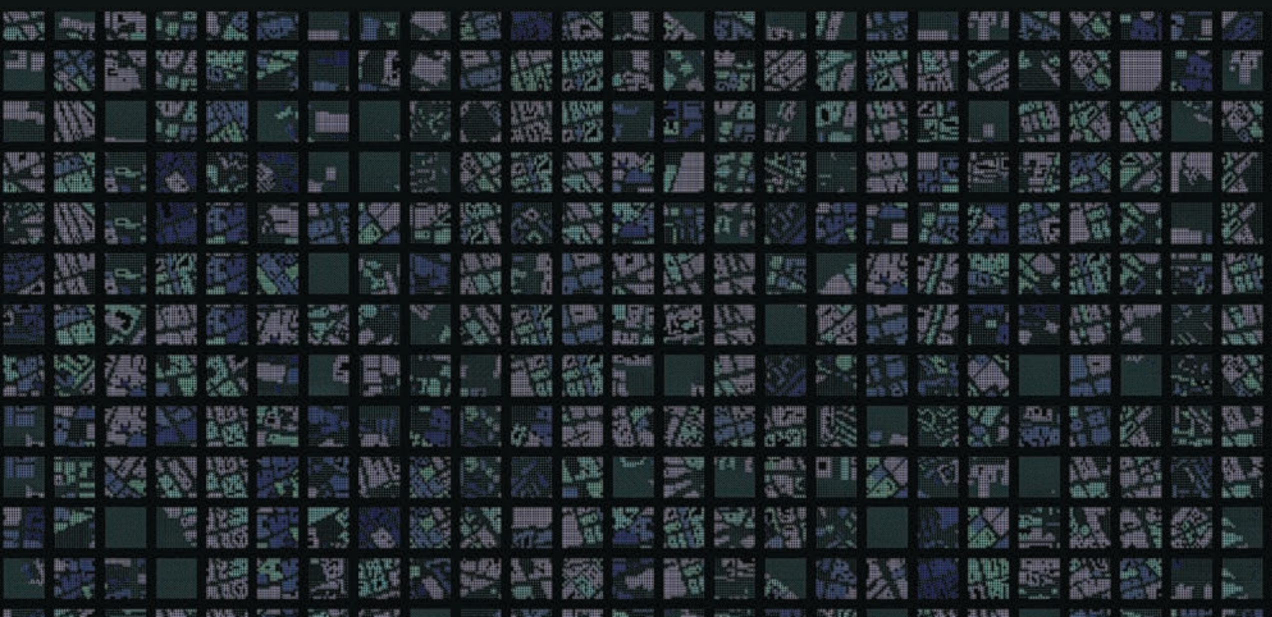
Edited by Silvio Carta

Machine Learning and the City

Applications in Architecture and Urban Design



WILEY Blackwell

Machine Learning and the City

Machine Learning and the City: Applications in Architecture and Urban Design

Edited by Silvio Carta

WILEY Blackwell

This edition first published 2022

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Library of Congress Cataloging-in-Publication Data
Names: Carta, Silvio, editor.
Title: Machine learning and the city: applications in architecture and urban design / edited by Silvio Carta.
Description: Hoboken, NJ: John Wiley & Sons, 2022. | Includes bibliographical
references and index.
Identifiers: LCCN 2021052711 (print) | LCCN 2021052712 (ebook) |
ISBN 9781119749639 (paperback) | ISBN 9781119749585 (pdf) |
ISBN 9781119749622 (epub) | ISBN 9781119815075 (ebook)
Subjects: LCSH: Architecture and technology. | City planning. | Machine
learning. | Artificial intelligence.
Classification: LCC NA2543.T43 C375 2022 (print) | LCC NA2543.T43 (ebook)
| DDC 711/.40285--dc23/eng/20211208
LC record available at https://lccn.loc.gov/2021052711
LC ebook record available at https://lccn.loc.gov/2021052712

Cover Design: Wiley
Cover Image: Neural Networks Public Places.
Authors: Roberto Bottazzi, Dr. Tasos Varoudis, Piyush Prajapati, Xi Wang, Bartlett School of Architecture, UCL, 2020.

Set in 9.5/12.5pt STIXTwoText by Integra Software Services, Pondicherry, India

10 9 8 7 6 5 4 3 2 1

For Emma and Oliver
and, in memoriam, Agnese

Contents

Preface

This book stems from the curiosity I have gradually developed over the past 10 years in data science and computational design, and the growing optimism with which I have been approaching this subject.

I grew up studying engineering, modern architecture, and architectural theory, and I've come to appreciate the intersection of functional architecture with the importance of ideas and ideology as driving forces in design. As someone who has always considered the social and human aspect of architecture before the seductive power of the form, and who has learnt to value the importance of the architecture of the 'guts' over the architecture of the 'brain', I approached formal methods in architecture and, later, computational design with scepticism. However, thanks to several influential colleagues I've had the fortune of meeting during my professional life, I learnt to appreciate the beauty of complexity, order, and logic; the elegance of certain mathematical models; and the satisfying feeling of finding elegant solutions to difficult problems. The more I worked in the field of computational design, the more my scepticism faded, leaving room for what I initially mistook for scientific indifference, but which later became enthusiasm and passion for rigorous logic processes (and computers) later on.

My epiphany was probably when I realised that computation can be used for more than just representing design ideas and the simplification of complicated design tasks. I realised that computation can be used to discover new knowledge, uncover hidden aspects of life, and create new things that can improve people's lives. This may seem obvious to readers with a background in mathematics, statistics, and computer science in general, but as architects and designers, our domain knowledge varies between the humanities (history of architecture or philosophy) and engineering (structural design, health and safety, or building performance). It is not uncommon for people to fall into either extreme of this spectrum, often favouring a nondeterministic view of the world.

I discovered that, by having a greater understanding of data and the techniques to manipulate them in my design, my control over the entire process and the outcome improved significantly. Machine learning (ML) methods (and the data wrangling that underpins them) in particular make the entire design process, from conception to execution, more open, transparent, and logically justifiable at any point.

I like to think of ML (and any other computational method in general) as a good colleague. To make the most of them, you must spend some time with them, trying to understand what they are good at and their limitations. You will soon learn when to ask for their help and for what task. Once you know them well enough, you will be able to reasonably guess how they think and operate, and why they come up with a certain solution. Since you understand your colleague's thought process, one can contextualise their choices and recognise when they are useful to your objectives. Simply put, if you understand them well enough, you can rely on them for those parts of your project that are difficult, laborious, or tedious. One can ask them to do the heavy lifting for you, such as handling extremely complex calculations and to suggest how to move forward when the project reaches an impasse. Your colleague can assist by providing intelligence and granular details as needed, they can find correlations between parts of the project that one has not considered, and they can expose new sides of the problem. Finally, they can generate multiple options and scenarios for you to evaluate and test; they can compare these projections based on relevant criteria and assist one in selecting the best design.

If such a colleague sounds a great asset to your design team, it is crucial to remember the importance of getting to know them well, understanding how they work, think, and the rationale behind their suggestions. Otherwise, ML and artificial intelligence (AI) will remain obscure yet fascinating presences around your work.

This book is intended to assist with this task: getting to know your powerful colleague. I honestly hope you will enjoy it as much as I do.

London, June 2021

Acknowledgements

This work has taken about two and half years to be completed. During this time, I have contacted and interacted with many interesting colleagues, researchers, and designers who generously offered their time, ideas, and contribution. I am extremely grateful to all contributors who kindly sent me their work and helped shape this ambitious project. As this book includes a large number of contributors, I am sure the reader would understand if I do not mention them all here. I also thank those contributors who took a bit longer than I initially hoped for to get back to me. To those, please accept my apologies for my gentle nudges that, in some extreme cases, have bordered on pestering. I am extremely glad that we eventually succeeded in having all contributions in time and in good order.

I would like to thank those who spent their time in reading drafts of certain parts of this work, providing insightful comments and great suggestions. In particular, I am grateful to Daniel Polani, David Leite Viana, Ian W. Owen, and Constantine Sandis.

I would also like to express my gratitude to those colleagues who, for various reasons, did not end up contributing to this book but whose ideas, comments, and suggestions have certainly helped me to refine it: Juan Carlos Augusto, Wassim Jabi, Matthias Scheutz, Alessio Malizia, Güvenç Özel, Charles Petzold, and Jesper Wallgren.

The idea for this book originated during some long and fruitful lunch breaks over the past few years with Alessio Malizia who, without probably realising it, nurtured my interest in data science and their human facets.

I am grateful to the School of Creative Arts at the University of Hertfordshire and, particularly, Steven Adams for the generous (and constant) support with this book and my work in general.

Many colleagues and friends offered their support with various parts of this book; in particular, I would like to thank Tommaso Turchi, who helped with advice in general and LaTeX in particular.

The final stage of this book would not have been possible without the great help of Cosmina Voicila, who ensured that each part (from figures' names to the bibliography) was in the right place and order.

I would like to thank the team at Wiley, who have been extremely supportive and helpful with this work. In particular, commission editor Paul Sayer, who believed in this project since the very beginning and provided useful comments and advice; Viktoria Hartl-Vida, for keeping everything in place and time; Nagavalli Kaipa, who helped greatly with leasing with all contributors; Tim Bettsworth, for the scrupulous and painstaking copy editing; Kanimozhi Ramamurthy for the help in the final stages of production; and Amy Odum, for ensuring my collection of material would actually become a real book.

Finally, my biggest thank-you goes to my family for their great support and patience during these years. They believed in the importance of this project from the very beginning and helped me in many ways to get through it.

Introduction

ML between Routines and Wonder

In recent years, the term machine learning (ML) has been increasingly used in many scientific fields and communication outlets to describe automated processes where computers can make relatively autonomous decisions. On the one hand, there are those who use (and have been using for many years now) ML routinely as part of their job (whether in computer science, finance, physics, astronomy, medical science, statistics etc.). These individuals generally have a solid understanding of both the potentialities and limits of this form of automation. The users of ML chiefly employ it as one of many tools in their skillset to solve the given problems (for example, genome sequencing, predicting market shares etc.). On the other hand, there are people who have a limited understanding of what ML actually is and how it works. For these people, an automation process is where a machine is able to make decisions independently. This often generates wonder as well as concern. ML can be associated with an incredible technological advancement, where tedious, repetitive or very complex human tasks can be assigned to computers that can easily resolve them with a level of speed and efficiency that humans can probably never achieve. When decisions are devolved to programmed automated systems (i.e. computers), it is almost logical to consider the possible implications of such choices. Computers have no responsibility (certainly not in social, ethical or legal terms), no ideological (or political) intent, no regrets and no guilt. From this perspective, it is not difficult to associate deciding machines with dystopian narratives where computers will eventually replace human jobs, making ruthless decisions without considering the social and human context that we normally include in such processes and, in the most pessimistic versions of this narrative, machines replacing humans as they will be deemed somewhat obsolete.

Between these two extreme positions (the daily users of ML and non-experts), lies probably the majority of people who use ML every day without even realising it and who have a general interest in intelligent systems with no real strong opinion about it. One of the underlining arguments of this book is in fact that ML has permeated many of our daily activities and routines already. Some are quite evident (think of the prediction methods used in stock exchanges around the world) and some others are more subtle, running in the background of our lives (for example, the standard predictive text that we have on our phones or the spam filters in our email software).

This book has been designed with this large category in mind, specifically people who would like to know more about ML, to understand the mechanisms by which it works and, hopefully, to take a more proactive approach towards it. The goal is to provide the readers with a robust set of cultural, theoretical and technical coordinates to enable them to be able to understand and contextualise ML approaches, and –in a more active stance—to perhaps start using ML in their work.

Why machine learning?

Artificial intelligence is a large area of research where scientists are trying to design intelligent systems that are able to make decisions independently. In order to decide autonomously, a machine needs to learn and create new links between the data through inference, association, correlation and classification. Computers need to learn from the existing data in order to predict new data and to be able to compute the best solution when more options are available.

Machine learning is the combination of several models and techniques that are based on probability theory, statistics, data science and, on a higher level, applied mathematics. It is these domains that make the computers' decisions possible. Given enough training data, a computer can learn to recognise patterns in the numbers and shapes, as well as in the resulting trends and behaviours. Unlike data mining, where patterns are discovered in the existing data, through machine learning, new information can be both found and predicted.

By having a clearer and deeper understanding of such models and techniques, designers can directly see what determines a choice in an intelligent system. This becomes particularly true in the case of complex systems like cities. The management and control of cities are increasingly characterised by a vast infrastructure of interconnected devices. This requires architects, urban designers and planners to use and design tools that are intelligent in order to handle the growing urban complexity. Machine learning can be increasingly considered the backbone of urban intelligent systems (any system controlled by interconnected machines where a degree of AI is present).

A few initial points

Before going into the details present in this book, it is important to establish a few anchor points that will help to contextualise and understand the role and importance of ML in its multiple applications.

ML is not new. As mentioned several times in this book, the term machine learning was popularised by the computer scientist Arthur Samuel at the end of the 1950s as a way to programme computers to self-learn (in one of the very first applications, this was in order to play checkers). Today ML is widely still associated with computer programming, data science and artificial intelligence (AI). However, it is important to consider that most of the principles underpinning the ML techniques that we use today have their origin in mathematics and statistics. Algorithms like linear regression (one of the simplest ML methods in any textbook) is a direct application of a statistics principle. Put simply, most of the algorithms used in ML are the direct application of statistical methods (Wilmott's Chapter 10 clarifies this point). Samuel may have initiated ML as we know it today, but some of the methods used have a significantly longer history in mathematics and statistics. For example, linear regressions have been in use since the early nineteenth century by either Carl Friedrich Gauss or Adrien-Marie Legendre.

ML is about correlations. As we will see in the chapters that follow, computers are able to learn by finding relationships in the data. At the most basic level, a learning algorithm is a function through which a computer can calculate (and therefore predict) the value of a variable (or multiple variables) based on the given conditions. This is the case of, for example, a linear regression, where the function needs to determine a line that describes how the data points are distributed in a given space. Another basic example is a simple classification problem, where the function that we evaluate is a line (or curve or surface) that is able to separate points (in a two, three or n-dimensional space) more or less accurately in order to determine the groups following certain assigned criteria (more details provided in Chapters 9, 10 and 11). On a more sophisticated level, a ML algorithm is able to infer the rules that underpin the behaviour of a given phenomenon. This applies, for example, to neural networks. A helpful way to consider correlations is to remember the Data-Information-Knowledge-Wisdom (DIKW) pyramid. In this diagram, the learning occurs in the passage from the information (data organised in a meaningful manner) to knowledge (synthesised information that generates new ideas, concepts etc.).

Garbage In, Garbage Out. As new knowledge (learning) happens as the consequence of algorithms finding correlations in the data, it is important to stress the relevance of the initial data used in the learning process. The idea that the input data play a key role in the type of output that is obtained has been highlighted since the time of Charles Babbage: "[...] *if you put into the machine wrong figures, will the right answers come out?*" (Babbage 1864, p. 67). This idea will help when considering questions about algorithmic bias and unfairness, specifically in cases where the inputted data is not scrutinised and the bias is generally attributed to the way in which the algorithm has been designed. Another important facet of this question is the fact that humans need to prepare the input data in such a way that a computer can process them. As we will see in the following chapters, this puts significant stress on programmers (in a general sense, whoever uses a ML method) where the real phenomenon as perceived by humans need to be represented by data and their attributes (i.e. measurable properties) that are machine-readable.

Description and prediction. The data used in ML can be thought of as an abstraction of a given real phenomenon. The first challenge is to find the correct way to represent a phenomenon (for example, through a set of attributes and properties that will result useful in the subsequent stages of the learning process). We call this the analytical phase, where reality is represented by a number of selected elements. The second challenge is to be able to describe the phenomenon through a model, specifically a representation of the reality that is adequate for the expected results. This is usually called modelling. Once the reality is represented through an artificial system (the model), new data is inputted and the algorithm will process the new information. The outcome of this computation is a prediction of how reality will be with the new data. In short, the algorithm learns from a real scenario to predict a new one if and when certain conditions vary. In most cases, prediction is the main reason why ML methods are developed and applied. Accuracy in prediction is one of the measures used to evaluate how efficient algorithms and models are. The final step in this process is the prescription of how parameters should be changed in order to obtain the desired result.

