The Kind of Problem a City Is

Luís M. A. Bettencourt

SFI WORKING PAPER: 2013-03-008

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
The kind of problem a city is:
New perspectives on the nature of cities from complex systems theory

Luís M. A. Bettencourt
Santa Fe Institute
1399 Hyde Park Rd, Santa Fe NM 87501, USA.

Cities are not organisms, any more than they are machines, and perhaps even less so. They do not grow or change of themselves, or reproduce or repair themselves. They are not autonomous entities, nor do they run through life cycles, or become infected. […] But it is more difficult, and more important, to see the fundamental ineptness of the metaphor and how it leads us unthinkingly to cut out slums to prevent their “infectious” spread, to search for an optimum size, to block continuous growth, to separate uses, to struggle to maintain greenbelts, to suppress competing centers, to prevent “shapeless sprawl”, and so on. Kevin A. Lynch, Good City Form (1984).

Jane Jacobs’ Challenge

The title of this essay is taken from the challenge posed by Jane Jacobs in her influential book the Death and Life of Great American Cities (1). In the last chapter of her book she provides a conceptual basis for her arguments about urban form by defining cities as problems in organized complexity (2). Ideas about complexity in science were brand new at the time, and Jacobs readily realized their importance for understanding cities through her serendipitous association with Warren Weaver. Weaver, in turn, was fresh from explaining the significance of Shannon’s new theory of information (3) to the world. These were exciting times in science, when a series of new integrative perspectives and concepts were sweeping like wildfire through many disciplines. But it was perhaps too early then to define in more precise terms what kind of complex systems cities really are. Clearer formalizations of social networks took another fifty years to emerge. The spectacular diversity and scope of new urban forms and growth patterns that we observe in cities worldwide today was also still unknowable then.

In my view, the challenge of defining the kind of problem a city is goes well beyond a principled rejection of the urban renewal planning practices of Jane Jacob’s day. It also asks for more than the consideration and adoption of her observations about the nature of urban spaces. Today some of these observations may even appear a little parochial, too close to what she observed in New York’s West Village of the late 1950s. What lives on as a challenge is the creation of new and better re-conceptualizations of cities as complex adaptive systems (to use more modern language) and the development of a more scientifically grounded practice of urban planning guided by such principles.
My main objective here is to address this challenge by discussing the essential character of cities with the benefit of over 50 years of research and of new insights from the emerging science of cities as complex systems.

I will show that taking a perspective of cities as integrated social networks imbedded in space and time, and requiring general properties for their open-ended land use and infrastructure development, provides us with a new unified model of urbanization. Such a perspective is supported quantitatively by an enormous body of empirical research, characterizing thousands of cities across the globe at very different levels of development, at present and in the past (4,5). In terms of urban design, this conceptualization of cities emphasizes the importance of generative models, where local structures remain to be developed by agents possessing particular goals and information, but must also be constrained by the function of the city as a whole, as an open-ended "social reactor".

**Conceptualizations of Cities as Complex Systems**

The realization that cities are complex systems still begs the question of what specific type of system they actually are. Are they strictly analogous to other systems in nature, such as biological organisms? Or ecosystems? Or are they a new kind of complex system? The problem with many of these analogies, made clear by Kevin Lynch’s quote at the beginning of this essay, is that they tend to focus on limited similarities of form rather than function: To make progress we must focus on the primary function of cities in human societies. From that standpoint many of the problems of organizing urban space and using technology will become clear.

Before I provide my answer to this fundamental question, I first take a very brief and very selective tour through the history of ideas in urbanism and planning. I do not intend here to be historically accurate or comprehensive, but simply to highlight a few important perspectives on the nature of cities, in order to discuss them later, in light of recent findings in the science of complex systems.

Cities have been conceptualized in analogy to other systems from the very beginning, at least since the debates between Plato and Aristotle about the nature of human societies. Aristotle in particular presciently noted in his *Politics* (4) that people are the most *political* (=urban) of all animals. What he meant by the greater scope of human sociality in cities, as compared that of other species, has remained hard to understand before recent quantitative studies of urban social networks.

