have brought to light two rather remarkable things. The first is that the grid configuration itself is largely responsible for the pattern of movement flows along streets. We call this the theory of natural movement. Second, the way we navigate spatially is guided not by metric distance as has been uncritically assumed, but by geometrical and topological factors.

The spatial form of cities

With this knowledge, then, we have a new tool for investigating the form and functioning of cities. First let us look at the emergent spatial form of cities. Applying the integration, or to-movement, measure to real cities, and using the least angle change, or geometrical, definition of distance, we find some remarkable emergent geometrical patterns, and again we find they are near invariant across different kinds of city. For example, we commonly find a pattern we call a deformed wheel: a hub, spokes and rim forming the main structure of public space, and the residential areas in the interstices of the wheel. This first came to light in the study of small towns in the South of France, and we found the same pattern in London’s urban areas with her ‘urban villages’ at the hub. But it was something of a surprise to find the pattern approximated in very large cities such as London, with a relatively weak rim, and Tokyo, with much stronger, and multiple rims (Figure 17).

Figure 17 Integration analysis of London within the M25 (left) and metropolitan Tokyo (right) showing the underlying approximation of the ‘deformed wheel’, in the Tokyo case with multiple rims

We seem to find this pattern emerging under very different geometric conditions. For example, we find this emergent structure in strongly geometrical Atlanta, and the very ungeometrical old city of Hamedan in Iran (Figure 18).

Figure 18 Integration analysis of Atlanta (left) and Hamedan (right) showing how the deformed wheel pattern emerges under very different geometrical conditions!
When we apply the choice, or through-movement measure, we find a different kind of structure reflecting some of the deformed wheel but more like a network (Figure 19)

Figure 19 Least angle choice (betweenness) analysis of London (left) and Tokyo (right) showing the network pattern in both cases

and again, this seems to work for interrupted grids as well as for deformed grids. We can also combine the two measures by simply multiplying one by the other, to give a combined picture of the to- and through-movement potentials of each street segment in the system with respect to all others.

Applying the restricted radius measures then allows us to capture much more detail of local structure, reflecting the fact that when we make large scale trips in the city we tend to use the global structure, but when we move locally we will often find ourselves prioritising spaces which are not part of global pattern, but which are locally important. We can use the colour range as a kind of microscope to explore these detailed local patterns. For example, in Figure 20, we reduce the radius of the choice measure from n (left) to 1000 metres (right) in an area of north west London, and then use the colour range to zoom in and begin to detect the urban villages, which of course are focuses of local, but not global, movement.

Figure 20 showing radius n (left) and radius 1000m choice analysis of part of north west London

We can use the same technique to detect London’s often surprising pattern of local shopping street and market areas, and to identify local area structures, often in the form of a local deformed wheel. The local deformed wheel is in fact the secret of London’s surprisingly strong local organization and the reason we name it as a system of ‘urban villages’.

But how do these patterns affect the functioning of the city? We already know the emergent structure of the grid reflect and shape movement flows. Does this have further consequences? It does, and by understanding these consequences we arrive at a new theoretical definition of the city. What we find is that the link between the network configuration and movement flows is the key to the dynamics and evolution of the system. Because the network shapes movement, it also over time shapes land use patterns, in that movement-seeking land uses, such as retail, migrate to locations which the network has made movement-rich while others, such as residence, tend to stay at movement-poor locations. This creates multiplier and feedback effects through which the city acquires its universal dual form as a foreground network of linked centres and sub-centres at all scales set into a background network of residential space. This is the space syntax definition of a city. Through its impact on movement, the network has set in train a self-organising processes by which collections of buildings become living cities.
The dual city

We have then found our dual structure, and we can explain it. The foreground structure, the network of linked centres, has emerged to maximise grid-induced movement, driven by micro-economic activity. Micro-economic activity takes a universal spatial form and this type of foreground pattern is a near-universal in self-organised cities. The residential background network is configured to restrain and structure movement in the image of a particular culture, and so tends to be culturally idiosyncratic, often expressed through a different geometry which makes the city as a whole look spatially different. We call the first the generative use of space since it aims to generate co-presence and make new things happen, and the second conservative since it aims to use space to reinforce existing features of society. In effect, the dual structure has arisen through different effects of the same laws governing the emergence of grid structure and its functional effects. In the foreground space is more random, in the background more rule-governed, so with more conceptual intervention.

![Figure 21](image.png) The old city of Nicosia (left) and its integration analysis, showing the deformed wheel core in spite of culturally differentiated residential space.