Prediction Vs Generation. Machine learning approaches can be characterised and classified in many ways depending on their architecture, the data structure that they use, the algorithms that underpin them, the methods they use to classify and discriminate data etc. Some of the basic categorisations of the most common methods are presented in this book (Chapters 10 to 12). However, we would like to point out the important distinctions between the methods covered in this book. As mentioned, one of the key principles underpinning the learning of machines is the ability of an algorithm to divide data into groups (classes, categories, clusters etc.). In very simple terms, these algorithms establish a line (or curve or surface) among the points of a dataset, generating different groups. This line is called the decision boundary. In other words, these are discriminative models, where machines learn how to separate new data based on the experience (training) of a given initial dataset. As an alternative to discriminative models, there are generative models, where –oversimplifying– algorithms learn how the data are generated in probability terms. Because of that, generative models are able to create new data based on the joint probability distribution of the dataset (or, more precisely, new configurations of the data). Among the many examples contained in this book, one of the most promising techniques is perhaps Generative Adversarial Networks (GANs). The progress made by many researchers (Goodfellow et al. 2014, Karras et al. 2017 among others) in developing this architecture and related methods is particularly promising for architecture and design. Prediction techniques are very useful for designers to use as they can help them analyse and understand how cities (or architecture in general) perform and how they can be improved. However, generative models can also be used as part of the creative design process, alongside many other traditional skills and tools.

Not all training is learning. The fact that an algorithm can apply probabilistic methods to predict the distribution of a certain dataset does not necessarily imply that the outcome will be knowledge or even useful in any way. One of the classic examples is the case of overfitting. In over-simplified terms, this happens when the accuracy with which we look at an initial dataset (used for training) is too high and the algorithm is "too precise" when classifying the training set. It "too well". When the same model is then applied to a new dataset (representing a new case, a new possible scenario etc.), the algorithm is not able to predict the distribution of the new data. The same applies when the generalisation applied to the initial dataset is too high, as the model does not have enough information to be able to compute a satisfactory prediction using the new data. This is called underfitting. A working ML model should be able to predict a new data distribution by considering a sufficient level of generalisation, despite the existence of incomplete data in the training set (that is usually significantly smaller than the new dataset) and the noise that comes with it. In other words, a good model should have a good "fit". This is a measure that indicates how efficiently the model is able to approximate a target function (i.e. the method needed to solve the initial problem given). Adding an extra level of sophistication, in the case of neural networks the model needs to predict how the data pass through the different layers and which neuron (node in the network) is connected to the others. In the case of neural networks (NN), in addition to the training data and the target function, the model needs to also consider the precision of the outcome (prediction). This level of accuracy is computed through a metric called the loss function. The more the loss function is minimised, the better the predictions (the model makes fewer errors in the prediction) and the fitter the model. The loss function is usually represented by a curve that, in a good model, tends to flatten. This means that the model achieved good prediction results and is learning correctly. There are many variables that determine the success of the model. One of them is the learning rate at which the model makes predictions, which is the speed at which the model learns). If the learning rate is too high, then the model has not had enough time to learn. If it is too low, then the model will take a long time to compute useful results.

Accountability and responsibility. Once the ML model works in a satisfactory way, it predicts new knowledge that can be helpful to researchers and designers, enabling them to better understand existing phenomena as well as to imagine new scenarios. These predictions are as good as the data that has been inputted into the model and the design of the architecture of the model as a whole. A helpful analogy is to think of a ML model as a car. This can be more or less powerful, spacious and fast but this does not relate to the start and the end of the journey that people make with it. As cars have no bearing on how and why people use them, so do ML models and algorithms have no accountability as to what predictions are made as the output. This responsibility is with the people who design them, use them and with those who have generated, selected and prepared the data used. There are a growing number of studies that have sought to help designers and researchers understand the importance of the ethical use of data. We have included some of this key work in the last section of this book to provide you with a solid starting point for your own work.

ML and the City

Cities are probably the most complex and sophisticated manifestations of collective human life. They are the place where a plethora of social and cultural values, needs, ambitions, and a certain degree of freedom converge. Each city can be defined by its own organisational structure and degree of indeterminacy where tensions emerge and balances are struck between the human intention of providing and following rules and the natural human tendency towards flexibility, interpretation and individual expression. One of the main reasons behind the inherited complexity of cities is the diversity that people bring with them when cohabiting a territory. It is exactly this richness that makes cities attractive to an increasing number of people.

The study of urban complexity and the mechanisms that underpin cities is nothing new. Scientific approaches to urban contexts can be traced back to various disciplines from the history of architecture to social studies, and from planning to geography. This latter has probably made the biggest contribution to the development of urban studies with "*quantitative geography and urban modelling, digital mapping and geographic information systems, and in urban cybernetics theory and practice*" as explained by Kitchin (2016, p. 4). In the field of architecture and urban studies, the seminal work of Lionel March (geometry and the spatial organisation of the built environment), George Stiny and James Gips (shape grammar), Bill Hillier and Julienne Hanson (space syntax) and Michael Batty (urban modelling) carried out in the 1970s and 1980s, and, to a lesser extent, Christopher Alexander (1964's on the synthesis of form) comes to mind. Their work paved the way for many of todays' computational and mathematical approaches to architecture including formal methods in architecture (see Leite Viana, Morais and Vieira Vaz 2018), urban informatics (Foth 2008), city information modelling (see Stojanovski 2018), the mathematics of spatial configurations (see Ostwald 2011, Ostwald and Williams 2015, Ostwald and Dawes 2018), sense-able cities (see Ratti 2010), and the connected city (Neal 2012) to name but a few.

These approaches and theories can be considered further developments that go in a sort of linear direction, naturally branching off from the main directions suggested in the 1970s and 1980s. This book suggests that there has been a major breakthrough in this development that has the potential to yield radical changes in the way that we understand and design cities and their complexities. We argue that the introduction of data science, ML approaches and, generalising, AI to urban studies and design can have a significant impact on cities in the coming years. It is true that urban complexities cannot be entirely reduced to data, a process often called datafication (see Cukier and Mayer-Schoenberger 2014). It is more generally known as dataism (see Harari 2016) by journalists and cultural commentators. However, it is also true that many aspects of our lives, certainly those pertaining to urban life, can be sampled, modelled and predicted through the data that represents them. This is a long-established practice in statistics, physics, mathematics, astronomy, engineering and related fields. Within the context of this book, we consider data science and ML to be a plane intersecting all these fields, at least in disciplinary terms.

By looking at the city through the lens of the data that represent its complexity and richness, researchers and designers are in an advantaged position to filter the noise that usually characterises urban questions (anything from subjective perspectives to ideological or political views etc) in order to be able to focus on the essential aspects of an observed phenomenon. We can look at this through the analogy of a tree. The ability to generate models based on abstraction (data) allows the viewer to look past the leaves, flowers and birds resting on the tree (symbolising all the beautiful and interesting aspects of cities) to focus solely on the branches and their structure.

The most important aspect of the application of ML and AI to cities is the fact that filtering through the leaves, the focus on the branches and the abstraction of the essential elements of each phenomenon is done through and by computers. The algorithms that we have designed allow for a certain degree of autonomy when making decisions and drawing conclusions which eventually need to be evaluated. Their meaning needs to be assigned to them by humans. As we can extensively see in this book, computers and the logic by which they operate enhance the work of researchers and designers by offering new and powerful ways of observing cities inclusive of their complexities and indeterminacies. Designers are now offered a new skillset with which they can find correlations (and sometime causations) that were not visible before. They can predict how a scenario may work in the future under certain given conditions, and they can simulate, forecast and anticipate how cities will react to certain changes. In short, we can now scientifically and quantitatively see what we so far have only predicted and imagined using our intuition and qualitative interpretation of cities and urban questions.

Aims of this book

The resources available to designers and researchers who want to learn more about ML with the aim of integrating it into their work can be categorised into two groups. On the one hand, textbooks and technical literature offer information about the statistical and mathematical principles and techniques underpinning the use of ML for general purposes. They may be more oriented to computer science students and thus contain statistical models and relative functions. Alternatively, they may be intended for the general public including people interested in applied mathematics and statistics. The first group addresses the general audience of people interested in mathematics, statistics and computer programming, and it does not generally refer to design, architecture or urban questions. Designers who are interested in ML may find it challenging to appreciate the depth of such specialised resources without a background and/or previous training in data or computer science. On the other hand, there are publications within the field of the built environment, specifically architecture and urban design, where the emphasis is on projects, design tools and the final results. Within these categories, designers may also find relevant resources from city-related subjects including urban geography, urban sociology and urban studies. Most of these have extensively covered the impact of automation and new digital technologies for use in city. Designers accessing resources of this second group will most likely find it interesting and stimulating yet lacking the necessary information to move beyond a general understanding of the subject.

In short, technical publications on ML may be inaccessible without previous training on the subject. Analytical work on the impact of AI and ML on the city does not usually contain clear guidance on how to start working with ML methods.

This book has been designed to support designers and researchers in their access to ML and to provide them with clear references to further their studies and practice in this field.

We hope that, with this book, readers will be able to: i) understand the ideas and techniques underpinning ML, ii) to start using some ML techniques and to be able to read the existing projects at a deeper level (for example, understating the statistical model used and the logic behind a certain application), iii) to have a solid framework of references to further their studies, allowing them to discuss and analyse ML-related topics in other fields (architectural criticism, design history, social and urban studies) with a greater understanding of the subject and finally, iv) to have a greater appreciation of the impact of ML and AI to the contemporary city. This last point is twofold. Firstly, there is the technological perspective whereby this book aims to help readers understand how new technologies underpin smart cities and how informational systems work. Secondly, this work addresses the social dimension of ML. This is where readers will be able to further their understanding of the mechanisms through which ML works. Readers will be able to see how the programmers and designers involved in the process make decisions and assumptions, and how these have an impact on the ways in which the technology works and therefore how people live their lives within the urban context.

More generally, Machine Learning and the City Reader aims to first provide a clear timeline of the development of machine learning techniques and its relationship with AI, robotics and computing in general (Sections 1, 2 and 5). Secondly, this book tries to demystify some of the common ideas that ML and AI are obscure black-box technologies where one inputs data into a computer to gain new knowledge as an output without any degree of control (Sections 3 and 4). By doing this and providing a clear explanation of how ML works when it is applied to an urban scale, this book aims to increase the number of designers and researchers willing to engage with ML and AI in their work.

Structure of the book

This book is organised into 5 sections covering the origins of machine learning, the description of how a machine can think and learn, some of the technical aspects that underpin ML, its application in a city and, finally, the human dimension of ML and its consequences for urban design and the city. Each section couples theoretical and technical contributions written by key scholars in their field with concrete examples and projects by designers who employ ML methods as part of their working routines. The aim for each section is to provide authoritative references with a direct link to their application in the urban context and design in general. This enables any readers to be able to understand the use of ML approaches and their possible results in design using spatial and societal terms.

The first section "**Increasing urban complexity**" suggests a possible starting point for the use of ML and computational methods in general in the city. As the level of complexity of urban questions (from infrastructures to people's cultural diversity) increases and computer-operated technologies become increasingly more pervasive, designers and thinkers have had to integrate new methods into their work in order to be able to better understand the growing complexity. This section describes the gradual increase in complexity through the work of **Sean Hanna** (Bartlett School of Architecture, UCL) where he describes the intelligibility of cities and their urban complexity through patterns and scales. In Chapter 2, **Cassey Lee (**Institute of Southeast Asian Studies -ISEAS, Singapore) explains how these patterns emerge in complex systems and how this emergence links to the notion of universal computation. One of the most common representations of urban patterns (and patterns in complex systems in general) are fractals. In Chapter 3, **Pierre Frankhauser** (University of Franche-Comté) and **Denise Pumain** (University Paris Pantheon-Sorbonne) provide insight int fractals (as an automated way for urban structures to grow) and their development and relevance in spatial practices: "*The irregular and fragmented forms of relief or, urban patterns, the ramifications of hydrographic or transport systems, the hierarchized structures of the world's territories and city systems all have properties, and fractal analysis could propose new interpretations"* (Chapter 3). Two projects are associated with this section. In Project 1, **Ljubomir Jankovic** (University of Hertfordshire) provides an example of the computational methods applied to emergence and urban analysis, while **Nahid Mohajeri** (UCL) and **Agust Gudmundsson** (Royal Holloway University of London) illustrate a method used to analyse the evolution and complexity of urban street networks. The first section lays the important foundation for the syllogistic idea behind this book. The degree of complexity of cities today is increasing exponentially. If everything in nature (and therefore in cities) is computable (Wolfram 2002) and computation is cognition (Scheultz 2002), we can only understand the city today in its complexity through computation.

The second section “**Machines that think**” introduces the key concepts in Artificial Intelligence and Machine Learning. The starting point is **John McCarthy**’s text ‘Artificial Intelligence, Logic and Formalizing Common Sense’ (Chapter 4) where the initial distinctions are drawn between human and machine behaviours. Originally written in 1989, this chapter explains the importance of the relations between artificial intelligence (AI), mathematical logic and the formalisation of common-sense knowledge and reasoning. It also approaches the other problems of concern regarding both AI and philosophy, as well as formalised languages. Following on from this, ‘Defining Artificial Intelligence’ by **David B. Fogel** (Chapter 5) considers AI to be a field of scientific research and human progress. This chapter offers key descriptions and explanations of the methods that allow machines to “*improve themselves by learning from experience and to explain the fundamental theoretical and practical considerations of applying them to problems of machine learning*” (Chapter 5).

Next, **Shelly Fan** describes the passage from the initial enthusiasm for AI during the 1950s and 1960s to the ebbs and flows that have characterised the history of AI from the 1970s to date, including the AI winter (1970s), the 5th generation computer systems (1980s) and the Good Old-Fashioned Artificial Intelligence (1990s) through to the strong reliance on ML techniques in AI systems in the last 2 decades. Thanks to machine learning, AI systems have become increasingly more reliable and precise in their categorisation, clustering, decision-making and predictions. In her text *AI: from copy of human brain to independent learner* (Chapter 6), Fan elaborates on this crucial passage by explaining how AI has moved away from being designed as a copy of the human brain to gradually becoming a system programmed to learn independently. **Keith D. Foote**, in Chapter 7, describes the history of the use of ML in computing and its progress towards becoming AI. In his *The History of Machine Learning and Its Convergent Trajectory towards AI*, Foote describes and comments on several key definitions and moments in the history of ML ranging from what an algorithm is and the Hebb’s Rule to the rise of the computer vision and advancements in Natural Language Processing (NLP). To conclude this part dedicated to the history and development of ML and AI, **Iyad Rahwan** (Massachusetts Institute of Technology, Cambridge, MA, USA and Max Planck Institute for Human Development, Berlin, Germany), **Manuel Cebrian** (MIT), Nick Obradovich (MIT), **Josh Bongard** (University of Vermont, Burlington, VT, USA), **Jean-François Bonnefon** (Toulouse School of Economics (TSM-R), CNRS, Université Toulouse Capitole, Toulouse, France), **Cynthia Breazeal** (MIT), **Jacob W. Crandall** (Brigham Young University, Provo, UT, USA), **Nicholas A. Christakis** (Yale University, New Haven, CT, USA), **Iain D. Couzin** (Max Planck Institute and University of Konstanz, Germany), **Matthew O. Jackson** (Stanford University, Stanford, CA, USA; Canadian Institute for Advanced Research, Toronto, Ontario, Canada and The Sante Fe Institute, Santa Fe, NM, USA), **Nicholas R. Jennings** (Imperial College London, London, UK), **Ece Kamar** (Microsoft Research, Redmond, WA, USA), **Isabel M. Kloumann** (Facebook AI, Facebook Inc, New York, NY, USA), **Hugo Larochelle** (Google Brain, Montreal, Québec, Canada), **David Lazer** (Northeastern University, Boston, MA, USA and Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA), **Richard McElreath** (Max Planck Institute, Leipzig, Germany and University of California, CA, USA), **Alan Mislove** (Northeastern University, Boston, MA, USA), **David C. Parkes**

(Harvard University, Cambridge, MA, USA), **Alex 'Sandy' Pentland** (Massachusetts Institute of Technology, Cambridge, MA, USA), **Margaret E. Roberts** (University of California, San Diego, San Diego, CA, USA), **Azim Shariff** (University of British Columbia, Vancouver, British Columbia, Canada), **Joshua B. Tenenbaum** (Massachusetts Institute of Technology, Cambridge, MA, USA), and **Michael Wellman** (University of Michigan, Ann Arbor, MI, USA) in Chapter 8, *Future development of ML – Machine Behaviour* provide a series of key points to understand the significant impact that ML has on intelligent systems and the Internet of Things (IoT). In this chapter, Rahwan and colleagues focus on the importance of understanding the behaviour of artificial intelligence systems in order for humans to be able to control their actions and behaviour, while "*reap*[ing] *their benefits and minimiz*[ing] *their harms*" (Chapter 8).