Although urban planning has existed in every civilization that needed to deal with the problems of cities, planning practices from antiquity to the time of the first industrial revolution remained mostly a practical means of organizing vernacular
architecture and informal building practices\(^1\) (7-10). Nevertheless, it is an important lesson to urbanism and planning that cities build before the advent of large scale formal planning remain universally attractive to us today (7). Christopher Alexander attributes this quality of buildings and urban spaces to the fact that they reflect long histories of social and natural events. This requires that the built environment of cities is able to change gradually, in response to social and economic needs and that, in this sense, it remains alive.

The management of cities changed dramatically in the late 1800s when urban planning, enabled by new mass production technologies (such as standardized construction materials and mechanized means of production and transportation), assumed a fundamental new role in the fast urbanizing industrial societies of its time. Urban planning then acquired a much more pervasive aspiration of designing and building the city at all scales. From this time on, two main lines of thought conceptualized the city very differently: Those that view cities as systems subject to optimization, and those that think of cities as open-ended processes, subject to gradual evolution. Metaphors of cities as representations of the cosmos, or as organisms or machines versus e.g. ecosystems or nervous systems (10-19) capture this dialectic.

The “evolutionary” view of cities was initially shaped by the writings of Patrick Geddes, who, as a biologist, saw the city in natural ecological terms and planning as a way to guide its evolution, not determine it (15). Lewis Mumford eloquently expanded this vision, emphasizing the historical and organic character of cities in his monumental *Cities in History* (10). However, he also misunderstood the role of new technologies (especially the car) and of modern city forms in fulfilling urban functions in new ways. More contemporary authors such as Jacobs (1,16) or Peter Hall (20,21) made some of these ideas of organic evolution of urban form more definite and practical, both in the sense of urban history and planning practices. The tension between the nature of spatial and social order in cities remained unsettled though, as exemplified by Mumford’s criticism (22) of Jacobs work.

This tradition of smaller scale, more organic design lives on today in some of the ideas of *New Urbanism, Smart Growth, Green Building*, etc. (23). In somewhat different forms this view of cities is also embraced by advocates of generative urban design (7,24), or by planers motivated by respecting historical traditions and vernacular architecture (8). Nevertheless, even if these models for urban design intuit a deep sense of what was successful in the past, they seem to me to lack a deeper and more functional justification from a more empirical and quantitative view of how cities operate in practice. Perhaps because of this, the problem of enabling urban function under economic and technological change, rather then

---

\(^1\) Some exceptions have always occurred, such as the impressive master plan of Teotihuacan (25), a pre-Columbian city in the Basin of Mexico built some 2000 years ago.
copying form, looms large as a general criticism to some of simplest applications of New Urbanism.

But planning has always been primarily motivated by the necessity of mitigating urban problems, such as poverty, disease, violence, congestion, etc. Out of this need arose another tradition that more directly views cities as a set of problems to be managed and, if possible, to be re-designed altogether. This tradition is intellectually inspired by engineering practices and more or less explicitly thinks of the city as a machine. Its problem solving strategies are largely defined by ideas of control theory. While this conceptualization of urban planning may seem outdated to readers in developed countries today, it is alive and well in fast developing nations, such as China, and is also the intellectual forerunner to Big Data strategies to manage modern cities. Clearly these practices have their place in city management: It is the scope of their use to control the city as a whole, and to constrain its evolution, that is questionable.

Planning as radical urban problem solving has had many manifestations in the history of urbanism from Ebenezer Howard’s Garden Cities (26) to the grand designs of modernist planning made iconic in Le Corbusier’s Ville Radieuse (13). These models of planning provided detailed and executable ways of (re-)designing and (re-)building cities on a large scale. Today we think negatively of the most extreme of these modernist experiments - we are still digging ourselves out of this “crater in our understanding” of cities, to use Koolhaas infamous expression (27). The intuition, gained painfully in the suburbs of Paris or the infamous Pruitt–Igoe Project in St Louis, is that such spaces don’t work well socially. However, they are at once typically dense as buildings and sparse as neighborhoods in ways that might have been more successful in different circumstances. How density at different scales in the city contributes to (or detracts from) successful urbanism remains a huge open question.

