PABLO JENSEN

YOUR LIFE IN NUMBERS:

MODELING SOCIETY THROUGH DATA

Springer

Your Life in Numbers: Modeling Society Through Data

Pablo Jensen

Your Life in Numbers: Modeling Society Through Data

Springer

Copernicus Books is a brand of Springer

Pablo Jensen
CNRS
École Normale Supérieure de Lyon and IXXI
Lyon, France

ISBN 978-3-030-65102-2 ISBN 978-3-030-65103-9 (eBook)
https://doi.org/10.1007/978-3-030-65103-9

This Copernicus imprint is published by Springer Nature
The registered company is Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 Introduction 1
Going Further 3

2 Three Simple Models 5
2.1 Do Segregated Neighbourhoods Imply Racist Residents? 5
2.2 Economic Competition on the Beach 7
2.3 Social Avalanches 8
Going Further 10

3 What Can Be Learnt from Simple Models? 11
Going Further 13

4 Reality Check for Simple Models 15
4.1 The Physics of Elections 15
4.2 Mathematical Smoke Curtains 17
4.3 The Shape of Cities 18
Going Further 20

5 A Physical Simple Model 21
Going Further 23

6 What Are Simple Models Worth For? 25

7 Complexity: Step by Step 27
7.1 Ants' Nests 27
7.2 Framed Human Action: Car Traffic and Pedestrians 29
7.3 What Are These Models For? 31
Going Further 32

8 Complex Models to Understand Complex Social Situations 33
8.1 Predicting Economic Growth 33
8.2 The Future Is Like the Past 35
8.3 Modelling the Job Market 37
Going Further 38

9 Modelling Epidemics 39
9.1 Individual Models 39
9.2 Brute-Force Big Data 41
9.3 Controlling Complexity 41
Going Further 43

10 Why Weather/Climate Forecasts Can Be Trusted 45
10.1 Predicting the Weather 45
10.2 A Tamed Virtual Earth 47
10.3 Predicting the Climate! 48
10.4 Why Did Climatologists Succeed? 49
Going Further 50

11 We Are Not Social Atoms 51
Going Further 55

12 Social Data Are Soaked by Social Complexity 57
12.1 Careful with Data! 57
12.2 Data on Social Regularities 58
12.3 Rich Data, Needs Only (Conceptually) Simple Modelling 59
Going Further 60

13 Machines that Learn How to Model 61
13.1 Neural Networks 61
13.2 World Chess Champion 63
13.3 Predicting the Success of Tweets 64
13.4 Predicting Life Trajectories 67
13.5 Computer Vision 67
Going Further 69

14 Starting from Data to Hunt Causes 71
14.1 Why Is This Happening? 71
14.2 Complex Causality: Whimsical Causes and Effects 72
14.3 Farewell to the Analytical Method? 74
Going Further 75

15 A Moral Thermometer? 77
15.1 How Hot Is It Today? 78
15.2 The Construction of the Scale 79
15.3 An Absolute Temperature 80
15.4 Temperature as an Institution 81
15.5 Building the Moral Thermometer 82
15.6 The Perceived Temperature 83
Going Further 84

16 Where Do Indicators Lead Us? 87
16.1 Compute the Profitability of Public Projects 87
16.2 Ranking Universities 88
16.3 Evaluating Researchers 89
16.4 The Origin of the Indexes 90
16.5 Indexes Are Not Perfect. 91
16.6 But Can Be Useful! 91
16.7 Where Does This Indicator Lead Us? 92
Going Further 93

17 Which Numbers for the Future? 95
17.1 Social Numbers Were Invented by Centralized Powers. 95
17.2 Lost Individuals 96
17.3 Markets to Replace Planning 98
17.4 Care for Public Numbers 99
17.5 Planning as a Common Destiny. 100
Going Further 102

18 Conclusion 103

Acknowledgments 105

Annexes 107

Chapter 1
Introduction

I'll always remember my 54th birthday. That very day, the French president announced an unprecedented decision: a general lockdown to prevent the spreading of COVID-19. A few weeks later, half of the world population would have to remain closed at home. How did governments justify such drastic measures, which took the world back to the Middle Ages? The short answer is a number, 500,000, as the number of deaths predicted by a mathematical model if the government took no action.