Section 2 includes 7 projects that introduce how ML and artificial intelligent systems can be applied in architecture and urban contexts. These include the works of Plan Generation from a Program Graph (Project 3) by **Ao Li, Runjia Tian, Xiaoshi Wang** and **Yueheng Lu** (Harvard GSD), Genetic Algorithms and Care Homes (Project 4) by **Silvio Carta, Tommaso Turchi, Stephanie St. Loe** (University of Hertfordshire) and **Joel Simon**, (Project 5) **Roberto Bottazzi** and **Tasos Varoudis** (Bartlett School of Architecture, UCL)' N2P2 - Neural Networks and Public Places, **Matias del Campo** and **Sandra Manninger** (SPAN)'s Project 6 on Urban Fictions: Lines, Surfaces and Quasi-Intelligent Machines and **Stanislas Chaillou** (Spacemaker AI)'s Latent Typologies. Architecture in Latent Space (Project 7), Enabling Alternative Architectures (Project 8) by **Nate Peters** (Harvard Graduate School of Design) and finally, Distant Readings of Architecture: A Machine View of the City (Project 9) by **Andrew Witt** (Certain Measures) are also included.

The third section "**How machines learn**" is dedicated to the description of the ways in which ML works from the computational, probabilistic, statistic and mathematical viewpoints. In Chapter 9, *What Is Machine Learning?*, **Jason Bell** introduces the key concepts of ML and basic examples of their application. In Chapter 10 *Mathematics for ML*, **Paul Wilmott** explains the key mathematical concepts including Principal Components Analysis (PCA), Maximum Likelihood Estimation (MLE), confusion matrix, cost functions, gradient descent, training, testing, validation and other fundamental notions that are recurrent in ML. In Chapter 11 *Machine Learning for Urban Computing* **Bilgeçağ Aydoğdu and Albert Ali Salah** (Utrecht University) explain how the methods introduced by Wilmott are applied in general terms, and to the city in particular. This chapter provides a second list of the key concepts in ML, including classification, artificial neural networks (ANNs), pattern discovery and clustering, and Bayesian approaches. In his *Autonomous Artificial Intelligent Agents* writing (Chapter 12), **Iaroslav Omelianenko** (NewGround) introduces the idea of Autonomous Artificial Intelligent Agents and how genetic algorithms can be designed and deployed in urban simulations. This section features four projects that illustrate how some of the techniques described can be applied in real design projects. **Sherif Tarabishy, Stamatios Psarras, Marcin Kosicki** and **Martha Tsigkari** (Foster and Partners) (Project 10) present the recent methods for Machine learning for spatial and visual connectivity. **Zhoutong Wang, Qianhui Liang, Fabio Duarte, Fan Zhang, Louis Charron, Lenna Johnsen, Bill Cai** and **Carlo Ratti** (MIT Senseable City Lab) describe their recent project: *Navigating indoor spaces using machine learning: train stations in Paris* (Project 11) This is where they used Deep Convolutional Neural Networks

(DCNN) using photographic images as the input. Project 12 describes the work of **Rolando Armas** (Shinshu University), **Hernán Aguirre** (Shinshu University), **Fabio Daolio** (University of Stirling) and **Kiyoshi Tanaka** (Shinshu University): Evolutionary design optimization of traffic signals applied to Quito city, where evolutionary computation and machine learning methods are applied to analyse transportation systems. Finally, this section includes **Patrik Schumacher**'s (Zaha Hadid Architects) *Constructing Agency: Self-directed Robotic Environments* (Project 13), where architectural and human agents are modelled using a Unity game engine to design the "*densification and transformation of a North London urban district into a creative industry hub via four incubator projects elaborated by four design teams working in parallel and with mutual awareness*" (Project 13).

The fourth section "**Application to the city**" describes how ML can be applied to urban projects in both analytical and design approaches. In Chapter 13, **Martin Dodge** (University of Manchester) and **Rob Kitchin** (National University of Ireland) introduce the notions of code/space and the transduction of space. These are two key concepts elaborated on within human and urban geography that are of relevance when seeking to understand how cities can be analysed in their growing complexity. Partially based on the idea of the transduction of space (Mackenzie 2002), code/space revolves around the analysis of the mutual relationships between software (code) and space (both physical and digital). This notion is considered to be key when it comes to understanding how new models of space and cities can be generated through computing and, more specifically, ML. **Mark Graham** (Oxford Internet Institute), **Matthew Zook** and **Andrew Boulton** (University of Kentucky) continue in Chapter 14 to explain the importance of the virtual aspects of urban spaces. Through an analysis of digital augmentations, they explore the ways in which our everyday lived geographies are changing. In Chapter 15, **Marcus Foth, Fahame Emamjome**, **Peta Mitchell** and **Markus Rittenbruch** (Queensland University of Technology) discuss the importance of urban analytics in: *Spatial Data in Urban Informatics: Contentions of the Software-Sorted City*. They generalise the idea of urban informatics, where intelligent systems, powered by IoT, ML and AI can provide new ways of designing, monitoring and living in the city. The chapters that follow in this section introduce the concrete applications of the approaches and methods explained so far. **Vahid Moosavi** (ETH) in Chapter 16 provides a clear example of how deep learning can be applied at the city level. **Snoweria Zhang** and **Luc Wilson** (KPF Urban Interface) (Chapter 17) present some recent computational approaches developed for large-scale urban projects in Computational Urban Design: Methods and Case Studies. **Diana Alvarez Marin** (ETH) presents her work *Indexical Cities*. Personal city models with data as the infrastructure (Chapter 18), shows how she investigated the methods used to infer the intrinsic characteristics of cities. In Chapter 19, *Machine Learning, Artificial Intelligence, and Urban Assemblages*, **Serjoscha Düring** (Austrian Institute of Technology AIT), **Reinhard Koenig** (*Austrian Institute of Technology AIT, Austria and Bauhaus-University Weimar, Germany*), **Nariddh Khean, Diellza Elshani, Theodoros Galanos, Angelos Chronis** (Austrian Institute of Technology AIT) discuss the importance of computation in data-driven projects and analytics. In particular, they present their recent work on "*generative methods for urban spatial configurations that integrate a number of different simulation engines, along with InFraRed, into one framework* [to] *quickly explore thousands of urban design alternatives by generating a diverse and informative design and performance dataset*"

(Chapter 19). Finally, in Chapter 20 *Machine Learning and Design Fiction* **Franziska Pilling, Haider Ali Akmal, Joseph Lindley** and **Paul Coulton** (Lancaster University) explain how AI and ML methods are used to "explore transparency around human-AI cohabitation in an urban environment" as a part of Lancaster City Council's AI for Lancaster programme. Section III includes ten projects that, taken as a sample, represent the current state-of-the-art regarding the application of ML methods to the city. Project 14 (*A Tale of Many Cities: Universal Patterns in Human Urban Mobility*) by **Anastasios Noulas** (University of Cambridge), **Salvatore Scellato** (University of Cambridge), **Renaud Lambiotte** (Université Catholique de Louvain), **Massimiliano Pontil** (UCL) and **Cecilia Mascolo** (University of Cambridge), and Project 15 by **Gwo-Jiun Horng** (Southern Taiwan University of Science and Technology) (Using Cellular Automata for Parking Recommendations in Smart Environments) illustrate how these techniques are used to address practical urban problems like urban mobility and circulation. A number of projects showcase how neural networks are employed to analyse urban conditions, to generate new urban forms and to transform existing topologies. These include **Sean Wallish**'s (University of British Columbia) *Gan Hadid* (Project 16), **Elizabeth Christoforetti** and **Romy El Sayah**'s (GSD Harvard) *Collective Design for Collective Living* (Project 17), **Erik Swahn**'s (KTH School of Architecture in Stockholm) *Architectural Machine Translation* (Project 18), **Jose Luis García del Castillo y López**'s (Harvard GSD) Project 20: *Style transfer/Boston landscape*, **Benjamin Ennemoser**'s (University of Texas A&M, USA)'s *ML-City* (Project 21), and Project 22 *GAN-Loci* by **Kyle Steinfeld** (UC Berkeley). In Project 19 **Hui Wang** (School of Architecture, Tsinghua University), **Elisabete A Silva** (Department of Land Economy, University of Cambridge) and **Lun Liu** (School of Government, Peking University) demonstrate their method that was used to evaluate large-scale areas and urban street views using deep learning-based models. Finally, **Iacopo Testi** (Rhea Group and Urban AI) presents his project *Urban Forestry Science* (Project 23), where he developed a method based on convolutional neural networks (CNN) to analyse urban forestry in Madrid, Spain.

While the use of ML in intelligent systems may suggest new and promising urban futures, it also includes some inevitable degree of uncontrolled effects for both people and the built environment that they live in. The fifth section **ML and Humans** offers a discussion about some of the less direct aspects that are inherited with any algorithmic approach that require a certain level of awareness and control, specifically when focusing on the human-machine relationship. This discussion starts with the seminal work that Danah Boyd and Kate Crawford carried out asking critical questions about the use of data (see, for example, Boyd and Kate Crawford 2012). Building on their work, in Chapter 21 *Ten Simple Rules for Responsible Big Data Research*, **Matthew Zook** (University of Kentucky), **Solon Barocas** (Microsoft Research), **Danah Boyd** (Microsoft Research), **Kate Crawford** (Microsoft Research), **Emily Keller** (Data & Society, New York), **Seeta Peña Gangadharan** (London School of Economics), **Alyssa Goodman** (Harvard University), **Rachelle Hollander** (National Academy of Engineering, Washington), **Barbara A. Koenig** (University of California-San Francisco), **Jacob Metcalf** (Ethical Resolve, Santa Cruz), **Arvind Narayanan** (Princeton University), **Alondra Nelson** (Columbia University) **and Frank Pasquale** (University of Maryland) provide a set of recommendations that can be helpful to any researcher and designer working with data and people. In Chapter 22 *A Unified*

Framework of Five Principles for AI in Society, **Luciano Floridi** (University of Oxford) and **Josh Cowls** (Alan Turing Institute) provide a set of ethical principles used for the adoption of AI (and ML) that all designers and researchers should consider in their work. If the conscious use of data is the key to improving the quality of our projects, it is equally important to be aware of the consequences of potential digital inequalities. **Matthew T. McCarthy** (University of Wisconsin-Milwaukee) explores the notion of algorithmic divide in Chapter 23 by highlighting possible "*complications relating to identity, social sorting, use, agency, and global development that are inextricably related to the issues above and to the study of big data*" (Chapter 23). Having set out some of the useful principles that we should all consider when working with data and people, we introduce some of the key concepts and strategies that may help to move the discussion around data and ML forward. The last part of this section is dedicated to the future of ML and, by extension, the future of the urban environment based on it. **Julian Bleecker** (Near Future Laboratory) in Chapter 24 discusses his idea of Design Fiction as a way to generate future scenarios in design, science, fact and fiction. Bleeker's work on Design Fiction is increasingly used by designers as a successful strategy to explore the possible consequences (both in a positive and negative light) of technology for people. Bleecker's work explains how design fiction can be used to influence the general public's understanding or expectations about new technology, thus providing this technique with an active design role. We then introduce an extreme position where the development of intelligent systems and AI may reach a point in the future when the current balance of human-machine is completely altered. In Chapter 25 *Superintelligence and Singularity*, **Ray Kurzweil** (Google) discusses this perspective, introducing a future where we will witness a "*merger of our biological thinking and existence with our technology, resulting in a world that is still human but that transcends our biological roots*" (Chapter 25). Finally, in Chapter 26, **Vincent J. Del Casino Jr** (San José State University), **Lily House-Peters** (California State University), **Jeremy W. Crampton** (University of Kentucky) and **Hannes Gerhardt** (University of West Georgia) examine one of the physical embodiments of AI: robots and their relationship with people (both designers and users), in a reflection on new social geographies. Although not specifically covered in this book, robots offer an interesting point of connection between human and machines, where ML and AI may have a closer encounter with humans. The last section includes 5 projects that embody some of the principles set out in the chapters. Project 24 *Experiments in Synthetic Data* by **Forensic Architecture** illustrates how ML can be used in open source investigations to deal with incomplete or challenging datasets. In Project 25 *Emotional AI in Cities: Cross Cultural Lessons from UK and Japan on Designing for An Ethical Life*, **The Emotional AI Lab –Vian Bakir** (Bangor University, UK), **Nader Ghotbi** (Ritsumeikan Asia Pacific University, Japan), **Tung Manh Ho** (Ritsumeikan Asia Pacific University, Japan), **Alexander Laffer** (Bangor University, **Peter Mantello** (Ritsumeikan Asia Pacific University, Japan), **Andrew McStay**, (Bangor University, UK), **Diana Miranda** (University of Stirling), **Hiroshi Miyashita** (Chuo University, Japan), **Lena Podoletz** (University of Edinburgh, UK), **Hiromi Tanaka** (Meiji University, Japan), **Lachlan Urquhart** (University of Edinburgh, UK)**–** showcase their approach to emotional AI where they "*explore how we may best live with technologies that pertain to sense, profile, learn and interact with people's feelings, emotions and moods*" (Project 25). In Project 26 *Decoding Urban Inequality* **Kadeem Khan** (Facebook) explains how ML methods can be used to serve noble purposes like providing useful insights on spatial inequality in cities like Nairobi, Kenya.

The last 2 projects show how design fiction can be used to examine possible scenarios. In Project 27, **Maria Luce Lupetti** (TU Delft) presents her work on *Amsterdam 2040*. This is a fictional future scenario that illustrates design fiction methods in action. **Jason Shun Wong** (Project 28) also presents his work *Committee of Infrastructure* where there is a fictional Los Angeles city council meeting where people need to address "*issue of agency, representation, and intention within the domain of machine learning and artificial intelligence* (AI)" (Project 28).

What is next?

This book collects some of the key texts and projects that represent the current use of computational methods in architecture and urban design with a specific focus on ML. This is underpinned by a growing level of interest in the application of not only computer science and digital technologies but data science as well. It is, in fact, becoming increasingly apparent that in order to use these new technologies and methods, designers and researchers need to become familiar with some of the fundamentals of the scientific approaches to data. In the 1990s and 2000s architects concentrated on the discovery and testing of new technologies (new 3D modelling and parametric software and CAD/CAM technologies for example) and their potentiality for architecture and cities. It was the time of form-finding and the first establishment of parametric architecture. In the 2010s, it became clear to designers that more of an understanding of the nature of data is needed. A number of architects and designers are increasingly characterised by a hybrid profile where the spatial practice of architecture is combined with data science and a deeper understanding of how the data should be used, analysed and read. Computational design is a growing discipline and this is evident in the number of new academic and training courses that include computation in their curriculum around the world, in the number of large and small architectural practices that have a computational branch (this extends to the AEC industry at large), and in the interest of individuals that want to include computational methods in their skillset. Fast-forward to 10 - 20 years from now, it is not difficult to imagine that data science and computational methods may be part of architectural training and the expertise of designers (Carta 2020). We hope that this book, with its insightful contributions from world-leading researchers and designers, will help all those who want to increase their understanding and knowledge of ML and AI and start using some of the techniques in their own work.

References

Alexander, C. (1964). *Notes on the Synthesis of Form* (Vol. 5). Harvard University Press.

Babbage, Charles (1864). *Passages from the Life of a Philosopher*. Longman and Co. p. 67 Available at: https://www.google.co.uk/books/edition/Passages_from_the_Life_of_a_Philosopher/-Mod6ePT28kC?hl=en&gbpv=0

Boyd, D. and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15 (5): 662–679.

Carta, S. (2020). Machine learning and computational design. Ubiquity *2020*(May). 1–10.

Cukier, K. and Mayer-Schoenberger, V. (2014). *The Rise Of Big Data: How It's Changing The Way We Think About The World* (pp. 20–32). Princeton University Press.

Foth, M. (ed.) (2008). *Handbook of Research On Urban Informatics: The Practice And Promise Of The Real-Time City: The Practice And Promise Of The Real-Time City*. IGI Global.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661*.

Harari, Y.N. (2016). *Homo Deus: A Brief History of Tomorrow*. Random House.

Karras, T., Aila, T., Laine, S. and Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint ar Xiv:1710.10196*.

Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083): 20160115.

Neal, Z.P. (2012). *The Connected City: How Networks Are Shaping The Modern Metropolis*. Routledge.

Ostwald, M.J. (2011). The mathematics of spatial configuration: Revisiting, revising and critiquing justified plan graph theory. *Nexus Network Journal* 13 (2): 445–470.

Ostwald, M.J. and Williams, K. (2015). Mathematics in, of and for architecture: A framework of types. In *Architecture and Mathematics from Antiquity to the Future* 31–57. Cham: Birkhäuser.