Looking back at these conceptual debates on the fundamental nature of cities and their massive practical consequences to the lives of billions of people worldwide we are left with one crucial question: To what extent do we understand cities better today, so that different models of urban design can be weighed on their objective merits?

**A New View of Cities as Complex Systems**

There is no longer much of an excuse to ignore many of the measurable properties of cities. Cities across the globe and through time are now knowable like never before, across many of their dimensions: social, economic, infrastructural and spatial.

The main empirical findings from a new body of work in geography and complex systems (4,5,28-32) can be summarized succinctly in the observation that all cities realize both certain spatial economies of scale as they grow and, simultaneously,
attain general socioeconomic productivity gains (24-26). For example, when we compare two cities of different population sizes in the same urban system (usually a nation) we tend to find that the larger is a little bit denser, and that its volume of infrastructure networks per capita (roads, pipes, cables) is smaller. We also find that the larger city is usually wealthier (as measured by its per capita GDP or wages), more expensive and more productive culturally and technologically, e.g. in terms of numbers of creative endeavors or patents filed. Thus, these measurements confirm general familiar expectations that larger cities are not only more congested and expensive, but also more productive and exciting culturally.

These effects are universal, in the sense that they are observed across cities in many urban systems, from the US to China, and from Brazil to Germany (4,5). They also apply over time, in contemporary urban systems, as well as in the spatial patterns of settlement of Pre-Hispanic Basin of Mexico measured over a 2,000 year period, which must have evolved independently from their old world counterparts (32).

More specifically, one finds that these properties of cities change gradually with city size and do not point to any special city size. With each doubling of city population, on average, a city contains about 10-20% less infrastructure volume per capita. It also displays a 10-20% increase in rates of wealth production, innovation and, if left unchecked, of other, less benign products of human socioeconomic interactions, such as violent crime. While definitions of cities vary, these findings apply to the city as a whole, conceived as mixing populations or unified labor markets. These definitions result in a view of functional cities as metropolitan areas (4,5), rather than their central cores or any other particular administrative divisions, which have no real socioeconomic meaning. For small cities this distinction is less important.

These findings and especially the generality of their applicability and magnitude, from developed to developing cities, are certainly a starting point but they still do not provide a new model of what cities are. To that end I have recently proposed a new theoretical framework (4) that derives these properties quantitatively, as well as many others, from a simple conceptualization of cities as urban social and infrastructural networks co-located in space and time.

The central idea is that cities are first and foremost large social networks. In this sense cities are not just large collections of people, they are agglomerations of social links. Space, time and infrastructure play a fundamental role in enabling social interactions to form and persist, and in allowing them to become open-ended in terms of increased connectivity, and sustainable from the point of view of energy use and human effort.

From this perspective the observed scaling properties of cities with population size are obtained from four simple assumptions:

1) Cities are mixing social networks, where every urbanite on a minimal budget afforded through his/her social connections can in principle connect with any other
person in the city\(^2\). This determines the cost of movement of people, goods and information. A city may be denser because it is larger, everything else being equal, and/or because the relative cost of movement is high.

2) Networks of infrastructure are decentralized, meaning that they are built gradually as the city grows, with amounts of new infrastructure added proportional to typical urban spatial densities\(^3\). This means that there is no single center (no heart to the city) where all traffic must be relayed through (33), and also that the infrastructure network is evolvable gradually as population size increases or technologies change.

3) The individual effort to stay socially connected is independent of city size. This is made possible because 1) and 2) imply that as cities grow they become spatially denser, so that even if distances traveled by individuals stay approximately the same (constant effort) more people and institutions become available for contact. In other words, the city comes to us as it grows. This removes an important objection about larger cities from sociology and social psychology (34-37) and allows cities, in principle, to be open-ended in terms of their size.

4) The socioeconomic outputs of cities, from economic production to innovation but also crime, are proportional to the number of social interactions realized per unit time. These in turn are determined on the average by the density of people in public networks\(^4\).

The mathematically inclined reader can find the translation of these statements into more formal language in Ref. (4), as well as the derivation of several other properties discussed below.