Epidemic models are but one example of the major role played by mathematical models in modern societies. The social and political importance of being able to create a virtual world, which can be manipulated at will, to see, beforehand, the consequences of our actions, is obvious. What is the effect of closing schools on COVID-19 propagation? What are the economical consequences of creating a carbon tax? The point, of course, is whether the answers provided by models can be trusted. Mathematical models draw great legitimacy from physicists' past achievements. Newton's deep understanding of planetary motion lead to startling forecasts. His equations allowed to build a realistic virtual scale model of the solar system. Using it, astronomers could predict not only future eclipses but also the very existence of an unknown planet, Neptune, which is billions of kilometres away from the Earth. The fascination with mathematics' successes has led some scientists to try to apply the same approach to social issues, and this trend has exploded with the recent digital revolution, which brought powerful computers and lots of social data. Some scientists, mainly coming from physics and computer science departments, fantasize about creating virtual societies in their machines, through a myriad of interacting virtual robots.

For example, the *FuturICT* project proclaimed in 2012: "Many problems we have today – including social and economic instabilities, wars, disease spreading – are related to human behavior, but there is apparently a serious lack of understanding regarding how society and the economy work… Combining complexity theory and social data analysis, *FuturICT* will develop a new scientific and technological

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_1

approach to govern our future”. The project managed to slip into the six finalists to win €1 billion in European funding. It proposed to build a “Living earth simulator”, powered by a “planetary nervous system”, a worldwide network of sensors recording and centralizing billions of individual and environmental data every second. These data would be fed into powerful computers and processed using “the hidden laws that underlie our complex society”. Such an ambitious approach has proved fruitful in physics. Thanks to their knowledge of the laws governing atoms, physicists have built virtual crucibles that make it possible to explore, quickly and at almost no cost, the properties of original materials. They can thus imagine original alloys and test whether they are capable of transforming CO_2 into fuel, which could solve two environmental problems at once. But can this approach be extrapolated to society?

My double experience as practicing physicist and social scientist allows me to critically review recent scientific contributions from the emerging field of computational social science. After reading this book, you’ll understand how numerical models work and how they can help governments to take tough decisions. We’ll first explore several simple models, which are easy to understand, but contain the essence of more complex simulations, such as those that deal with epidemics, or the economy. To understand why the numbers spit out by complex social models are generally unreliable, we’ll compare them to the virtual Earth built by climate scientists, which is able to provide climate predictions that hold out in the face of powerful opposing interests.

However, social numbers can be used in many other ways to build shared knowledge. Since the nineteenth century, statistics has developed a set of mathematical techniques to enable governments to analyse real data, in order to understand the causes of events, identify responsibilities and intervene. Are wage inequalities between men and women due to gender discrimination, or are they simply the result of differences in working hours or qualifications? How can researchers claim that “fine particles kill half a million people in Europe every year”, when none of these deaths is directly observable? Finally, it is possible to share knowledge about society by transforming a complex phenomenon into a number. The gross domestic product (GDP), the ranking of a university, the number of crimes solved by a police department or Amazon’s rating of a book only retains from reality certain aspects deemed to be relevant, in an attempt to construct an “objective” point of view. A detailed comparison of GDP, which represents a kind of “moral thermometer”, and physical temperature will allow us to understand why physical indicators are far more reliable than social ones. Of course, the ongoing digital revolution only exacerbates the role played by numbers, as these are transformed into a fundamental pillar of our social life. Global platforms have replaced oil companies in the top market capitalizations and are profoundly changing our economy and social organization. I will conclude this book by analysing the history of social numbers, and the present struggles for data control, and by wondering how number could be used for planning a common future, capable of mastering the ecological crisis.