Ostwald, M.J. and Dawes, M.J. (2018). *The Mathematics of the Modernist Villa: Architectural Analysis Using Space Syntax and Isovists*. (Vol. 3). Birkhäuser.

Ratti, C. (2010). June. The senseable city. In *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application* (pp. 1–1).

Scheutz, M. (ed.) (2002). *Computationalism: New Directions*. MIT Press.

Stojanovski, T. (2018). City Information Modelling (CIM) and Urban design-morphological structure, design elements and programming classes in CIM.

Viana, D.L., Morais, F. and Vaz, J.V. (eds.) (2018). *Formal Methods in Architecture and Urbanism*. Cambridge Scholars Publishing.

Wolfram, S. (2002). The principle of computational equivalence. Ch. 12 In: *A New Kind of Science* 5–6 and 715–846. Champaign, IL: Wolfram Media. https://mathworld.wolfram.com/PrincipleofComputationalEquivalence.html

Section I

Urban Complexity

1

Urban Complexity

Sean Hanna

Bartlett School of Architecture, UCL, UK

Cities are arguably the most complex things we have ever built. Most of humanity now lives peacefully and productively in groups vastly larger than our natural social capacity would allow, and as this proportion only continues to grow (United Nations 2019), the problems it poses for those responsible for designing and managing the city are equally great. In part this is due to the increasing complexity of the city itself. We humans are equipped with intuitions about spaces we can see at a glance and easily traverse on foot, and about social groups of fewer than 200 people (Dunbar 2014), but cities are well beyond our human scale. Historically, we have often managed urban growth gradually, building by building and street by street, but we can now construct environments for millions of people almost overnight, without feedback. In part this is due to unprecedented technological change. We do not have sufficient precedents for self-driving cars, smart homes, or the setting of our social and commercial interactions away from the streets and into virtual space, yet we need to understand the nature of the city so that we don't plan counter to it.

What we do have is access to data, more than ever before, although there are different approaches to how we can use it. Traditional scientific research, on which much of our understanding of cities rests, begins with theory or hypotheses. The field of Space Syntax, for example, is a scientific discipline through which accurate predictions can be made about the effect of spatial configuration on traffic density, crime, social interaction, property values, and other complex phenomena (Penn 2003; Hillier 2007; Silva 2017). It consists of a set of theories about the relationship between space and society, including the influence of space on the natural movement of people. From these, along with associated representations of space and related analytical methods, specific hypotheses or predictions are made, which can then be tested against observations. Data on real human movement and behaviour enter this scientific process last, to test the hypothesis, which is sometimes corroborated, and sometimes refuted, thereby advancing the core theory. Importantly, theories themselves are valued for their clarity, and in some cases the stated theory necessarily simplifies what patterns are in the data for the sake of this clarity and understanding. Sometimes, complexities in the pattern are missed.

Machine learning (ML) approaches instead begin with the data, and attempt to discern patterns from within them, with the intent that these patterns will reliably inform decisions we make about the city. The complexity of these patterns determines how they do so: where they are clear enough for us to articulate they may guide general theory and policy,

Machine Learning and the City: Applications in Architecture and Urban Design,
First Edition. Edited by Silvio Carta.

and where they are not they may be used to make highly contextual predictions. The essential problem of ML is that of picking out these patterns. As Jacobs (1961) articulated in the early days of complexity science, there is a profound difference between the 'disorganised' complexity if millions of independent actions, which might be treated statistically, and the 'organised' complexity of systems (people, spaces, data, goods) for which their mutual interactions matter. Cities are the latter.

Phenomena that matter most in a city do not lend themselves to treatment by gross statistical analysis, precisely because of these interactions between many parts. While statistics isolate one or two variables, for example a particular demographic of the population, ML can be valuable just because it may find the patterns among many variables. Such variables might include the population's distribution in space, movement in time, connection via technology, and so on. But each new variable potentially interacts with all the others, potentially increasing the complexity of the system to be studied exponentially. The essential problem in such an analysis is in understanding whether there exist any points, scales, or levels of representation at which a meaningful pattern can be extracted. If ML begins with the data, is this even possible with something so complex as a city? How is the city intelligible to the machine?

1.1 How Can a Machine Understand the City?

For some phenomena, patterns are predictable because they converge with increasing scale or time. Agent-based modelling often owes its effectiveness to the fact that a large population is used. Turner and Penn's (2002) exosomatic visual architecture (EVA) agents, for example, are extremely simplified models of pedestrians in space, which make random navigation decisions based on a probability weighted by how far they can see in any given direction. Their individual paths are entirely unrealistic, appearing often to walk in circles, but over an extended time, a population of agents will converge very closely towards the distribution of real people in a space, correlating approximately 76% (Turner and Penn 2002) with observed pedestrians. The result illustrates clearly that individual people are unpredictable, but, in some ways at least, the aggregate is predictable.

This suggests one factor that makes the complexity of cities intelligible: the fact that useful patterns will appear at a sufficient scale. A rank-size relationship is one 'law' that has been found in many aspects of cities: if the population size for a group of cities is plotted against their rank on a logarithmic scale, the result is nearly linear (Batty 2006). This change in population is directly relevant to social factors, in that total measured values of variables like economic output, income, patents, as well as crime, all scale reliably with population, and increase at a rate of about 15% more than linear; people even walk faster in larger cities (Bettencourt et al. 2007). The same scaling pattern appears also in properties closer to the domain of the architect and planner, such as the heights of tall buildings (Batty 2008) and the degree of connections between streets in a city (Jiang 2009). All of these are high-level patterns that apply to the aggregate only; Batty (Batty 2006, p. 12) has shown that there is no discernible regularity as to where an individual city will appear in the size ranking over time, as their dominance rises and falls randomly over centuries.

Another factor is that some of these high-level patterns involve a relationship between the variables that we can design and plan, such as spatial configuration, and the social or economic factors we might desire in cities. The angular betweenness centrality of street segments within a given urban network gives a theoretical measure of how much traffic is likely to pass through any given street segment, and observations confirm that greater centrality corresponds to greater pedestrian and vehicle count, and that the network of major roads can be reliably found in any network (Hillier 2007). But, depending on the scale used, the same measurement can clearly pick out the location of local high streets, the known centres of local neighbourhoods, and where commercial activity is actually located, all as a function of the street geometry (Hillier 2007). Local shops, for example, will tend to be successful in zones of maximum centrality measured at a radius of about 1 km; we can know this even for streets yet to be built, and use this information to plan. The same applies across far larger scales, even to the extent of indicating the major locations of commercial and economic activity on international street networks, in which the pattern of centrality values correlates with the economic output of nations – the higher a country's total centrality, the higher its GDP (Hanna et al. 2013). An analysis of such centrality across the range of scales (Krenz 2017) indicates that the scale hierarchy seen among cities, and in distances between them, is also a property of the network itself. It is not matched by randomly generated networks, which suggests that these human spatial networks may be optimised for patterns of particular human activity.

1.2 Cities Are Optimised to Make Some Patterns Clear

Some patterns of cities appear to be discernible simply because cities are optimised to make these patterns evident. Of the many possible ways of arranging roads, spaces, and buildings across a surface, real cities are a quite constrained subset. This is useful in any machine search for regularities, as it drastically reduces the space to be searched.

The agent models discussed in Section 1.1, just like individual people, have a view only of their immediate surroundings, yet the movement they predict strongly resembles that of global centrality measures. Is the pattern one of the small scale or the large? Is it of the cognitive, phenomenological properties of moving pedestrians or of the structural properties of roads and space? Causally, these appear to be entirely independent of one another, and it has been noted of methods that study the structure of space that they 'cannot account for the dynamics of movement' (Batty 2001). Those who analyse street structure have argued that activity in a city is driven solely by the properties of the network (Ma et al. 2018), whereas others who look at individual path choice see relevant visual cues and cognitive factors (Turner 2007; Emo 2014). The evidence supports both claims, not because they are necessarily related but quite possibly because cities are so often shaped such that the same patterns appear at both scales.

The reasons for this are evident when considering what it means to navigate an unfamiliar part of the city without a map. Immediate visual cues frequently lead us towards longer, wider streets where our visibility is greater, and most of us can normally rely on these to lead efficiently to our destination because, if they did not, we would be lost in a labyrinth.

Where natural footfall on a street does not bring many people in contact with a shop, that shop is more likely to fail as a business and disappear. The degree to which the large-scale properties are conveyed by the small can be measured, as in the space syntax measure of *intelligibility* (Penn 2003), which gives an assessment of how effective an area is at conveying this essential information.

Where we see evidence that large-scale structural information correlates with small-scale geometry, it need not take the complexity of the human visual system to reveal it. Agent models far simpler even than those discussed in Section 1.1 can be constructed to simulate random walks through a city by taking the street segment network as graph, with connections weighted by the angle at which streets meet; agents walk randomly but with a greater probability of continuing straight than making acute angled turns. The distribution converges quite rapidly, within 20 to 30 'steps', to approximate that of observed pedestrians and vehicles at distances ranging from neighbourhood to regional scales (Hanna 2020). This simple random process appears to mimic all the complexity of a city full of real human travellers (Figure 1.1).

This is significant because the route decisions made are entirely local, with agents unable to 'see' beyond their current intersection, yet the model predicts real movement similarly

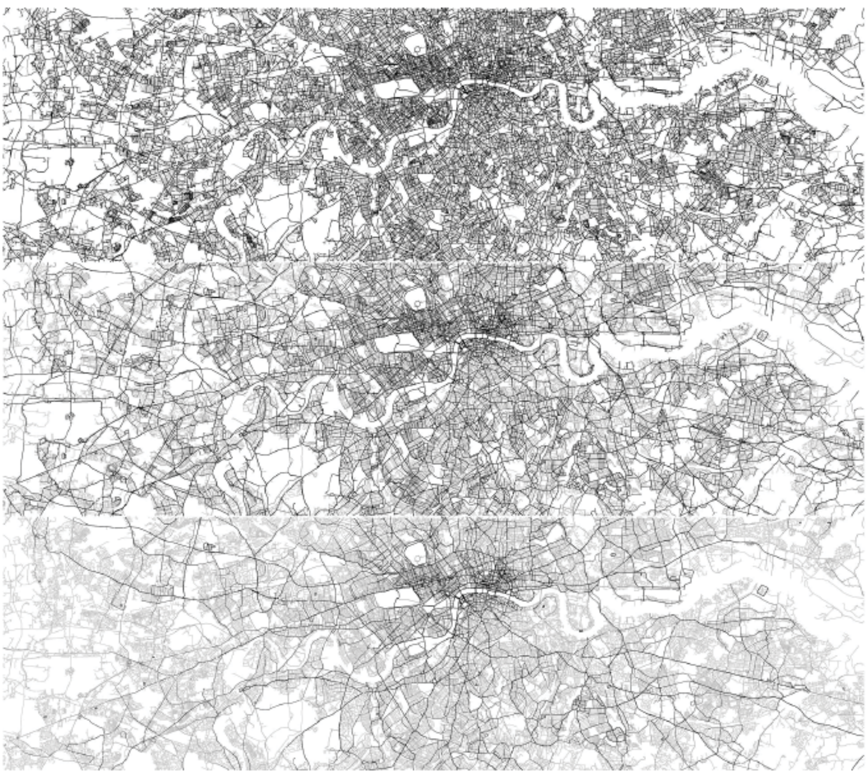


Figure 1.1 A simulated random walk on the street network of London, at 5, 26, and 100 iterations. Agents begin uniformly distributed (top) but aggregate on the more highly trafficked roads after a short time (bottom), revealing the major routes and areas of traffic density.

to others that explicitly optimise longer-range routes. Methods such as betweenness centrality (Hillier and Iida 2005), which calculates optimal paths through graph nodes, and network analyses of continuity lines (Figueiredo and Amorim 2005) or natural streets (Jiang et al. 2008), which group together sequences of segments with minimal angles of turn, exploit information at some distance across the city to model movement. The implication is that some longer-range knowledge of the network is necessary for navigation, and that people optimise their routes accordingly. But both these longer-range methods and the locally informed random walks correlate well (with Pearson coefficients > 0.7) with movement, and with one another, which suggests, at least for the urban networks studied, that in real cities there is a rough equivalence between navigating based on knowledge of the street map and navigating based on immediate visual cues, and that the geometry of the street network is optimised such that it conveys the relevant information about distant routes to a naive traveller at any intersection (Hanna 2020). The same patterns are clear at both large and small scales.

Explicit optimisation is rarely likely to have been the cause of such intelligibility. In the case of New York's Central Park, for example, centrality analyses of a hypothetical street grid indicate that if the park did not exist the streets in its place would be less central and rarely used, simply due to asymmetries in the shape of the island (Al Sayed et al. 2009, pp. 1–12). Manhattan's planners placed the park exactly where modern computational methods would recommend, but without any such methods being available at the time. The causes of such decisions in real cities are often too complex to be known for certain: market forces and competition may determine the location of commercial property or of parks; cultural precedents may suggest resemblances to other known cities; the political pressures on design and planning are numerous. But to the extent that cities are intelligible, they are so because their patterns are obvious even when we are not certain of their underlying cause.

1.3 Nondiscursive Features Also Appear in Data

Many of the qualities relevant to us in our own experience of the city are more complex even than we can precisely articulate. The style of buildings in a particular neighbourhood is intuitively recognisable to us as different from another, yet the precise features of those buildings that determine the difference are not easily described. It may seem that these nondiscursive properties are elusive, or impossible to quantify, but the patterns are no less real for their complexity, and also there to be found by the machine.

To investigate properties of building form in Athens and London, Laskari et al. (2008) used the shape of the combined footprint of buildings within an urban block as the unit for comparison, which captured essential properties of building width, density, and uniformity in the shape and size of the internal courtyards hidden away from the street facade (Figure 1.2). Fourteen different measurements were taken for each unit, including straightforward values of perimeter and area, and more complex ones, such as fractal dimension and quantities derived from the lines of sight within the courtyard spaces. All such measurements are legible automatically, using no more than rudimentary machine vision, to yield a fourteen-dimensional data point for each building block. When these are compared,

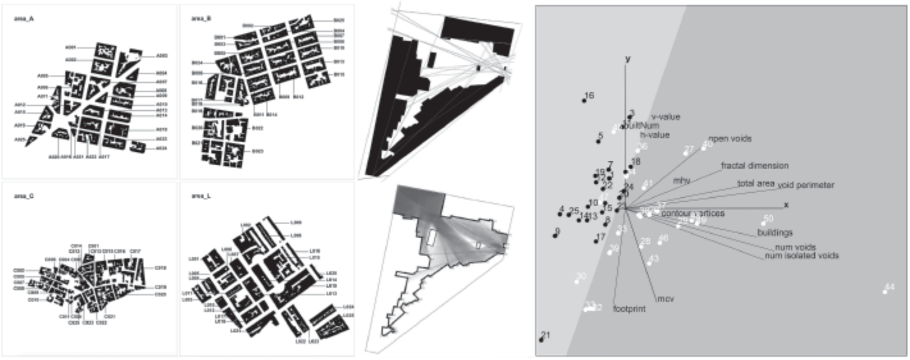


Figure 1.2 Building footprints, which differ from neighbourhood to neighbourhood, naturally form distinct clusters associated with these neighbourhoods in the space given by various automated measurements of their shape. (*Source*: redrawn from Laskari et al. 2008).

25 points for each of four neighbourhoods in Athens, in addition to Bloomsbury in London, clear clusters of points were seen to differentiate one neighbourhood from the next. While these are not perfectly separable, with some points overlapping, the differences coincide well with our own intuitive assessments of building style or type. The clusters within Athens are closer to one another than any of them is to London, and those neighbourhoods of a similar age are closer than those built a century apart. The result quantifies, for the machine, just those complex aesthetic properties that we would find so hard to describe.

Although its judgements coincide with our own perception of stylistic categories, most of the measurements do not resemble something a human observer would notice. The machine 'sees' the buildings in a plan view of the entire block, a view which is not given to its occupants or passers-by. The measure of fractal dimension might be thought of as a degree of complexity of this plan, but only approximately. Some measures of the lines of sight of the courtyards could be considered as a degree of convexity, but only approximately. The fact that machine and human judgements of the categories agree despite this difference between human vision and machine measurement suggests that such patterns are not dependent upon the selection of particular features to be used as inputs, and that they are likely to be found with relative ease regardless of which method is used.