What are the general implications of this conceptualization of cities?

The first is that the view of cities in terms of social networks emphasizes the primary role of expanding connectivity per person and of social inclusion in order for cities to realize their full socioeconomic potential. In fact, cities that for a variety of reasons (violence, segregation, lack of adequate transportation) remain only incipiently connected will typically underperform economically compared to better

\(^2\) This derives a relationship between total land area, \(A\), and population \(N\), of the form \(A(N) = aN^\alpha\), where \(a\) depends on transportation technology and the elasticity exponent \(\alpha = D/(D+H)\), where \(D\) is the dimension of space \((D=2)\) and \(H\) is a fractal dimension measuring how accessible the city is to each individual. In practice, \(\alpha \approx 2/3\).

\(^3\) This means mathematically that \(A_n \sim d^{-1/2}N \simeq A_n N^{5/6}\), where \(A_n\) is the volume of infrastructure (road surface) and \(d=N/A\) is the overall population density.

\(^4\) This means that social outputs \(Y\) (total GDP, number of violent crimes) increase with population size according to a scaling relation of the form \(Y \sim N / A_n \sim N^{7/6}\).
mixing cities. It has been argued that expanding social connectivity (in cities) is indeed the key to development and civic participation (38), for example by increasing opportunities for the division and coordination of labor, and the creation of more complex and sophisticated social and economic organizations (39). It has also been noted that planning that manages to increase connectivity (e.g. by coordinating transportation and housing policies (40)) is typically more successful than initiatives that address these issues separately. Certainly, on a small scale this is a central tenant of models for Smart Growth (41), but what these results emphasize is the need for social integration in huge metropolitan areas over their largest scales, not only at the local level, such as neighborhoods.

There are several important consequences for general land use in cities. First, the price of land rises faster with population size than average incomes. This is the result of per capita increases in both density and economic productivity, so that money spent per unit area and unit time, i.e. land rents, increases on average by 50% with every doubling of city population size! It is this rise in the price of land that mediates, indirectly, many of the spontaneous solutions that reduce per capita energy use and Carbon emissions in larger cities. Cars become expensive to park, and taller buildings become necessary to keep the price of floor space in pace with incomes, thus leading to smaller surface area to volume. This, in turn, presents an opportunity for reducing heating and cooling costs per person. These effects may also create the conditions for public transportation to be a viable alternative to automobiles, even when the price of time is high and fast point-to-point transportation becomes increasingly more desirable. Thus, larger cities may be “greener” paradoxically as an unintended consequence of their more intensive land use and their higher economic productivity. Policies that increase the supply of land per capita or reduce transportation costs (such as urban renewal), while addressing other problems, will tend to create cities that are less dense and that require higher rates of energy consumption in buildings and transportation to perform the same social functions.

This raises the issue of population density. When we compare cities at the same time but across different sizes, we tend to find that larger cities in the same nation are denser. Nevertheless, in the United States and increasingly all over the world, we also find many modern urban forms, and especially many low-density large cities, such as Atlanta or Dallas. Are these lesser cities than the West Village that Jane Jacobs knew, or the walkable towns that Smart Growth planners advocate? The perspective of cities as interaction networks tells us how these urban forms can coexist as the result of variations in transportation technology and levels of economic development. The overall spatial extent of the city is determined by the interplay between interactivity and the relative cost of mobility. Personal preferences for certain types of neighborhood may also play a role, of course. In general, when it is possible to move fast across space, cities tend to become much more diaphanous and are able to spread out spatially while preserving their social connectivity. It is, in fact, the diffusion of fast transportation technologies (modern mass transit, but especially cars and road infrastructure), in developing world cities, that is
promoting rapid increases in their land area, sometimes faster than their population growth (42). This, of course, creates possible vulnerabilities: For example, if the cost of transportation relative to incomes suddenly rises (e.g. because it is tied to oil prices) then cities may not be able to stay socially connected, leading to expected decreases in their socioeconomic production. Ideas for “shrinking” cities that have lost population and economic activity, such as Detroit and other former industrial cities, apply the same ideas in reverse: By reducing the spatial extent of the city, social connectivity can be realized more pervasively and at a smaller cost, with the expectation that such places should “heal” their social fabrics by becoming more reconnected and revitalized.