Going Further

The presentation of the *FuturICT* project: "Communication and Information Technologies for the Future", http://www.futurict.eu/. All quotes are from this site, accessed in September 2012.

Chapter 2
Three Simple Models

2.1 Do Segregated Neighbourhoods Imply Racist Residents?

Let's start with a simple model proposed in 1969 by economist Thomas Schelling. He addressed a very sensitive question: How can we explain "segregation by colour" in the United States? More specifically, Schelling wanted to know whether it necessarily takes racist dwellers to create a segregated town. Our intuition suggests that the overall configuration of a city merely reflects, by aggregation, the characteristics of individuals. If people wished to live in mixed neighbourhoods, then districts should bring together people from different categories. To put it another way, does a segregated global state necessarily imply an individual willingness to segregate? Schelling guessed that "the interplay of individual choices [...] is a complex system with collective results that bear no close relation to individual intent". To test his hunch, Schelling proposed the following model. The city is represented by a chessboard, each square representing a dwelling, which can be occupied by a red or green agent, or be temporarily empty. We assume that all people have a clear preference for a mixed neighbourhood, with as many reds as greens. If someone has more neighbours of her colour, her satisfaction decreases, and it becomes zero if all the neighbours are of the opposite colour (Fig. 2.1). To complete the creation of his world, Schelling specified the dynamics. Each day, someone is chosen at random and is offered an empty dwelling also chosen at random. The person first calculates whether this move increases her satisfaction. If it does, she agrees to move; if not, she stays at her place. Then, another person and another empty square are chosen at random, and the process is repeated. Our intuition tells us that, since individuals move to improve their satisfaction, and since satisfaction is highest when neighbourhoods are mixed, the city should tend towards mixing. The interest of the simple virtual city invented by Schelling is to show that this implicit modelling is incorrect, and that the dynamic leads to a segregated city, where most inhabitants are dissatisfied.

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_2

Fig. 2.1 The light agent placed in the centre of the square has maximum satisfaction in the figure on the left (mixed neighbourhood), low satisfaction in the central figure (neighbourhood entirely of its colour) and zero satisfaction on the right (neighbourhood entirely of the other colour)

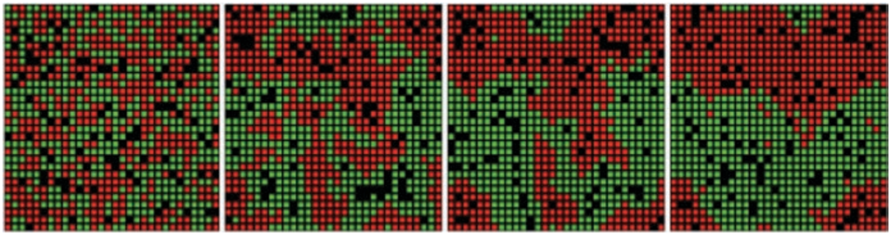


Fig. 2.2 Initially, the inhabitants are randomly distributed in the different squares. As they move to increase their satisfaction, the colours separate and the city ends up being totally segregated

Let's start from a city in which people are randomly distributed and run the simulation according to the rules described (Fig. 2.2). The city becomes more and more segregated, and this result is robust, as it is observed for any initial distribution of agents. A few years ago, we even managed to prove this with a few mathematical equations: after a few dozen moves of each person, the city inevitably segregates in two zones of different colours. How can we explain this collective race to disaster? What vicious mechanism happily leads all these people towards a segregated city where they are all… unhappy?