Such a result is exactly what we would hope if we were concerned with a machine's capacity to pick out these relevant clusters, because we needn't be too concerned about choosing the correct input features. The best strategy, in this case at least, seems simply to have as many different features as possible. When analyses are compared using different groups of features (Hanna 2011), the correct clusters become more clearly differentiated as more features are used. This is not a case of having more dimensions in which to divide the classes, as is done in supervised learning; a fixed number of principal components is used to ensure each clustering is made in a space of equal dimensions. The results show that different machines classifying the sets of buildings converge both with one another and with the correct identification of neighbourhoods, with greater numbers of inputs. The relevant patterns appear readily in the data drawn from the buildings, regardless of the particular representation used.

1.4 More Complex Patterns Can Be Learnt

When the relevant patterns do not so readily fall out of the data but are a more complex function of the input dimensions available, ML can be used to find them. Like the form of building plans shown in Figure 1.2, the local configuration of roads differs depending on the land use, but the precise features relevant to these differences are not obvious. In recent analyses of the UK road network, Tasos Varoudis has used a type of neural network known as an auto-encoder, which is trained to extract the principal nonlinear variances in the data from local samples of streets, and thereby map them to a subspace in which the most relevant differences are clear (Figure 1.3). The data are not preprocessed by taking any predetermined measurements but are instead presented directly to the network as street graphs,

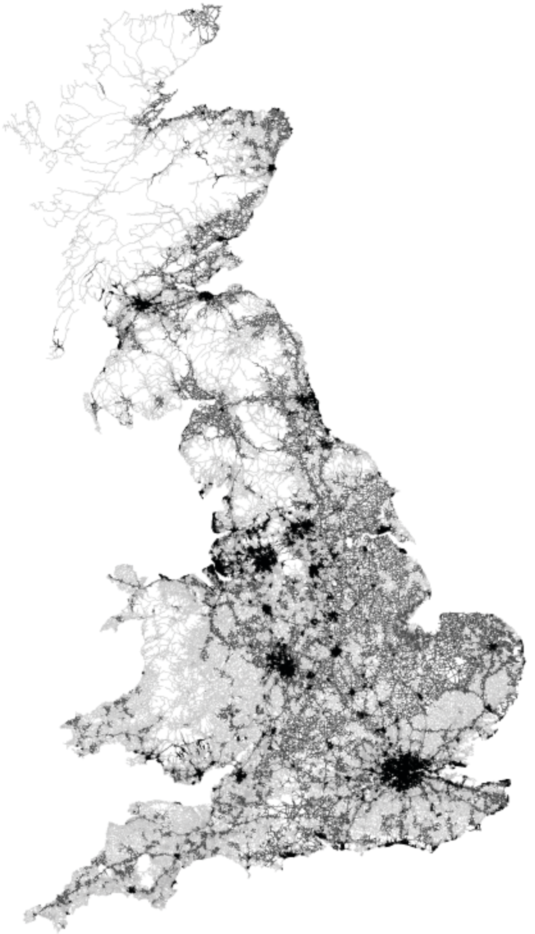


Figure 1.3 Differences in road morphology are clearly distinguished by a neural network, here revealing natural clusters which correspond with actual land use designations. (*Source*: redrawn from Varoudis and Penn 2020).

represented as adjacency matrices. The natural clusters of similarity that appear correspond almost exactly with the land use regions designated as urban areas, farmland, or natural landscapes in surveys such as the Corine Land Cover Inventory (2018).

If the relevant classes or features we are looking for are known beforehand, but there is complexity in the input, supervised methods can be used to train the ML algorithm. Much larger urban graphs have been classified using their graph spectra, which represents the entire city as a vector in many more dimensions than the local samples shown in Figure 1.3. Clustering cities in this high-dimensional space results in very little discernible pattern, but the geographical location of the cities can be used to tell the supervised learning algorithm what to look for. In Hanna (2009), a training set of cities is presented as input to a support vector machine, identifying each one as, for example, a European, or Asian, or North American city. Once trained, the algorithm correctly classifies new cities with an accuracy of between 75 and 85%, based entirely on their form.

Complexity often comes not from the scale of the sample but from a considerable overlap in the input dimensions. In Thirapongphaiboon and Hanna (2019), centrality measures at varying scales were seen to correspond to different types of urban land use. Commercial buildings tend to be located on street segments with high values of closeness centrality at low radii, under 1.8 km, whereas business and industrial buildings correspond to higher radius measures from 1.8 to 7.2 km, and node count, or density of streets, is also relevant measure. Residential use, by contrast, is marked by low centrality across the full range of radii. With much overlap, no single measure, scale, or selected group of such makes the distinction between these uses clear in itself, but supervised learning uses the known classes (in this case commercial, business, or residential) to derive a particular spatial signature that best describes each class. With this, the proportion of land use can be predicted using a multilayer perceptron for street segments with an accuracy of more than 80% (Thirapongphaiboon and Hanna 2019).

Much more specific land uses can also be identified by such spatial signatures, such as particular business types, or even locations of an individual chain. Silva (2017) used a random forest algorithm to predict, for a range of centrality measures of a given street segment, whether it was likely to contain types of business, such as pubs, cafes, or travel agencies, each of these being correctly identified more than 70% of the time. Some particular business chains, such as Waterstones bookshops, could not be placed any better than chance, but Starbucks' locations (and solicitors) were positively predicted at a rate of more than 80%.

1.5 Putting the Patterns to Use

If the examples in Section 1.1–1.4 have focused entirely on the search and understanding of regularities in the data rather than the task of managing, intervening in, or designing the city, it is in part because this pattern recognition is the strength of ML. But it is also because pattern recognition is such a natural and intuitive part of our own cognition its importance is overlooked, and because the necessity of coping with novel, larger, and more complex patterns in cities has never been more acute. The rapidly changing requirements of cities and the speed of their construction mean that decisions are more costly than ever – not only in the present but also over the long term, socially and economically. The ability to project these patterns into new scenarios allows them to be tested *in silico* before committing, to use this knowledge to place a business where natural footfall will mean it will thrive, to target changes to streets so that the city can be navigated effectively, or build new sections of the city that remain naturally connected and vibrant.

The apparent limitlessness of complexity may seem to be a problem in that we will always find more of it, if we look deeper, if we have more data. The examples in this chapter predict only the long-term behaviour of many individuals, but individual behaviour (thankfully) and many lower-level patterns may be forever beyond our ability to determine. It is fortunate, then, that the scale of the regularities we are able to discover happens to coincide with the scale of our intervention. We design for aggregates of many people, not for single individuals. We design for the climate over the span of years, not for the weather of a single day. Even to the extent that we could in principle forecast individual behaviour, as we may with increased access to large sets of personal data, this is not the level at which our design and planning decisions are made. The aim, for example, to design cities that bring diverse individuals into contact with one another is a higher-level goal, which requires descriptions of many people in many spaces over extended times, and this is the level of description of complexity with which we need to contend. These complex patterns of the aggregate, which are most important, are also those which are most easily found in the data.

This trait suggests the reason why ML is useful in the context of the complex city, a reason too easily overlooked in the day-to-day training of learning models, which are judged on their success in prediction. The problem with prediction, in the sense of foretelling the outcome of specific events, is that it is not possible in the context of the wicked problems of the complex city, nor should we aspire to it. What is more useful to us is understanding, which, in the best case, is what the patterns extracted by ML will provide. Like the theory-led approach, the data-led approach can give us models that generalise sufficiently to tell us how the phenomena we care about – including social interaction, economic activity, movement, and more – will occur in new and different urban environments. When faced with complexity, the recognition of patterns, otherwise invisible owing to the scale or form of data in which they appear, is a way of seeing the city more clearly, of extending the limits of our own natural understanding, and, ideally, a means to inform better decisions.

References

Al Sayed, K., Turner, A., and Hanna, S. (2009). Cities as emergent models: The morphological logic of Manhattan and Barcelona. In: *Proceedings of the 7th International Space Syntax Symposium* (eds. D. Koch, L. Marcus, and J. Steen), Stockholm: Royal Institute of Technology (KTH). http://www.sss7.org/Proceedings/05%20Spatial%20Morphology%20and%20Urban%20Growth/001_AlSayed_Turner_Hanna.pdf (accessed 17 October 2021).

Batty, M. (2001). Agent-based pedestrian modeling. *Environment and Planning B: Planning and Design* 28 (3): 321–326. 10.1068/b2803ed.

Batty, M. (2006). Rank clocks. *Nature* 444 (7119): 592–596.

Batty, M. (2008). Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies. In: *Encyclopedia of Complexity and Systems Science* (ed. R.A. Meyers), Vol. 9, New York: Springer.

Bettencourt, L.M., Lobo, J., Helbing, D., Kühnert, C., and West, G.B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences* 104 (17): 7301–7306.

Corine Land Cover Inventory (2018). Corine Land Cover. Copernicus Program. https://land.copernicus.eu/pan-european/corine-land-cover (accessed 17 October 2021).

Dunbar, R. (2014). *Human Evolution: A Pelican Introduction*. Penguin.

Emo, B. (2014). Seeing the axial line: Evidence from wayfinding experiments. *Behavioral Sciences* 4 (3): 167–180.

Figueiredo, L. and Amorim, L. (2005). Continuity lines in the axial system. In: *The Fifth Space Syntax International Symposium*, 161–174. Delft, The Netherlands: Delft University of Technology. http://spacesyntax.tudelft.nl/media/Long%20papers%20I/lucasfigueiredo.pdf (accessed 17 October 2021).

Hanna, S. (2009). Spectral comparison of large urban graphs. In: *Proceedings of the 7th International Space Syntax Symposium* (eds. D. Koch, L. Marcus, and J. Steen), Royal Institute of Technology (KTH).

Hanna, S. (2011). Design agents and the need for high-dimensional perception. In: *Design Computing and Cognition '10* (ed. J.S. Gero), 115–134. Dordrecht, The Netherlands: Springer.

Hanna, S. (2020). Random walks in urban graphs: A minimal model of movement. *Environment and Planning B: Urban Analytics and City Science* 48 (6): 1697–1711.

Hanna, S., Serras, J., and Varoudis, T. (2013). Measuring the structure of global transportation networks. In: *Proceedings of the Ninth International Space Syntax Symposium* (eds. Y.O. Kim, H.T. Park, and K.W. Seo), Seoul: Sejong University.

Hillier, B. (2007). Spatial sustainability in cities: Organic patterns and sustainable forms. In: *Proceedings of the 7th International Space Syntax Symposium* (eds. D. Koch, L. Marcus, and J. Steen), Royal Institute of Technology (KTH), K01:1–K01: 20.

Hillier, B. and Iida, S. (2005). Network and psychological effects in urban movement. In: *Proceedings of Spatial Information Theory. COSIT 2005. Lecture Notes in Computer Science* (eds. A.G. Cohn and D.M. Mark), Vol. 3693, 475–490. Berlin, Heidelberg: Springer. https://doi.org/10.1007/11556114_30.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Random House.

Jiang, B. (2009). Ranking spaces for predicting human movement in an urban environment. *International Journal of Geographical Information Science* 23 (7): 823–837.

Jiang, B., Zhao, S., and Yin, J. (2008). Self-organized natural roads for predicting traffic flow: A sensitivity study. *Journal of Statistical Mechanics: Theory and Experiment* 2008 (07): 10.1088/1742-5468/2008/07/P07008.

Krenz, K. (2017). Regional morphology: The emergence of spatial scales in urban regions. In: *Proceedings of the 11th International Space Syntax Symposium*. Instituto Superior Técnico, Departamento de Engenharia Civil: Arquitetura e Georrecursos, 74.1–74.23.

Laskari, S., Hanna, S., and Derix, C. (2008). Urban identity through quantifiable spatial attributes: Coherence and dispersion of local identity through the automated comparative analysis of building block plans. In: *Design Computing and Cognition '08* (eds. J.S. Gero and A. Goel), Springer. https://doi.org/10.1007/978-1-4020-8728-8_32.

Ma, D., Omer, I., Osaragi, T. et al. (2018). Why topology matters in predicting human activities. *Environment and Planning B: Urban Analytics and City Science* 46 (7): 1297–1313.

Penn, A. (2003). Space syntax and spatial cognition or why the axial line? *Environment and Behavior* 35 (1): 30–65.

Silva, J.P. (2017). The syntactic signature of Starbucks' locations: Toward a machine-learning approach to location decision-making. In: *Proceedings of the 11th International Space Syntax Symposium* (eds. T. Heitor, M. Serra, M. Bacharel, L. Cannas da Silva, and J.P. Silva), 1–8.

Thirapongphaiboon, T. and Hanna, S. (2019). Spatial distribution of building use: Recognition and prediction of use with machine learning. In: *Proceedings of the 12th International Space Syntax Symposium*, 194:1–194:30.

Turner, A. (2007). To move through space: Lines of vision and movement. In: *Proceedings of the 6th International Space Syntax Symposium*. Istanbul Technical University, 1–12.

Turner, A. and Penn, A. (2002). Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment. *Environment and Planning B: Planning and Design* 29 (4): 473–490.

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. New York: United Nations.

Varoudis, T. and Penn, A. (2020). Spectral clustering and integration: The inner dynamics of computational geometry and spatial morphology. In: *Formal Methods in Architecture: Advances in Science, Technology & Innovation* (eds. S. Eloy, D.L. Viana, F. Morais, and J.V. Vaz), Springer, 243–250.

2

Emergence and Universal Computation

Cassey Lee

Institute of Southeast Asian Studies (ISEAS), Singapore

Original publication details: Lee, Cassey. (2004). Emergence and universal computation. Metroeconomica 55(2–3): 219–238.

Emergence refers to the spontaneous formation of higher-level (macro) structures or patterns in complex systems. Attempts to formalise the notion of emergence via algorithmic complexity theory runs into the problem that the Kolmogorov complexity function is not computable. The reason for this motivates a closer examination of the link between emergence and universal computation. Following Wolfram's pioneering work in the classification of cellular automata (CA) behaviour, the research programs of Langton and Crutchfield, while incomplete, provide important insights to economists seeking to understand the relevance of emergence and universal computation to their discipline. They lead to questions on the emergence of institutions and the concomitant changes in rule-based behaviour on the part of economic agents.

2.1 Introduction

The twentieth century is replete with scientific and mathematical discoveries that have profoundly changed our worldview. In physics, within a mere century, our view of the cosmos changed from a classical (Newtonian) to a relativistic one following Einstein's relativity theory in the early 1900s. In mathematics, Hilbert's faith in the closure of formal axiom systems fell apart with Gödel's incompleteness theorem in the 1930s.

Out of these ashes of lost deterministic foundations arose the sciences of complex systems, first in the study of nonequilibrium thermodynamics (under the intellectual leadership of Ilya Prigogine in Brussels) and later, in broader interdisciplinary terms, in New Mexico with the establishment of the Santa Fe Institute. Interestingly, a core element of this new paradigm – the notion of emergence – reflects the passage of the sciences and mathematics from a focus on closed and deterministic systems to open and dissipative systems, where order, structure, or patterns arise seemingly out of nowhere (at least as far as initial and boundary conditions are concerned).[1]

Machine Learning and the City: Applications in Architecture and Urban Design,
First Edition. Edited by Silvio Carta.

The study of emergence began first with empirical observations followed by attempts at theoretical formalisations. Thus far, the later endeavour has not yielded a unified and universal theoretical treatment of emergence. Instead, the study of emergence has extensively been carried out via computer simulations. Some have even conjectured that computer simulation is the only method of studying true emergent phenomena.[2]

Aside from the extensive use of computer simulations to study emergence, scholars are also studying emergence from a computational theory point of view. It is in this area of study that we find a confluence of two concepts – emergence and universal computation – that found place in the fertile mind of Alan Turing some years ago.[3]

In this paper we seek to understand some of the key questions that arise from the study of emergence and its relationship to universal computation. A fundamental question pertains to the possibility of characterising emergence formally. This leads to an examination of the relationship between universal computation and emergence. The rest of the paper reflects on the relevance of these issues to economics.

Why should economists be interested in emergence and universal computation? Economists have long understood that the economy is a complex system. Recent research in complex systems, particularly in the physical and biological sciences, offers novel analytical technologies and frameworks to economists to renew their study of the economy as a complex system. This is particularly true in the study of emergence in the economy which includes diverse areas such as macroeconomics (macroeconomic fluctuations), transportation economics (traffic jams), urban economics (hierarchy of cities), industrial organisation (motion picture industry), and political economy (mass protest behaviour). The diverse methodological approaches undertaken in each of these areas point to the need to define emergence. An effort to define emergence formally, while difficult, is important in order to set formalisation on the 'right' path in this area.[4] This brings us to computable economics and the tantalising links between emergence and universal computation.