We can illustrate this most clearly in a city with more than one culture (now unfortunately separated): Nicosia (Figure 21). Top right is the Turkish quarter, bottom left the Greek quarter. Their line geometry is different. In the Turkish quarter, lines are shorter, their angles of incidence have a different range, and there is much less tendency for lines to pass through each other. Syntactically, the Turkish area is much less integrated than the Greek area. We can also show that it is less intelligible, and has less synergy between the local and global aspects of space. Yet in spite of these strong cultural differences in the tissue of space, we still find Nicosia as a whole is held together by a clear deformed wheel structure. This shows how micro-economic activity spatialises itself in a universal way to maximise movement and co-presence, while residence tends to reflect the spatial dimension of a particular culture, and the expression is in the first instance geometrical. Since residence is most of what cities are, this ‘cultural geometry’ tends to dominate our spatial impressions of cities.

A cognitive conjecture: how do we acquire non-local knowledge of the city?

But whatever their geometry, space networks in cities have further unexpected property. Although the form of the system has evolved bottom-up, its functioning is top-down, in that the movement flows which drive the evolution of the system reflect the position of each space in the large scale configuration, not the local properties of the space. In this sense, the properties of spaces which are critical to its functioning are non-local and reflect a large number of remote connections. This poses a very interesting cognitive question. In order to produce the patterns of flows we find, people must be using some kind of non-local internal representation of the space network, with both geometrical and topological properties. Since people see cities only a bit at a time, it seems that they somehow synchronise discrete experiences into a non-local picture. What might this be like?

In space syntax, non-local patterns are visualised through a graphical device we call the justified graph, or j-graph (Hillier & Hanson 1984) We illustrate this through Figure 22 below. Each element in the system — in the case in Figure 22 each ‘street’ - is represented as the root of a graph with all other elements to which it connects aligned in the first layer above,
those that connect to this layer in a second layer, and so on. In this way we can picture, for example, the complexity of routes from each street to all others, a critical non-local property. A graph shallow from the root, called an integrated $j$-graph in space syntax, means little route complexity to all other line, a graph deep from the root much complexity.

Figure 22 A notional grid and two of its justified graphs

The representation people are using for navigation looks as through it might be some kind of justified map of connected lines: a justified line graph, perhaps. What might cognitive science say about this? First it says of course that human beings have egocentric route knowledge and allocentric map knowledge (O'Keefe & Nadel 1978 – but for a recent review see (Nadel & Hardt 2004), and that these are actually dealt with in different parts of the brain. It is the latter we need to solve navigational problems beyond the scope of simple route recall. But cognitive science also says something else very interesting. In memory people routinely correct bends to lines, and near right angles to right angles (Allen 1981). These are instances of what we might call the 'Kantian simplification': people impose more geometry on the situation than it actually has (O'Keefe & Nadel 1978). But in terms of our putative line graph, correcting lines is not so much a geometrical error as a topological simplification. Its effect is to turn two line elements into one and so simplify the line topology. It also eliminates an asymmetric relation by which we must go through this space to get to that one, so making the $j$-graph shallower.

We might then conjecture that going from route knowledge to map knowledge is a matter of going from the $j$-graph (in fact its spanning tree) to the graph. The 'Kantian simplification' may be part of the means by which we pass from an egocentric to an allocentric understanding of space, since it makes the topo-geometric $j$-graph simpler and less asymmetric, and so less complex in two critical ways. The more we straighten lines and correct turns, the more the graph become shallow from its root, symmetric and bipartite. This makes the line graph easier to transform to see other viewpoints, and it is perhaps this transformability that permits the passage from an egocentric to an allocentric model. Could this be how we learn cities?

Further cognitive reflections on the objective city

Two further aspects of this account of cities stand out as raising cognitive challenges. Both reflect the fact that the city and its functioning, as we have descried them, are products of human agency, and inconceivable without the active role of human cognition. The first is that the complex emergent patterns and processes we have described as near-invariants in the form and functioning of cities are constructed from the metric and geometric properties of the system of linear spaces that links the city into a single system in the first place, that is the lengths of lines and the angles of incidence at the intersections between lines. The line-graphs which have been the basis of our analysis are composed essentially of no more that this.

We find, in effect, that the city has a pervasive two level geometry. There is a geometrical consistency in the way in which the relations between line lengths and angles of incidence at the intersection are formed into local patterns to give the local differentiation between the more public and more residential parts of the system. And there is a geometrical consistency in the way in which these patterns are scaled up to give the emergent structure of the whole city. It is these geometrical consistencies that govern the form-function relation in cities and which relate the spatial patterning of the city with the emergent spatial patterning of city life. The city cannot have acquired the pervasive geometricity of its form without the equally
pervasive involvement of human geometric intuition. In this sense, human geometric intuitions seem embedded in the city itself.