Another interesting consequence of the theory is that total volume of infrastructure is expected to increase with city size faster than land area. This is subtler mathematical prediction (see footnotes) that follows from the requirements that infrastructure networks are decentralized and are built gradually. In practice, this means that in larger cities infrastructure network volumes become a larger and larger part of space. This is intuitively clear in large and/or developed cities where cables, roads, pipes, etc become a ubiquitous feature of urban landscapes, before they eventually are moved into the third dimension, either above or below ground in order not to crowd out other more fundamental uses. Thus, planning for incommensurate changes in land and infrastructure footprints with city size is necessary to enable growth, especially of large dense cities.

Another aspect of this problem is that we can predict the amount of energy spent in transport processes (not only of people, but also of goods, information, services, etc) necessary to maintain the city connected. Perhaps surprisingly, given spatial and material economies of scale, energy dissipated in transportation increases faster than population and behaves, in fact, just like incomes or measures of innovation. This is also the expected cost of operating (but not, typically, of building) urban infrastructure: it has been measured, e.g. in terms of resistive losses in the power grids of German cities (5). It may also appear in the form of higher congestion in larger cities, which is always worse in the larger arteries (highways) than in small roads (43). This suggests some original solutions for the problem of managing urban infrastructure costs and promoting sustainability in cities. Investing in efficiency improvements that reduce energy dissipation in transportation processes, especially in the wider tracts of urban infrastructure, should be done strategically, as a (nonlinear) function of city size.

In addition, the theory gives us some specific clues about how energy use in transportation may be minimized, as well as some of the limits to this effort. In particular it tells us that the ratio of social outputs to energy transportation losses, a measure of urban efficiency, is independent of city size. This is probably one of the most important reasons why cities appear across a huge diversity of scales5. The

---

5 Cities are scale invariant, meaning that the distribution of city sizes does not contain any specific population number at which the properties of cities change
socioeconomic agglomeration advantages of larger cities are proportional to their dissipative costs. Increases in congestion costs mirror socioeconomic advantages of agglomeration. Cities that work well maximize the difference between net social benefits and infrastructural costs. Such measures provide an overall set of metrics that can guide adaptive urban planning and policy (4).

Thus, I have shown how the form and function of cities can be determined, at least overall, by a set of network principles that more fundamentally captures the kind of problem a city is. These ideas will continue to require empirical testing and further development, but the recent direct detailed measurement of human social interactions through mobile phone technologies (44) and the observation of spatial scaling in urban systems of Pre-Hispanic Mexico analogous to that observed in modern cities lend further support to their generality (32).

Two further problems remain in our quest to understand the importance of cities in people’s lives and in the development of nations. The first is to understand more directly how the structure of human interactions produces effects well known to economist and sociologists, who study, for example, crime or innovation within cities. Some advances were recently made, through the direct measurement of individual connectivity statistics via mobile communications (44), but much more needs to be done. Better knowledge of this kind should also help plan for more successful public spaces (45). The second is the interaction between cities in an urban system. Measures of growth over time, from population to economic development, are properties shared by all cities in an urban system, but how they are established and what determines their magnitude is perhaps one of the most important open problems in economics and urbanism. These questions seem to be now open to new theoretical efforts as data across scales, nations and time establishes a clearer theoretical view of these problems.

Planning for the Future

The entire world is now urbanizing at breakneck speed and, as city growth takes place spontaneously in India, Africa or Latin America, it is following no architectural master plan. Perhaps as much as a billion people worldwide today live in slums and more than that build their own shelter and organize their neighborhoods without the services of architects or urbanists or knowledge of any science of cities. The ongoing great urban transition, which entails the building of more urban infrastructure in the next few decades than over the entire history of our species, is dramatically. The statistics of city sizes in an urban system have been intensely studied since the 1930s. They follow approximately “Zipf’s law”, which states that if we order cities by size \( N \), according to their rank, \( r \), we obtain a statistical rank-size rule of the form \( N(r) = N_0/r \), where \( N_0 \) is the size of the largest city, at \( r=1 \).