Since the theory of computation is generally not part of the standard tools of trade of the economist, we begin with a basic discussion of what we mean by computation and universal computation in Section 2.2. This is followed by a discussion on emergence in Section 2.3. The relationship between emergence and universal computation is explored in Section 2.4. The theory of self-organised criticality is discussed in Section 2.5. The relevance of emergence and universal computation to economics is explored in Section 2.6. Section 2.7 concludes these discussions.

2.2 Universal Computation

We begin with the notion of computation. What is computation? A simplistic answer would be computation involves the mapping of a set of numbers to another set of number:

$$\mathbb{R} \to \mathbb{R}$$

We are familiar with the notion of a function that is defined in terms of the mapping of numbers:

$$f : \mathbb{R} \to \mathbb{R}$$

But computation is not merely a mapping of numbers. Computation involves the use of finite procedures or algorithms to generate number mappings. These procedures or algorithms comprise a series of steps or instructions to generate a set of numbers from an initial set of numbers. Velupillai (2002) notes that the notion of 'function' in mathematics is historically associated with the notion of a rule, a procedure, a set of instructions to perform a task.

A computation model is an entity that is capable of carrying out computations. One famous example of a computation model is the Turing machine. The model consists of a hypothetical/conceptual machine with a device (head) that reads/writes on an infinite tape (i.e. infinite memory) with discrete cells that move back and forth. Computation takes the form of writing symbols (e.g. 1s and 0s) on the cell it is reading based on the existing symbol on the cell it is currently reading as well as the machine's present state. This is followed by reading either the present cell or the cells to its left or right.

For our purpose, it is sufficient to note that the basic elements of the Turing machine are[5]

1) the machine's present state and its future state
2) the present and future symbols
3) movements of the tape (no movements, movement to the left, and movement to the right)
4) the program (collection of algorithms) which determines the future actions of the machine (i.e. the result of computation).

An interesting variant of the Turing machine is one that is capable of carrying out anything computable by an algorithmic process, i.e. capable of universal computation. Such a machine, which is called the universal Turing machine (UTM), is a Turing machine that uses as its input the computation results from another Turing machine, thus enabling it to simulate the computational processes of the other Turing machines.

Can other models of computation exhibit universal computation? Interestingly, the Church–Turing thesis suggests an affirmative answer to this question. The Church–Turing thesis states that any computable quantity (effective process) can be computed on a UTM. True enough, a variety of systems have been shown to exhibit universal computation – ranging from abstract models such as quantum computers, CA, and systems of partial differential equations to collections of physical entities, such DNA molecules, hard sphere gases, and lattice gases.

2.3 Emergence

2.3.1 Definitions

Emergence has been studied in a variety of disciplines (physics, biology, chemistry) and at different levels of sophistication. What is emergence? The following definitions give us some idea of what emergence is.

> Emergence is understood to be a process that leads to the appearance of structure not directly described by the defining constraints and instantaneous forces that

control a system. Over time something new appears at a scale not directly specified by the equations of motion. An emergent feature also cannot be explicitly represented in the initial and boundary conditions.

(Crutchfield 1994)

In (complex adaptive) systems, agents residing on one scale start producing behaviour that lies one scale above them: ants create colonies, urbanites create neighbourhoods... The movement from lower-level rules to higher-level sophistication is what we call emergence (Johnson 2001).

The above definitions suggest that emergence is a phenomenon that involves the spontaneous formation of a higher-level (macro) structure or pattern in a system that is brought about by local interactions between a large number of components of the system.[6] The formation of hierarchy is clearly an important element of emergence. In a hierarchical scheme, interactions at the different levels or scales (e.g. companies, countries) play a different role. It is also often emphasised that such higher-level structures cannot be deduced (or predicted) by looking at the components alone. A consequence of this is that reductionist methods are not useful in studying emergence. Whether this implies that a formal (mathematical) approach is also ruled out is an interesting question. But which 'mathematics' are we referring to? Formalist or constructivist?

2.3.2 Formalising Emergence

How can the notion of emergence be formalised?

Let S be the initial state space, I the input space, and S' the new state space. Let function f represent a process that transforms state S into a new state S':

$$f : S \times I \rightarrow S'$$

where $S \neq S'$

At first glance, we can surmise that there is nothing special about this function. This could be a function characterising any dynamic process. How would a function exhibiting emergence differ from f? It is perhaps useful to review the characteristics of emergence.

In a process that exhibits emergence, the new state $s'' \in S'$ is not merely just different from the initial state $s \in S$. There is a 'qualitative' difference between the states that do not exhibit emergence ($s' \in S'$) and those that exhibit emergence ($s'' \in S'$, where $s'' \neq s'$).

More specifically, as discussed in Section 2.1, emergence results in a macrolevel structure that is not present either in the initial state (s) or in other new states (s') that do not exhibit emergence. Furthermore, the difference between emergent processes and a nonemergent process is a hierarchical one:

$$\phi(s'') > \phi(s')$$

where ϕ denotes a measure of complexity (or hierarchy). But how can we measure complexity? One can measure complexity by using the algorithmic complexity theory. In the theory, the *Kolmogorov complexity K* of an object (in this case the emergent state s'') is measured by the smallest program that can be used to compute it.[7] This can be defined more precisely in the context of a specific computation model. The UTM is one such candidate.[8]

Definition 1: The Kolmogorov complexity $K(x)$ of a state x with respect to a UTM is

$$K(x) = min\, l(p)$$

where $l(p)$ is the minimum program size that can print x and halt.

Using the above definition, we can then state that an emergent state s'' is more (Kolmogorov) complex than a nonemergent state s':

$$K(s'') > K(s')$$

There are several problems with this approach.

First, it is possible that two UTMs might disagree on whether s'' is more complex than s'. Which UTM should then be used as the reference machine? Standish (2001) does not consider this a problem if we can posit an observer who can compare both descriptions (by the different UTMs) and decide whether they belong to equivalent classes. Second, a nonemergent state s' might be more 'complex' than an emergent state s'' in an algorithmic information sense. For example, longer strings would be required to describe a purely random state. Standish (2001) overcomes this problem by comparing the length of the description (l) with the corresponding measure of an equivalent class (determined by the observer).

A more fundamental problem with the above approach is that $K(\cdot)$ is not computable. In other words, we cannot even be sure that any program size is the smallest. This relates to the fact that the output domain (the program) is larger than the input domain (axioms).[9] Ian Stewart makes this point in the context of Gödel's incompleteness theorem:[10]

> From Chaitin's viewpoint, Gödel's proof takes a very natural form: The theorems deducible from an axiom system cannot contain more information than the axioms themselves do.

I believe this is also the essence of Velupillai's (2002) discussion of Berry's paradox within the busy beaver framework. It is useful to note that the negative result is not a consequence of using the UTM to characterise complexity. For example, in the case of a cellular automaton (which is also capable of universal computation), the question of whether it is in any of the four states identified by Wolfram (1984) is also undecidable.

Even though we cannot determine emergence via the notion of Kolmogorov complexity, the above discussion is not entirely wasteful. The nature of our 'failure' provides some clues on how to tackle the notion of emergence.

First, the difficulty in formally characterising emergence comes from the fact that interactions at the lower level produce structures at the higher level. Why can't we know a particular program that produces emergence is the shortest program? It must be because there are potentially many other programs that can bring about emergence that we cannot know of until we run these other programs. Hence, the 'openness' of the system (for lack of a better word) may be the source of the failure of using Kolmogorov complexity to characterise emergence. The corollary of this openness is the importance of dynamics, i.e. the history of interactions. In other words, a formal understanding of emergence is not possible without understanding the dynamics of emergence.

Second, if the output domain is larger than the input domain, we are led to examine whether emergence involves a change in the computational nature of the system. More specifically, an analysis of the computational classes along Chomsky's (1956, 1959) hierarchy may be a more viable alternative to the minimal program approach. Instead of determining the minimal program to reproduce a given sequence of integers, Chomsky's hierarchy approach seeks to determine the minimal computational capability necessary to reproduce it.[11] In Chomsky's hierarchy, formal languages are arranged in the order of their computational power. The four classes (in increasing computational power) are regular language, context-free language, context-sensitive language, and recursive enumerable language (e.g. Turing machines).[12]

Finally, inquiries into computational aspects of emergence have also led to the tantalising conjecture that a complex dynamic system exhibits universal computation when emergence occurs. This is examined in Section 2.4.

2.4 Emergence and Universal Computation

2.4.1 Wolfram's Classification

An early hint on the link between emergence and universal computation can be found in Wolfram (1984). In that paper, Wolfram shows that the behaviour of one-dimensional CA models falls into four distinct classes, namely:

1) Class I: the system evolves towards a homogeneous state with probability 1, i.e. regardless of the initial values (this is analogous to limit points in a dynamical system).
2) Class II: the system evolves towards simple separated structures or periodic structures when initial values fall in a limited region (this is analogous to limit cycles).
3) Class III: the system evolves towards chaotic patterns with varying degrees of structure in a way that is dependent on an ever-increasing number of initial sites (this is analogous to chaotic attractors).
4) Class IV: the system evolves towards complex localised structures, sometimes long lived.

Wolfram conjectures that CA exhibiting the Class IV behaviour are capable of universal computation His definition of universal computation is similar to the definition discussed in Section 2.3, i.e. based on its ability to simulate any other computational system:[13]

> Computational universality implies that a suitable initial configuration can specify arbitrary algorithmic procedures. The system can thus serve as a general purpose computer, capable of evaluating any (computable) function. Given a suitable encoding, the system may therefore in principle simulate any other system, and in this sense may be considered capable of arbitrarily complicated behavior.

While this definition is standard, Wolfram draws our attention to an interesting property of a system exhibiting universal computation, namely its inherent unpredictability:[14]

> There are important limitations on predictions which may be made for the behavior of systems capable of universal computation. The behavior of such systems may in

> general be determined in detail essentially only by explicit simulation of their time evolution ... No finite algorithm or procedure may be devised capable of predicting detailed behavior in a computationally universal system ... Not only does the value of a particular site after many time steps potentially depend on the values of an increasing number of initial site values; in addition, the value cannot in general be determined by any 'short cut' procedure much simpler than explicit simulation of the evolution.

The last sentence carries the connotation that there is no way of knowing if a program (minimal in an algorithmic complexity sense?) exhibits universal computation except to simulate it. The problem with this 'proof by simulation' method is that we would not know how long or how many iterations are needed before the system exhibits the property of universal computation! If by emergence we mean the transformation of a system from a nonuniversal computational one to a universal computational one, this implies that we simply cannot know a priori when emergence will occur. But can we know whether a system is capable of universal computation even though we may not know when this will occur? This is a question that is difficult to answer. The issue is not entirely resolved in Wolfram (1984) but there is a suggestion that it may depend on the number of possible states (k) and the extent of local interactions (number of neighbours r).

Finally, two other interesting properties of class IV behaviour are highlighted by Wolfram (1984), namely the statistical characteristic of the system and irreversible evolution of the system. Unlike class I–III behaviours, the degree of fluctuations in statistical quantities in a system exhibiting class IV behaviour does not decline with an increase in the number of sites. Irreversibility in class IV behaviour is evidenced by the generation of a small set of persistent structures which dominate the statistical property of the system.[15]

In the subsequent work of Langton (1992), Crutchfield and Young (1990), and Crutchfield (1994) attempts were made to go beyond Wolfram's initial effort of identifying universal computation by trying to understand the dynamics underlying emergence and its relationship to universal computation. The two approaches by Langton and Crutchfield are quite different.

2.4.2 Langton's Extension

Langton (1992) extends Wolfram's work based on CA by applying an order parameter to derive the four behavioural classes identified by Wolfram. The order parameter λ incorporates both the number of possible neighbourhood states (K^N) as well as the number of possible transition paths:

$$\lambda = \frac{K^N n_q}{K^N} \tag{2.1}$$

where K is the number of finite cell states, N the neighbourhood size for each cell, and n_q the number of transitions to any given state in the assumed transition function.

The variation in λ produces a spectrum of CA behaviour that corresponds to that outlined by Wolfram (1984):

I $\rightarrow$	II $\rightarrow$	IV$\rightarrow$	III	(2.2)
fixed-point	periodic	complex	chaotic	

At the critical value λ_c that lies between class II and class III, there is a sharp increase in the number of iterations (time) taken by the CA to complete a cycle (for periodic CA) or achieve statistical convergence (for chaotic CA).[16]

How can we quantify these transitions to better understand what is happening? Langton (1992) uses two measures to locate λ_c – the Shannon entropy to measure information capacity and mutual information to measure interactions between CA cells.[17] Of the two measures, the mutual information approach is more interesting. Formally, the mutual information *I(A; B)* between two cells *A* and *B* is the correlation between them and is given by

$$I(A;B) = H(A) + H(B) - H(A,B) \tag{2.3}$$

where *H(A)* and *H(B)* are individual entropies while *H(A, B)* is the entropy of the two cells as a joint process. At λ_c, the value of *I(A; B)* is neither too low (random) nor too high (order) but sufficient to bring about 'cooperation' between cells to support computation. When this occurs, there is some information transmission between cells:[18]

> Correlations in behavior imply a kind of common code, or protocol, by which changes of state in one cell can be recognized and understood by the other as a meaningful signal. With no correlations, there can be no common code with which to communicate information.

Langton interprets the intermediate value of λ_c as indicating a trade-off between information storage (which entails lowering of entropy) and information transmission (which entails raising entropy). This trade-off is optimised at λ_c.

Finally, Langton also conjectures that the various classes of CA behaviour suggested by Wolfram have correspondences to computability classes in the following way:

I $\rightarrow$	II $\rightarrow$	IV $\rightarrow$	III	(2.4)
halting	halting	undecidable	nonhalting	

What explains undecidability at phase transition (class IV)? Langton attributes this to the phenomenon of a critical slowing down, where transients can diverge to infinity (transients grow exponentially with system size). What about universal computation at the same locus? Langton opines that (1992, p. 80):

> the existence of universal computation is explained by the fact that the dynamics of physical systems in the vicinity of critical transitions exhibit a divergence in their 'susceptibility', that is, in their sensitivity to minute details of their internal structure and to external perturbations. This self-sensitivity in the vicinity of a critical transition is manifested in universal computers as their 'programmability'.

Langton's approach has been critiqued in two ways: its dependence on the application of an order parameter (λ) and a lack of discussion of the intrinsic computation capability. To some extent, James Crutchfield's approach attempts to address these issues.

2.4.3 Crutchfield and Intrinsic Emergence

Crutchfield is interested in examining 'intrinsic emergence' – emergence of coordinated behaviour the existence of which does not merely rely on the presence of an external observer. Instead, intrinsic emergence is accompanied by changes in the (intrinsic) computational capability of the system. To analyse this, Crutchfield and Young (1990) use machine complexity theory based on the construction of ε-machines that are essentially stochastic automata of minimal computational power yielding a finite description of a data stream.[19] In other words, this approach entails the reconstruction from a given physical process of a computationally equivalent machine. The quantification of the amount of information processing at phase transition is given by the difference in the complexities of ε-machines above and below the transition, i.e. from periodic (solid) state to chaotic (gas) state. Aside from being able to analytically demonstrate the maximal complexity in ε-machines at phase transition, several interesting side results are obtained. First, the ε-machine at phase transition has an infinite number of states.[20] Second, there is self-similarity in machine structure at phase transition. Third, the choice of machine class or language becomes an important issue (i.e. Chomsky's hierarchy).[21]

In a later paper, Crutchfield (1994) discusses the dynamics of emergence in terms of upward movements in the hierarchy of computational model classes. Such changes are made possible by factoring out regularities (or equivalent classes) in the state transition structure within lower-level models. It is this reconstruction process that provides the dynamics of change in information processing structure.[22]

> The essential idea of moving up the hierarchy is that the symmetries assumed by the agent are broken by the data when reconstruction leads to an infinite model at some level of representation ... The key step to innovating a new model class is the discovery of a new equivalence relationship.

This interpretation provides a more elaborate definition of emergence:[23]

> A process undergoes emergence if at some point the architecture of information processing has changed in such a way that a distinct and more powerful level of intrinsic computation has appeared that was not present in earlier conditions.

Crutchfield admits that his dynamic computational model of emergence is not fully endogenous. Without such a model, it is difficult to understand the dynamics of emergence. For example, under what circumstances would a system exhibit emergence? It is not well understood whether emergence is always accompanied by universal computation. As Crutchfield (1994) noted, there is absolutely no general need for high computational capability to be near an 'edge of chaos'. Additional insights on these problems may be obtained from the study of self-organised criticality.

