Chapter 2
Introduction to Urban Science

Michael Batty

Abstract This introduction outlines a portfolio of theory and methods in the chapters that develop a basic urban science for urban informatics. Inductive and deductive methods for generating data, analytics, and urban simulation, form the focus. In this first Part of the book, the emphasis is on mobility, space-time theory, energy and infrastructure, the spatial economy, and the role of modelling in understanding and planning the smart city.

There are many different but related disciplinary perspectives underpinning urban informatics, and each of these brings a different science to bear on the tools and techniques which form the core of this new domain. In this introduction, we will not sketch all of these different approaches, for many of these will be developed throughout this book. Here, we will simply outline some of the basic physical theories that pertain to the structure of cities, in particular how the form of the city and its functions influence the location of different activities and the ways in which these activities are linked together. We call this "urban science," which is a little more comprehensive than particular sciences relevant to cities, which relate to ecology, energy, social structure, economic development, and so on, and which develop theories and concepts of these particular subsystems in greater depth. Urban science deals with generic theories of how cities are structured and how they grow and evolve in time, how they change qualitatively with respect to growth, and how their populations organize themselves in space. These features often reveal the kinds of problems that urban planning is designed to alleviate, and in this context, the ways in which urban informatics might progress physical planning can be rooted in some of the theories and principles which urban science is able to elucidate.

Like any science, urban science articulates relationships that define the components of the city using quantitative methods which are generally validated by observations that are drawn from actual cities. In short, the conventional scientific method is key to developing the best tools and techniques that comprise urban informatics. The

M. Batty (✉)
Centre for Advanced Spatial Analysis, University College London, London, UK
e-mail: m.batty@ucl.ac.uk

tool set that is evolving rapidly is based on the classic distinctions between methods that are used to infer order and pattern in data drawn from the city, as well as testing hypotheses that are framed about this order and pattern with respect to data about the city. In short, these tools are based on generating theory through induction or testing theory through deduction. The scientific method usually involves both induction that generates ideas, often alongside deductions from these ideas which in turn are tested. The loop that defines this method is continuous as new ideas are evolved, improved, or discarded, revealing whether or not they are fit for purpose. But at any point in this cycle, these theories need to be translated into forms that are useful in applying the methods of urban informatics. Indeed, the first substantive chapter by Daniel Zünd and Luis Bettencourt illustrates how we can capture data in real time from various objects in the city and by using machine learning, can generate patterns that define how the form of the city can be interpreted. In a later chapter, Shih Lung Shaw illustrates how a series of models about the dynamics of the city can be defined in terms of how the city changes in space and time, with the models then validated in classic deductive terms. Thus, induction and deduction are both brought to bear on the development of urban informatics.

This entire area is dominated by many new methods emanating from computer science, which in turn have developed as computers have scaled down to the point where we can use them to sense any movement and change in the built environment. These sensors may be fixed or mobile, but they have given rise to new data sets that measure how different components in the city change through time. This has led to very large data volumes that tend to produce highly unstructured data that we can only interpret using new methods of pattern recognition and statistical analysis that search for pattern and order in the data. These data are often called 'big' in that they pertain to individual movements and decisions in real time and are only bounded by the time the sensors are active. In this way, data streams can be continuous, and if they grow to terabyte or petabyte levels, we need new and different techniques to explore them, that is, to find the pattern in such data. This is in stark contrast to traditional data sets in cities that usually do have structure, as they are collected in one-off fashion through interview or census. The focus in this book on techniques that involve machine learning and data search has emerged primarily from the need to find structure in data that in their raw form are often completely unstructured. At the same time, increasing amounts of data which might become big can be fashioned from individuals generating their own data either individually or through crowdsourcing. Crowdsourcing has always been used to collect some data, but the existence of new information technologies to support such sourcing has given a new momentum to this kind of data collection.

The elements of urban science that the chapters in this first part of the book address deal with urban morphology, which defines the form and function of the city in terms of location and interactions. Morphology is developed in terms of a threefold characterization of the size, scale, and shape of the city, and much of urban informatics addresses ways in which we might improve the city by changing and manipulating these dimensions. Mobility is the generic area that has grown to encompass the relations between locations and interactions, and this immediately

raises the role of networks at different hierarchical levels in the city, as well as the flows that are directed by these networks. Transportation modeling encompasses the best-developed set of tools in this domain, and many of the chapters here allude to such modeling. The relations that bind all these ideas together and are the essence of urban science are scaling, which formalizes the way the hierarchy of elements of different sizes and scales, such as neighborhoods and districts, function within the city. The classic signature of such scaling is the power law, which is ubiquitous as a measure of nonlinearity in urban systems; and in the next chapter, these ideas are spelt out in more detail. In absorbing the contents of this book, readers will find that they emerge in many different guises.

With respect to what follows in this first part, Daniel Zünd and Luís Bettencourt illustrate how it is possible to sense the most obvious objects in a small town in the Galapagos Islands using a blanket coverage and street-view-like cameras. This produces data that can be mined for the more abstract morphology of the place, showing how a judicious mix of user-generated content can be used to sense the spatial structure of the town. Shih Lung Shaw then provides a detailed review of different dynamic models of cities based on urban systems dynamics, cellular automata, and agent-based simulations, setting this in the wider context of human dynamics at the individual person level, and space-time theory as originally developed by Torsten Hägerstrand. The use of new technologies in unpacking individual movements is explored by Martin Raubal, Dominik Bucher, and Henry Martin, who show how personalized tracking can be scaled to look more generally at mobile decision-making, complementing the two previous chapters, with the focus very much on urban dynamics, spatial structure, and individual mobility.

The argument then changes direction. Sybil Derrible, Lynette Cheah, Mohit Arora, and Lih Wei Yeow explore urban metabolism that they articulate using input–output relations and flows of energy and materials that define linkages between many different components of the urban system. These models are static in that they simulate flow at a cross section in time, and although the authors provide an example based on Singapore, they illustrate how problematic it is to generalize these kinds of models to embrace the fine spatial scale. Ying Jin then explores a simple spatial econometric model which looks at GDP in Guangzhou province in China, where he uses the classic measure of gravitational potential or accessibility to relate this to the way the urban system functions with respect to innovative economic activities. This has important implications for future planning of industrial development in the region. Helen Couclelis then concludes this part by standing back and speculating on how all these trends in digital modeling at different scales pertain to the planning of future cities, particularly smart cities. This serves as gentle closure to the ideas in this first part of the book, which establishes many of the theoretical concepts to be picked up and operationalized in the chapters that follow.