The explanation of the paradox is that when a person chooses to move, he only takes into account the variation of *his own* satisfaction, and not at all that of his neighbours. It is easy to understand that if a red inhabitant leaves a predominantly green neighbourhood to move to a mixed one, because his satisfaction increases, he penalizes all the other people who remain in his original neighbourhood, as it is now even more dominated by greens. In the same way, the satisfaction of the new neighbours decreases, because the neighbourhood, initially balanced, is now slightly red. The overall satisfaction decreases because the losses of the many agents affected by a person's move are not compensated for by the gain experienced by the mover. However, in this world governed by the economists' selfish individuals, this particular gain dominates the dynamic, leading inexorably to a social situation that harms everyone. It is however possible to change this result by forcing people to take into account the impact of their moves on their neighbors, by including the overall satisfaction variation into their calculations. These "altruistic" agents then reach an optimal result, i.e. a mixed city that maximizes overall satisfaction.

Schelling's segregation model is very useful. First, as all mathematical models, it forces us to make our assumptions explicit. Our intuition suggests that the observation of a segregated city implies the racism of individuals. By formalizing this somewhat vague idea, the model allows us to state that, logically speaking, this link is not rigorous. We can end up with segregated configurations even when all the individuals are looking for diversity. More generally, this model shows that economic agents pursuing their own interests may not achieve an optimal situation: the invisible hand is not that powerful. Finally, the small number of ingredients gives these models an undeniable pedagogical interest. One can understand their results and the causalities at work in depth, without being overwhelmed by the complexity of the real world. We understand that selfishness leads to awful situations because individuals take decisions without taking into account how their neighbours, old and new, feel about. We can study these effects in detail and show why they are stronger than the additional satisfaction obtained by the person moving, leading to an overall negative effect. In a nutshell, the collective effects of individual decisions are not always intuitive.

2.2 Economic Competition on the Beach

Let's analyse another simple model proposed in 1929 by economist Harold Hotelling. He also wanted to clarify a theoretical point, regarding the theory of competition "among a small number of entrepreneurs". More precisely, he wondered whether there exists an "undue tendency for competitors to imitate each other in quality of goods, in location, and in other essential ways", leading to too high product uniformity, at the expense of consumers. Hotelling's simple world is actually a beach, filled by bathers evenly distributed along the shore (Fig. 2.3). Two vendors, Paul and Susan, sell the same ice cream at the same price. Hotelling's question is: where should Paul and Susan locate their ice cream trolleys, for each to maximize his/her sales?

Since swimmers are supposed to prefer the closest seller, each seller tries to be closer than the other to the majority of buyers. For example, if Paul positions himself on the far right, Susan has to position herself just on his left, to catch most of the buyers. But then Paul can imitate her strategy and position himself just to her left… and so on until they both reach the middle of the beach (point B in Fig. 2.3). It can be shown mathematically that this position, where neither of the two sellers can improve their gain by moving, is the only stable location. Hotelling generalized this result into a "principle of minimal differentiation". When there is competition between two agents to share a market, there will be a tendency towards a high degree of product similarity in order to maximize sales. The beach can represent a more abstract space, that of product differentiation. For example, two candidates in the presidential election will try to position themselves in the centre to attract a maximum number of voters: if one is positioned on the left, the other candidate can position himself just to his right to attract a majority. Hotelling also discusses the

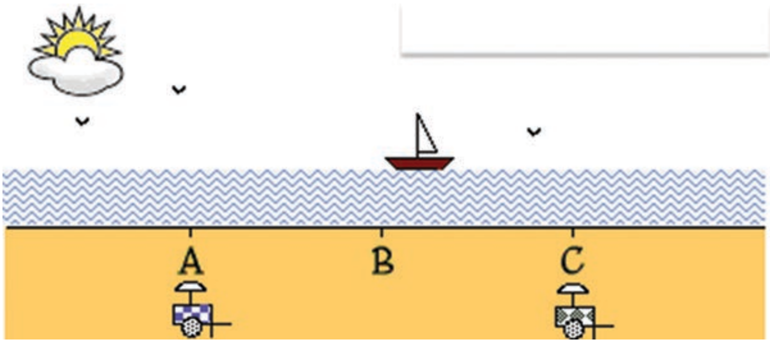


Fig. 2.3 Two vendors, shown here at positions A and C, try to find the locations that maximize their profits