2.5 Self-organised Criticality

The search for endogenous and dynamic models of emergence leads directly to the theory of self-organised criticality (SOC). SOC is a study of systems in which short-range (local) interactions lead to the formation of complex and coherent structures (i.e. long-range spatial and/or temporal correlations) without the tuning of any external parameter.[24] SOC combines two notions, namely self-organisation and criticality. Jensen (1998, pp. 2–3) explains these notions in the context of thermodynamic systems:

> Self-organization has for many years been used to describe the ability of certain non-equilibrium systems to develop structures and patterns in the absence of control or manipulation by an external agent ... The word *criticality* has a very precise meaning in equilibrium thermodynamics. It is used in connection with phase transitions. When the temperature of the system is precisely equal to the transition temperature, something extraordinary happens. For all other temperatures, one can disturb the system locally and the effect of the perturbation will influence only the local neighborhood. However, at the transition temperature, the local distortion will propagate throughout the entire system. The effect decays only algebraically rather than exponentially. Although only 'nearest neighbor' members of the system interact directly, the interaction effectively reaches across the entire system. The system becomes critical in the sense that all members of the system influence each other.

From the above description, the SOC literature is clearly consistent with the work of Wolfram, Langton, and Crutchfield in at least one important aspect, namely its emphasis on the link between emergence and phase transition. Scale invariance, an aspect of emergence emphasised in Crutchfield's (1994) later work, is also an important characteristic of self-organised critical systems. In scale-invariant critical systems, the distribution functions describing the emergent spatial and temporal structures exhibit power laws in the following form:[25]

$$f(x) = cx^{\alpha}$$

where c and a are constants. The scale-invariant property of power laws can be illustrated by the following derivation of the relative change in the distribution function of x (which is shown to be independent of x):[26]

$$\frac{f(kx)}{f(x)} = \frac{c(kx)^{\alpha}}{cx^{\alpha}} = k^{\alpha}$$

Another term that is often used to refer to scale invariance is self-similarity. Statisticians use the more technical term of stability.[27] A variety of real-world phenomena have been found to exhibit power law behaviour. They include the distribution of city size, word frequency, and energy releases in earthquakes.[28]

The SOC literature has not attempted to link SOC to universal computation. There are many potentially important questions that remain unanswered. For example, if not all emergent states exhibit universal computation, what makes an emergent system universally computational? Is scale invariance related to universal computation? (In other words, is an emergent system universally computational if it exhibits scale invariance?) There is also the deep question of time irreversibility. Within the context of SOC, what is the link between the power law distribution of stopping times (time-scale invariance) and time irreversibility?

2.6 Emergence and Universal Computation in Economics

2.6.1 Emergence in Economics: A Brief Tour

Economists familiar with nonequilibrium thermodynamics and evolutionary theory have attempted to find references to emergence in the works of the founding fathers of economics. The list includes Adam Smith, Alfred Marshall, Carl Menger, Joseph Schumpeter, Friedrich Hayek, and Armen Alchian. Different scholars, however, tend to emphasise different aspects of emergence. For example, Foster (1993) discusses Marshall's struggle with the irreversibility of time, while Kilpatrick (2001) focuses on Hayek's writings on the benevolent aspect of 'spontaneous order'.[29] Schmitz (2001) discusses Menger's work on the emergence of money as a medium of exchange. Most of these works are descriptive in nature, thus making them 'open' research programs (i.e. amenable to various interpretations). More recent work on emergence in economics continues to be as varied in terms of both methodology and area of application.[30] Most of the recent literature on emergence is based on multiagent and/or evolutionary models.

Another related strand of literature is that which studies power law distributions in economics.[31] This line of literature does not study emergence per se, but the power law distributions could well take place within an emergent system, such as a hierarchy of cities.

2.6.2 Emergence and Universal Computation: Implications for Economics?

There are very few studies in economics that attempt to link emergence to universal computation.[32] Thus, to begin our exploration of this issue, it is perhaps useful to recall some of the key points from the previous review of the research programs of Wolfram, Langton, and Crutchfield. They are as follows.

1) Complexity can emerge in a system with local interactions.
2) A system can be conceived of as one that is capable of a range of dynamical behaviours (fixed-point, periodic, complex, chaotic) with the actual range determined by the rules of local interactions (e.g. number of neighbours, number of states, and number of transition paths).
3) While it may be possible to determine the capability of such a system to exhibit complex behaviour (Wolfram's class IV), it is not possible to determine when this will occur (undecidability?).
4) The computational capabilities of a system change as dynamical behaviour changes.
5) The emergence of complexity (class IV) is associated with universal computation.
6) Complexity is associated with optimal trade-off between information storage (order) and information production and transmission (randomness).

In what follows, we discuss what these points imply for economics.

Most of the work in the study of emergence involves local interactions between agents (components) that are bounded-rational in the sense that each agent follows a set of behavioural rules. This leads us to the question of the nature of economic agents.

If emergence could only result from 'simple' local (decentralised) interactions, and if we take 'simple' to mean bounded-rational, we have the following naive conjecture on the link between emergence and rationality.

Conjecture 1: Emergence is only possible in an economic system if agents are bounded-rational.

This is a very strong conjecture. Heiner (1983) provides an argument that appears to support it.[33] Heiner argues that the emergence of a complex economy implies greater uncertainty for economic agents. This is accompanied by rule-governed behaviour becoming more predictable. This is made possible by the emergence of new institutions that allow agents to increasingly use local information.[34]

> In general, further evolution toward social interdependence will require institutions that permit agents to know about successive smaller fractions of the larger social environment. That is, institutions must evolve which enable each agent in the society to know less and less about the behaviour of other agents and about the complex interdependencies generated by their interactions.

Our earlier discussions on the relationship between emergence and universal computation lead us to question the relationship between universal computation and bounded rationality. This issue is dwelt upon in a slightly different manner by Velupillai (1999), who postulates a link between rationality and computation universality

Theorem 1: Only adaptive processes capable of computation universality are consistent with rationality 'in the sense that economists use that term'.

This leads us to the following question. Does rationality 'in the sense that economist's use that term' correspond to bounded rationality? The notion of rationality in the former sense is clarified by Velupillai (2000). Here rationality is understood in terms of agents that are capable of choosing a set of maximal alternatives.[35] Elsewhere, Rubinstein (1998) notes that the two notions need not be incompatible. We leave this discussion for a future paper.

Perhaps a more useful approach to rationality would be one in which rationality exists in different degrees. This would certainly be consistent with Crutchfield's discussion of the possibility of agents with evolving computational capability. Heiner's (1983) work implies that the complexity of the agent and its environment are interrelated. One cannot speak of the complexity of an agent without at the same time discussing the complexity of the agent's environment. This is particularly true because the amount and type of information that are available locally and globally and their transmission mechanisms may change over time. This point becomes more important when the system in question is continuously perturbed. Pushing this line of thinking further, it is then plausible that rationality in the sense of 'adaptive processes capable of computational universality' is not something that exists all the time in the economy.

The discussion on information storage (memory) and information product and transmission in both Langton's and Crutchfield's work highlights the interdependence between the two aspects. Adaptation by agents occurs when memory is updated by new information that is transmitted to agents. Thinking in these terms provides opportunities to look at dynamical choice theory from a different perspective. The starting point would be to identify what represents memory in choice theory and what represents information production and transmission. We then proceed to ask why a trade-off between the two exists. This then leads us to examine in what sense this trade-off is optimised when we have universal computation.

2.7 Conclusion

More questions on emergence and universal computation have been raised than answered in this paper. The difficulties in formalising the notion of emergence via complexity measures such as Kolmogorov complexity is well known. The insight from the nature of such difficulties motivates a closer look at the computational capacity and structure of systems exhibiting emergence. Wolfram's conjecture on the relationship between emergence of complexity (class IV behaviour) and universal computation has spawned two lines of research on the relationship between emergence and universal computation. Langton focuses on the use of an order parameter (l) to map out the locus of emergent behaviour and explain its existence and links to universal computation. Crutchfield takes a different route by focusing on the intrinsic computation capacity and structure of emergence. These research programs on emergence and universal computation are not without weaknesses. Part of the problem with the study of emergence (and universal computation) is related to the problem that we cannot have a totally endogenous field of inquiry (to paraphrase Velupillai). The weakness of Langton's research program is its dependence on the order parameter (l), and the lack of endogenous dynamics in Crutchfield's research program are a reflection of this problem.

The weaknesses of Langton's and Crutchfield's research programs aside, they provide important insights to economists struggling to understand emergence and universal computation. An important insight is the interrelation between the complexity of economic agents and the complexity of their environment. What Crutchfield's hierarchy of machines or languages hints at is that rationality in the sense of 'adaptive processes capable of computational universality' is not something that is expected to exist all the time in the economy. The trade-off between information storage and information production and transmission may also offer significant new ways to look at economic institutions and processes.

Notes

1 The common practice is to use the term 'emergence' interchangeably with 'self-organisation' and 'spontaneous order'.
2 Darley (1994): 'A true emergent phenomenon is one for which the optimal means of prediction is simulation.'.
3 See Turing (1936, 1952).
4 By this, we mean the avoidance of the axiomatic method, along the lines of arguments put forward by Velupillai (2000).

5 For a more comprehensive treatment see Velupillai (2000).
6 The *Oxford Dictionary* defines 'spontaneous' as 'occurring as a result of a sudden inner impulse or inclination and without premeditation or external stimulus'.
7 The application of Kolmogorov complexity is merely an exercise in formally defining emergence. If the approach works, one still lacks a theory of emergence.
8 For an alternative approach using the busy beaver function, see Velupillai (2002).
9 Chaitin (2002) states this in the following manner: 'If you have *n* bits of axioms, you can never prove that a program is the smallest possible if it is more than *n* bits long'.
10 Quoted in Robertson (2000).
11 Badii and Politi (1997).
12 Sawhill (1995).
13 Wolfram (1984, p. 150).
14 Wolfram (1984, pp. 152–53).
15 Wolfram (1984, p. 155).
16 The term critical slowing down is used to describe this. See Langton (1992, p. 57).
17 Recall that the Shannon entropy *H* for a discrete process of *K* states is given by $H(A) = -\sum_{i=1}^{K} p_i log p_i$ where p_i is the probability of state *i*.
18 Langton (1992, p. 65).
19 This approach is analogous to Kolmogorov complexity in terms of using the minimal criteria. However, the ε-machines reconstruction approach measures the amount of information and how information is processed (and not just information per se, as in Kolmogorov complexity).
20 The growth of the machine size is positively correlated with the reconstruction cylinder size. See Crutchfield and Young (1990, p. 251).
21 It does appear that Crutchfield's work is inspired by Chomsky's hierarchy.
22 Crutchfield (1994, p. 31).
23 Crutchfield (1994, p. 49).
24 Interest in SOC began with the seminal paper by Bak et al. (1987) based on a sandpile model. Subsequent work has explored SOC with other models, such as earthquake models, lattice gas models, and forest fire models. For a comprehensive survey of SOC see Jensen (1998). Introductory treatments of SOC include Bak (1996) and Sawhill (1995).
25 Note that not all power law behaviours are related to SOC. See Sornette (2000, p. 322).
26 Jensen (1998, p. 4, fn. †).
27 Stable distributions such as the Gaussian and Paretian distributions have the property of exchangeability. McCall (2004) defines 'exchangeability' in the following manner. The random quantities x_1, x_2, ..., xn are called finitely exchangeable if every permutation of the finite set (x_1, x_2, ..., xn) has a joint distribution identical with every other permutation. Hence, if *F* is the joint distribution, $F(x_1, x_2, \ldots, x_n) = F(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)})$ for all permutations π on the set (1, 2, ..., *n*).
28 For a readable treatment of power laws see Schroeder (1991).
29 See Foster (2000) for Schumpeter and De Vany (1996) for Alchian.
30 The literature is obviously large and growing. Owing to time and space limitations, we can only list some of these in the bibliography for the reader to follow up. They include Schelling (1978), Arthur et al. (1987), Sugden (1989), Lesourne (1992), Bak et al. (1993), Scheinkman and Woodford (1994), Vriend (1995), Krugman (1996), Axtell (1999), and Axtell et al. (2001).

31 The seminal contributions in this area come from Herbert Simon and Benoit Mandelbrot. These are collected in Mandelbrot (1997) and Ijiri and Simon (1977). More recent contributions come from a broad range of scholars, including William Brock, Jean-Philippe Bouchaud, Arthur De Vany, Masahisa Fujita, Xavier Gabaix, Paul Krugman, and Didier Sornette.
32 A rare exception is Peter S. Albin's work. The author thanks Håkan Holm for pointing this out.
33 The author is grateful to Håkan Holm for pointing out this reference.
34 Heiner (1983, p. 580).
35 Velupillai (2000, pp. 35–36).

References

Arthur, W., Ermoliev, Y., and Kaniovski, Y. (1987). Path-dependent processes and the emergence of macrostructure. *European Journal of Operational Research* 30: 294–303.

Axtell, R. (1999). *The Emergence of Firms in a Population of Agents, Technical Report, Center on Social and Economic Dynamics*. Brookings Institution.

Axtell, R., Epstein, J., and Young, P. (2001). The emergence of classes in a multi-agent bargaining model. In *Social Dynamics*, (ed. S. Durlauf and P. Young), Cambridge, MA: MIT Press.

Badii, R. and Politi, P. (1997). *Complexity: Hierarchical Structures and Scaling in Physics*. Cambridge: Cambridge University Press.

Bak, P. et al. (1993). Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche* 47: 3–30.

Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of $1/f$ noise. *Physics Review Letters* 59: 381–384.

Chaitin, G.J. (2002). Computers, paradoxes, and the foundations of mathematics. *American Scientist* 90: 164–171.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory* 2: 113–124.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control* 2: 137–167.

Crutchfield, J. (1994). The calculi of emergence: Computation, dynamics and induction. *Physica D* 75: 11–54.

Crutchfield, J. and Young, K. (1990). Computation at the edge of chaos. In: *Complexity, Entropy and the Physics of Information*, (ed. W. Zurek), Cambridge, MA: MIT Press.

Darley, V. (1994). Emergent phenomenon and complexity. In: *Artificial Life IV*, (ed. R. Brooks and P. Maes.), Cambridge, MA: MIT Press.

De Vany, A. (1996). Information, chance, and evolution: Alchian and the economics of self-organization. *Economic Inquiry* 34: 427–443.

Foster, J. (1993). Economics and the self-organisation approach: Alfred Marshall revisited. *Economic Journal* 103 (419): 975–991.

Foster, J. (2000). Competitive selection, self-organisation and Joseph A. Schumpeter. *Journal of Evolutionary Economics* 10: 311–328.

Heiner, R. (1983). The origin of predictable behavior. *American Economic Review* 73 (4): 560–595.
Ijiri, Y. and Simon, H. (1977). *Skew Distributions and the Sizes of Business Firms*. Amsterdam: North-Holland.
Jensen, H. (1998). *Self-Organized Criticality*. Cambridge: Cambridge University Press.
Johnson, S. (2001). *Emergence: The Connected Lives of Ants, Brains, Cities and Software*. New York: Scribner.
Kilpatrick, H.E. (2001). Complexity, spontaneous order, and Friedrich Hayek: Are spontaneous order and complexity essentially the same thing?. *Complexity* 6 (3): 16–20.
Krugman, P. (1996). *The Self-Organizing Economy*. Oxford: Blackwell.
Langton, C. (1992). Life at the edge of chaos. In: *Artificial Life II*, (ed. C. Langton. et al.), Addison Wesley, Redwood City, CA.
Lesourne, J. (1992). *The Economics of Order and Disorder*. Oxford: Clarendon Press.
Mandelbrot, B. (1997). *Fractals and Scaling in Finance*. Berlin: Springer.
McCall, J. (2004). Induction: From Kolmogorov and Solomonoff to deFinetti and Back to Kolmogorov. *Metroeconomica* 55 (2–3): 195–218.
Robertson, D. (2000). Coedel's theorem, the theory of everything, and the future of science and mathematics. *Complexity* 5 (5): 22–27.
Rubinstein, A. (1998). *Modeling Bounded Rationality*. Cambridge, MA: MIT Press.
Sawhill, K. (1995). Self-organized criticality and complexity theory. In: *Lectures in Complex Systems*, (ed. L. Nadel and D.L. Stein). Redwood City, CA: Addison Wesley.
Scheinkman, J. and Woodford, M. (1994). Self-organized criticality and economic fluctuations. *American Economic Review* 84 (2): 417–421.
Schelling, T. (1978). *Micromotives and Macrobehavior*. New York: W. W. Norton.
Schmitz, S. (2001). Carl Menger's money and the current neoclassical models of money. Forschungsstelle für Institutionellen Wandel u. Europ. Integration, IWE. University of Vienna.
Schroeder, M. (1991). *Fractals, Chaos, Power Laws*. New York: W. H. Freeman.
Sornette, D. (2000). *Critical Phenomena in Natural Sciences*. Berlin: Springer.
Standish, R. (2001). *On Complexity and Emergence*. University of New South Wales.
Sugden, R. (1989). Spontaneous order. *Journal of Economic Perspectives* 3 (4): 85–97.
Turing, A. (1936). On computable numbers with an application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society*. 2. 230–265.
Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B*. 237. 37–72.
Velupillai, K. (1999). Undecidability, computation, universality and minimality in economic dynamics. *Journal of Economic Surveys* 13 (2): 653.
Velupillai, K. (2000). *Computable Economics*. Cambridge: Cambridge University Press.
Velupillai, K. (2002). *Perplexed in the Tangled Roots of the Busy Beaver's Ways*. University of Trento.
Vriend, N. (1995). Self-organization of markets: An example of a computational approach. *Computational Economics* 8 (3): 205–231.
Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D* 10: 1–35.