6 This is known as Gibrat’s law: (population) growth rates are constant, on average, independent of city size, \( N \).
largely an unplanned and spontaneous process. Against this chaotic but vital urban backdrop, what is the role of urbanism and of a science of cities?

For those involved in planning practice it may look like the complex systems perspective on the spatial city leaves too much unspecified. In particular, it says nothing about some of the elementary choices in planning such as the shapes of streets or neighborhoods, houses and buildings, specific uses of space, zoning, etc.

That planning should leave many of these choices unspecified, to be developed locally by individuals, organizations and communities, is an altogether more radical statement. However, both urban history and fundamental scientific concepts about how complex systems are created and evolve suggests just that (47,48). This is sometimes known and the "planner’s problem". The problem rests on the fundamental properties of information in complex systems made up of an immense number of heterogeneous agents: The planner cannot possibly know in practice all the myriads of ways in which people would like to develop urban spaces over the history of a city. Better choices are usually made by agents with more specific information, adequate to their goals and aspirations, so far as these are constrained not to limit similar choices made by others and their integration across urban scales. Thus, ensuring general constraints, such as those discussed in the previous section, together with basic rules at the local level, such as those inspired by vernacular architecture over many centuries now, or by some forms of new urbanism or generative design, may provide a practical model for planning, especially in cities that are largely being built informally anyway. By helping create the city from the general to the particular, through designs that are generative of the whole but not prescriptive of the small parts, planning for the city as a complex system may finally fulfill Geddes’ aspiration "to undo as little as possible, while planning to increase the well-being of the people at all levels, from the humblest to the highest."

I hope to have demonstrated that cities are natural systems that evolve spontaneously in human societies under very general circumstances, whenever there are open-ended advantages to human sociality across scales. In this sense they are as natural as beehives or coral reefs, and should not be thought of as arbitrary human artifacts to be redesigned at will. At the same time, they are a different kind of complex system from other more limited forms of social organization in nature and can be, in fact, immensely more complex in the many forms of information they can embody and generate. Cities reveal at once the best and the worst aspects of humanity, in terms of our creativity and imagination but also our tendencies for violence or discrimination. Because of this enormous potential for human development cities should not be seen as systems to be controlled or resisted, but encouraged to evolve spontaneously in the direction of achieving the best open-ended expressions of our collective nature.

That then is our challenge. We are living the last few decades of the great urban transition and finally fulfilling our global potential as the most social of all species to create something altogether new in Earth’s history. We have within sight age-old
human aspirations, such as to eliminate extreme poverty, to end most injustice, to gain access to good health for all, and to do all that sustainably, in balance with the Earth’s biosphere. All this will have to happen in cities and it can now happen very quickly. Bigger data and a more scientific approach to cities will certainly help. But the ultimate challenge for all of us involved in influencing and practicing urban planning is to translate, apply and further develop these new ideas to promote types of urban environments that can encourage and nurture the full potential of our social creativity, targeted at sustainable and open-ended human development.

Acknowledgements
I am especially indebted to Nicholas de Monchaux for several discussions on the history of planning and urbanism and their relation to complex systems concepts that helped me frame the ideas developed here. I also thank Michael Mehaffy and Dietmar Offenhuber for comments and suggestions on earlier versions of the manuscript. This research was partially supported by the Rockefeller Foundation, the James S. McDonnell Foundation (grant no. 220020195), the National Science Foundation (grant no. 103522), the John Templeton Foundation (grant no. 15705), and by a gift from the Bryan J. and June B. Zwan Foundation.

References

5.  Santa Fe Institute Working Paper 12-00-014. Available online at: http://www.santafe.edu/media/workingpapers/12-09-014.pdf
7.  Aristotle, Politics (Book 1, II).
8.  C. Alexander, The timeless way of building
10.  Michael Smith
26. Teotihuacan plan
35. G. Simmel Die Grosstädte und das Geistesleben [The Metropolis and Mental Life], (Petermann, Dresden, 1903).