Second, we have also described a system in which spatial laws, implemented through human agency, intervene between the patterns of micro-economic and social activity that animate the city, and the emergent spatial form of the city itself. We have no concept for such a human-mediated but law-governed system. But what is clear is that the relation between human activity and space is mediated by spatial laws, and it is only through the lawfulness of space that economic and social activity is able to express itself in space. On reflection, it must be that way.

An unavoidable consequence of this is that cities are in a profound sense cognitive – and so human - objects before they are economic and social objects. Paradoxically, it is because they are so that they can absorb, without strain, massive changes in the patterns of activity that animate them. The relation between the form of the city and its functioning is in effect generic not specific. It is not specific patterns of activity that shape space but the way in which the relation between space and activity is mediated by the need for movement and co-presence. New patterns of activity, like the old, will require the full continuum of spaces, from integrated to segregated, and so will discover how to fit into and perhaps modify slightly an existing urban pattern. This capaciousness of the city for the absorption of new functional patterns comes from the underlying form of the city, that is the generic cognitive city, and its priority over the more specific social and economic city.

A general mechanism: description retrieval?

The signs of minds that we detect in the city, then, suggest the pervasive involvement of both geometric intuitions and spatial laws in both the formation of the city and its functioning. Is there a general mechanism governing this link? Here it is proposed that there is, and that it depends on the proposition that our mental interaction with the spatial world engages abstract relational ideas as well as concrete elements. In general, spatial relations are ideas with think with rather than of. The classic case is the prepositions, like between or beyond all of which embody bundles of relations in abstract form, and we use them routinely to structure our picture of the world. We propose there is a generic mechanism we call description retrieval (Hillier & Hanson 1984) through which we extract abstract information from concrete events and re-emboby it in real time.

Suppose, for example, as in Figure 23, one person builds a house and another person builds a house next to it (left above). In one case subsequent builders ignore the first relation (the top left-right sequence) and create a random pattern, in another (the bottom left-right sequence) other builders ‘get the idea’ of the initial relation and re-emboby it in subsequent actions, and a regular form emerges.

![Figure 23](image)

Figure 23 The generation of simple forms from simple rules

This process is interesting at two levels: that of the process and that of the emergent form. The relation ‘next to’ has an interesting abstract property: it is symmetrical, in that if a is next to b then b is next to a, unlike, for example, above or below, or behind and in front which are asymmetrical. In reproducing the relation, then, builders are reproducing the abstraction, and in this sense emulating a rule following behaviour. We can think of this as a kind of embedded rule: the rule is inherent in the concrete behaviour. Human behaviour is full of embedded rules of this kind, and often what is embedded is an abstract scheme of spatial relations. It is hard to avoid the inference that we routinely interact intuitively with the abstract spatial schemes in
the real world. The emergent forms (on the right) are also interesting. In the regular case, where the rule-following has put similar things in similar relations, we easily retrieve a description of the high level object as a line or perhaps ‘terrace’, but in the random case we do not. In the regular case, it seems we are able to synchronise the discrete objects into an overall shape to create a higher level template for further embodiment. In the other we cannot. In general this seems to be the case.

By following two embedded rules rather than one (Figure 24) we find other cases where local rule following leads to the emergence of other regular global forms, one we would name a courtyard, the other a street. In fact, the basic forms for which we have names can all be seen as products of a local rule following leading to global forms in which similar things are put into similar relations.

But exactly what is it that we recognise at this template level? Is it just the recursion process of locally similar events? Or is the upper level recognition in some sense independent of the lower level constructive process? The cases in Figure 25, in which the same objects are rearranged left to right to allow more global linear connections, provide the answer: the upper level is independent of the lower level. The left figure is no less recursive than the right figure, but we see it only as a recursive local pattern, not as a whole. The centre one we are more or less able to synchronise as a whole because lines have appeared at the level of the whole. With the one on the right, we are far more aware of the global pattern that of any recursive process. We have a kind of global takeover and we retrieve an upper level template.

So the process of synchronisation is the upper level of the description retrieval process seems to require certain formal properties to be satisfied at that level, and will not happen without those properties. We begin to see why the emergent spatial complexity of the city requires a certain emergent geometry if human beings are to interact with it effectively.

References


Hillier, B. (1999) The hidden geometry of deformed grids: or, why space syntax works, when it looks as though it shouldn’t. Environment and Planning B: Planning and Design 26(2) 169–191


Kostof S (1991) The City Shaped Thames and Hudson