3

Fractals and Geography

Pierre Frankhauser and Denise Pumain

University of Franche-Comté, France
University Paris Pantheon-Sorbonne, France

Original publication details: Frankhauser, Pierre, and Denise Pumain. (2007). Fractals and geography. In: Sanders, Lena (ed.), Models in Spatial Analysis. © John Wiley & Sons Ltd, pp. 281–300. First published in France in 2001 by Hermes Science/ Lavoisier entitled Modèles en analyse spatiale. First published in Great Britain and the United States in 2007 by ISTE Ltd. doi.org/10.1002/9780470612255.ch10
Reproduced with permission from John Wiley & Sons Ltd.

3.1 Introduction

Is the craze for fractals only a fad? Over the last two decades, fractals have fascinated us with the images that they have produced or reproduced but, in social sciences applications, they have fallen short of the expectations of researchers. The measures that have been generated by them have not established themselves alongside more traditional indicators and their use in dynamic models of complex systems has proven to be difficult to implement. It is natural though that they still inspire research in geography, in particular spatial analysis. The irregular and fragmented forms of relief or urban patterns, the ramifications of hydrographic or transport systems, and the hierarchised structures of the world's territories and city systems all have properties, and fractal analysis could propose new interpretations. The self-similar morphology of fractal objects, reproducing the same structures at different scales, is an important feature of the spatial organisation of several geographical objects. This essential property has been used in explanatory theories of hierarchical systems, as with central place theory for city systems, but with spatial models that were based on traditional geometry. Introducing fractal geometry as a reference in geographical models is therefore a way to demonstrate certain specific processes of the spatial organisation, particularly cities and systems and to find new expressions, especially for dynamic interpretations.

3.2 Fractality and Structuring of the Geographical Space

The spatiality of human societies is a paradox. On the one hand, when societies humanise and space the land's surface (Pinchemel and Pinchemel 1988), their development

Machine Learning and the City: Applications in Architecture and Urban Design,
First Edition. Edited by Silvio Carta.

generates homogeneity on the land's rough surface in terms of farming and lifestyle, for example, but also in a more general way, in terms of circulation conditions, at least within the boundaries of each large territorial political system. On the other hand, even within a given culture, the human implementation creates new disparities, in particular in terms of spatial distribution of the population and activities that increase as urbanisation progresses. The centre–periphery duality established itself and became more pronounced with the development of cities, as well as with the expansion of the interactions which contribute to the sizing of territories actually interacting. Its cumulative construction results in an increasingly strong hierarchisation of the area, depending on its degree of social complexity (Pumain 1997).

The fundamental heterogeneity of the humanised space is discernible at different observation levels, with a reproduction of similar structures at several organisation levels. These level nesting effects have often been reported. The self-similarity of the geographical space has been interpreted by Philbrick (1957), for example, as the result of an almost systematic alternation of polarisation effects, which differentiate a centre and its periphery, and the similarity effects, which define homogeneous regions at all geographical scales: homogeneous parcels of an operation are polarised by a farm, the union of several operations forms a small homogeneous agricultural region polarised by a market town, several agricultural regions form the periphery of a small regional capital, and the capital itself becomes, with others of its kind, a region polarised by a major city, etc. This functionalistic and static interpretation must be enhanced by the explanation of the reach of spatial interactions (Grasland 1999), as well as of the speed of movements between regions at the different scales and their historically differentiated evolution (Bretagnolle 1999).

3.2.1 Density: A Traditional but Unsuitable Measure

The measure of human presence most widely used by traditional geography is the measure of density. Compared with usual social indicators that report incomes, services, etc., for an individual, this measure is specific to geography since it indicates quantities on a surface and puts the population at the numerator and not at the denominator of the measure. In reality, population density, long considered the geographical indicator of choice, measures what we would call today the ecological capacity of territorialised socioeconomic systems. Density is an analogue measure, in its construction, to performance indexes. As with hundredweight of wheat per hectare, we count people per square kilometre. In this capacity, the conceptual efficiency of the notion of density is limited to the situations where there is an actual ecological relation between a surface and the population that develops the resources, and between the land and the human mass that it supports. This relation makes sense in the characterisation and comparison of agrarian economies. In the case of cities, which by definition do not practise agricultural activities, the connotation of productivity that the density indicator carries does not apply.

Besides, density necessarily refers to a situation of homogeneity. In physics/ chemistry, from which this measure is borrowed, density is a measure particular to a homogeneous distribution of particles for organisms in balance in self-contained systems. In geography, it can only help in characterising relatively homogeneous spatial systems, such as, for example, regions where the level of a specific activity would be uniform. By proposing point maps of town population quantities, George (George 1950) had already criticised the inability of density maps established at this level to explain the heterogeneity of the distribution of the population in its adjustments to local surface textures.

The insufficient density measures are completed by disparity indexes of surface quantity distribution conducted on a given pattern of areal subdivisions (a regular quadrat or any irregular grid). Most of these indexes, such as those based on concentration curves, measure the deviations from an equidistributional situation (Bretagnolle 1996). This model implicitly assumes a linear type relation of proportionality between population and surface. However, this relation is almost never verified: when it is empirically determined, the relation between population and the surface of units of certain administrative sections often assumes the form of a power law with an exponent that is lower than 1, in general around two-thirds (Haggett 1973). In other words, the most populated units are the smaller areas, or the density decreases as the size of the administrative units considered increases. It is due to these systematic variations that the concentration measures give different results depending on the geographical level at which they are measured (Isard 1960). In this respect, Le Bras (1993) rightly denounces the so-called space occupation intensity measures: '80% of the population lives on 20% of the land', which do not specify the aggregation level of the observations (in this case, for example, we would say that a fifth of the most populated French communes hold four-fifths of the total population). Brunet and Dollfus (1990) carry the image to the absurd by emphasising the arbitrary character and the problem with this notion of space reference by saying that 'to the standards of the Parisian subway, the whole world population could be contained in the Territory of Belfort'.

The measures that refer to a homogeneous distribution model lose much information by forgetting the systematically heterogeneous character of geographical distributions and especially by not integrating the knowledge acquired about the general form of these disparities, which are always distributed in a geometric progression. We give two examples relative to urban localisations. At two observation levels, that of the city and that of urban network levels, we measure a fundamental heterogeneity of the spatial distribution of urban mass indicators (people, built-up surfaces, activities, or flows) in the considered surfaces. At the level of a city, Clark (1991) has shown for a long time that the distribution of resident population densities, as that of the land rents and real estate costs, is organised with a strong gradient decreasing from the centre to the periphery, based on a model of exponential or negative power law of the distance from the centre. This model, which is still relevant for costs, maintains its descriptive power, even if the competition exerted by tertiary jobs for the more accessible central locations leads to the formation of a central 'crater' in the hyperbolic cone, representing population densities in three dimensions. This fundamental model is not called into question by the recent evolution which, owing to residential relaxation of the centres on the one hand and of the densification of suburbs and the dispersion of outlying suburbs on the other hand, has considerably decreased the gradients of density distributions in most of the large cities in the world and in Europe, even in small cities with as little as 20 000 residents.

At the urban networks level, the territory is unequally occupied by the extremely hierarchised system of cities. There again, the most widely used models for the analysis are not the most adapted, because they refer to the notion of uniformity. The measures aiming to test the form of the distribution of the spread of the cities in a territory have used a Poisson distribution as reference (Dacey 1967), which presumes an equal probability of occupation by cities and does not consider the effects of the accumulation characteristic of urban systems. Similarly, the regular hexagon models of Christaller (1933) take into account a hierarchy of the sizes of the cities, but not the disparity of the resulting density. Recently it has been shown that fractal geometry makes it possible to modify Christaller's model, by articulating two spatial systems: one consists of urbanised areas which are concentrated along transportation axes and the other one is a hierarchical axial system of non-built-up, rural zones (Frankhauser 2005). Finally, the attempts to use spectral analysis to characterise the scale components corresponding to the different levels of the hierarchical organisation of urban networks (Dacey 1967; Cauvin et al. 1985) have come up against the major irregularities of this organisation.

3.2.2 The Fractals: References Adapted to the Space of Human Societies

Knowing that the distribution of urban density, regardless of the level considered, is never homogeneous, it can be interesting to replace the model of density by a fractal reference that would contain from the outset the information relative to the heterogeneity and to the form that it most often takes. Not only would it be possible to directly compare degrees of heterogeneity or to integrate this property into models, but we could also hope to discover something new instead of treating hierarchisation phenomena as residuals with respect to a homogeneous model. Besides, if the space occupation by cities is similar to the images produced by models of reference using fractal geometry, we could then try to understand why it is so by imagining plausible processes that simulate the genesis of such configurations. Plausible in this case means compatible with the urban theory, or rather the urbanistic or socioeconomic theories of the formation and growth of cities. This reference to the fractal model has the advantage over the density model to return more directly to a dynamic conception (Pumain 2004).

This conceptual evolution moves the centre of interest with respect to densities. It is not so much the intensity of the occupation of space that will be considered but its structure, built from an underlying ordering principle, which is represented at different geographic levels and concerns the connection of these levels. This implicit ordering principle can be revealed despite the disparities caused by random fluctuations. For instance, in the hierarchy of cities, the absence of discontinuity between dimensional or functional levels seems instead to be the rule and this contrary to what central place theory predicts. At a finer scale, and at least in what concerns the built-up surfaces, a discontinuity seems to persist between the continuous constructed space of cities and their periphery not yet belonging to the urban cluster (Frankhauser 1993). A recent demonstration of this dual fractal structure of the urban field has been made for European urban areas described with Corine Land Cover data on built-up land use by Marianne Guérois (2003 and 2008). These observations could help theoretically justify the use of multifractals by calling for different genetic processes between those that organise the space occupation at the level of the city and those that structure the spatial thread at the territorial scale of city systems.

Finally, the traditional analysis methods place the population or the activities with respect to a space support containing them and whose properties are those of Euclidean geometry. For a long time now, Harvey has addressed the fruitfulness of the conception of a relative space that is defined by the historical and social practice: 'It is the activities and the objects themselves which define their spatial field of intervention' (Harvey 1969, p. 209). Authors such as Hägerstrand, Cauvin, or Muller have declared its non-Euclidean character, deliberately heterogeneous, and anisotropic, in analyses of space perception, or in the research of cartographic representations more adapted (Tobler 1979; Rimbert 1986) or still in the research on theoretical geography on the properties of geographical space (Brunet and Dollfus 1990). In intra-urban space we have noted that the reference to a homogeneous space does not really fit with distributions that are extremely contrasting and organised into very strong gradients, which express and infer simple or multiple, highly polarised centre–periphery fields. Similarly, at the level of city systems, we have observed that the geometric models that come from central place theory seem incompatible, owing to their reference to a homogeneous distribution of population, with the existence of polarisation fields which are defined by the methods of space occupation by the cities (Frankhauser 1993; François 1996). Besides, we have been noticing for a long time that the distributions of urban hierarchies were generally well represented by statistical models – Pareto model, still named 'rank size distribution' by Zipf (Brakman et al. 1999), or lognormal distribution studied by Gibrat (Bee et al. 2017) – which seem compatible with fractal geometry.[1] Thus, the multifractal generators used by Le Bras (1993) to simulate the spatial distribution of demographic growths, which are not explicitly linked in his book to territorial processes, actually present an analogy with a stochastic process of growth distribution, of which we know that it generates lognormal distributions (Pumain 1997).

In geography, the most important fractal geometry applications involve the morphology of cities (Batty and Longley 1994; Frankhauser 1993) are only a few of the major studies) or, more broadly, the organisation of the population on a territory (Arlinghaus 1985; Le Bras 1993, 1996). The configuration of transport networks was also the subject of descriptions with the help of fractals (Benguigui and Daoud 1991; Genre-Grandpierre 2000). The work of Dauphiné (1995) is one of the first to have proposed a panorama of applications in geography, including models in physical geography. A recent application in hydrology was presented by Hauchard et al. (1999).

First, we will present a few types of fractal objects, as well as the measures used to characterise them, before examining some applications in detail. These applications were most often chosen in urban geography because they might enable us to specify the interpretation of fractal structures in geography.

3.3 Fractal Models of Spatial Structures

Fractal geometry enables us to analyse a spatial structure from a reference other than Euclidian geometry. Different types of 'ideal' fractal objects serve as reference for the description of an observed reality. These constructed fractals play a role similar to that of basic figures like the circle, square, etc., in Euclidian geometry. These objects are differentiated according to criteria inherent to fractal geometry and we introduce specific measures that will characterise these following a fractal logic.

3.3.1 Surface Models

Figure 3.1 shows several of these theoretical models used for applications in geography. These objects are obtained by repeating a specific operation called generator. The first example is the teragon (see Figure 3.1a). The initial figure is a square the length of $l_0 = L$. The generator operates on the perimeter: each side is replaced by a polyline made up of $N = 8$ elements of length $l_1 = r\,L$, where r is the reduction factor. In the case of the teragon, $r = ¼$. We verify that the total surface of the object remains invariable. This procedure is repeated for each element of length l_1 during the next step. In this way, the border of the object lengthens, becomes more complex, and finally tends towards infinity, whereas the surface remains constant. Such a behaviour does not exist in Euclidean geometry, the perimeter having lost the usual characteristics of a line.

Two other examples are shown in Figure 3.1. In both cases, the generator reduces the initial figure, a square, by a factor of r and we place N of these reduced reproductions, according to a chosen diagram, inside the initial figure. These copies must not intersect. For the Sierpiński carpet (see Figure 3.1b), we obtain a chequerboard where the squares touch each other by their tops, whereas in the Fournier dust (see Figure 3.1c) the squares are isolated. By repeating this operation for each of the N elements generated, we can verify that in both cases the surface of the fractal object approaches zero, whereas the length of the perimeter diverges. The Fournier dust is finally made up of a set of points distributed unevenly: they form masses separated by gaps which are generated during the iteration

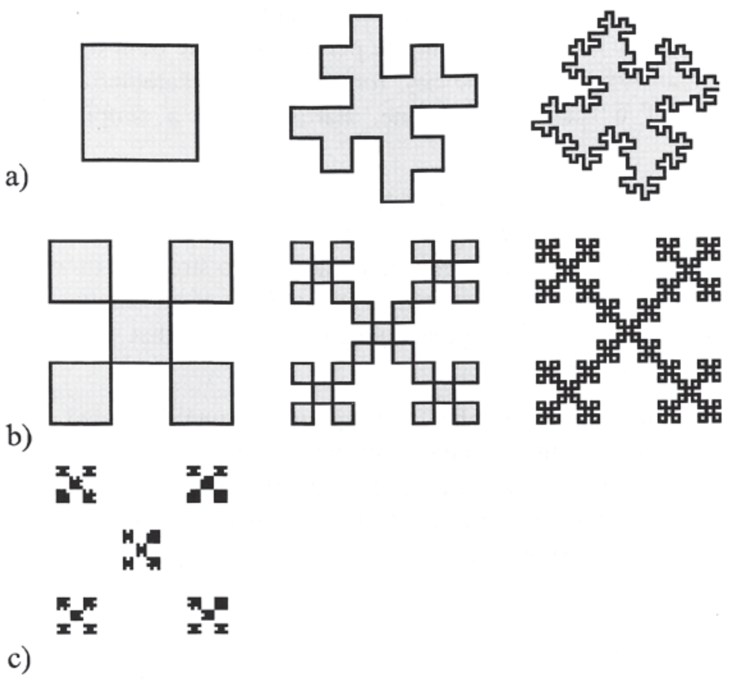


Figure 3.1 Fractal models of surface occupation.