sale, by two merchants located next to each other, of more or less bitter cider. Differentiation no longer takes place in the physical space, but in the space of consumer taste. If we assume that tastes are uniformly distributed on the bitterness scale, then both shop owners are lead to produce ciders of average bitterness. The principle is also valid in the case of non-uniformly distributed tastes: the average taste must be replaced by the "median" taste, i.e. the one that divides the market into two equal parts (equal number of consumers preferring the two options). Enthusiastic about the simplicity and apparent generality of its principle, Hotelling used it to explain "the tremendous standardisation of our furniture, our houses, our clothing…", and even as an "argument in favour of socialism against capitalism". Indeed, competition leads to homogenization, which is far from socially optimal. Take the case of ice cream sellers: competition drives both Paul and Susan to the middle of the beach, and all consumers have to walk that far to get their ice cream. Instead, if sellers could cooperate, they would locate at the first and last quarter of the beach (points A and C in Fig. 2.3). They would each retain half of the demand, but consumers would then be much closer to their seller. You can compute the average distance travelled by buyers and check that "capitalist competition" doubles the distance that consumers have to travel to get their ice-cream!

2.3 Social Avalanches

Let's explore a last example, before a general discussion about the value of simple models. In the two preceding cases, agents have fixed preferences, for neighbourhood color composition, or for minimizing travel distances. This is the general case for standard economic models: each agent has preferences for each of the goods and a budgetary constraint that forces him to make choices in order to maximize his overall satisfaction. This "atomic" economic agent only interacts with other agents via markets, which set the prices of the different goods, leading to a balance between supply and demand. Once these prices are known, each economic agent buys the goods in the quantities that maximize his well-being given his budgetary constraints, and everyone is satisfied – including the economist who finds his ideal world beautiful. Real social interactions are of course much richer than these interdependencies induced by price formation. Some physicists and economists are now proposing to complement the vision of standard economic theory by introducing these interactions, drawing inspiration from models of magnetism.

This strange analogy comes from the fact that the magnetism of many materials (such as iron) is created by the direct interaction between their atoms, which behave like small magnets. Let's start by considering an isolated iron atom. When a magnetic field is switched on, the atom tends to align with this external field, like a small compass. But physicists know that this model of an isolated atom does not account for the actual behaviour of magnets, which contain many interacting atoms. The key point is that an atom tends to imitate its neighbours, by aligning its compass with theirs. This explains the fact that, even in the absence of an external field, iron can become magnetic, as a result of the spontaneous alignment of all these small magnets. The interactions cause a collective effect, qualitatively different from the behaviour of an isolated atom. Indeed, as soon as, through some fortuitous fluctuation, a small group of atoms aligns into the same direction, their common alignment propagates to their neighbours, creating an avalanche throughout the magnet.

The analogy with economic models is through this imitation effect. Agents are endowed, in addition to their intrinsic preferences, a force that drives each of them to imitate the choices of their friends. If they have bought the latest model of mobile phone, I have a stronger preference to do the same. Adding imitation leads to collective behaviours that are qualitatively different from those observed in classical economic models. In addition to the imitative avalanches that result from the "snowball effect", direct interactions mean that agents' intrinsic preferences for goods no longer uniquely determine the overall state. Indeed, when the imitation force is stronger than the intrinsic preference, personal choice is dictated by that of others, uncoupling collective choices from the simple aggregation of intrinsic preferences. In the following example, we will see that this can lead to the existence of several outcomes, some of which are unfavourable to all agents, who have nevertheless chosen them collectively.

How does a restaurateur choose his prices? The standard theory suggests to find the price that maximizes profit. Since profit is given by the product of price and demand, and demand is generally assumed to gradually decrease as the price increases, the optimal price should be well defined. The gradual decline in demand is explained by the diversity of agents' preferences: some love the chef's kitchen and are willing to go at any price; others only if prices are low. Everything changes when interactions between customers are introduced. They may, for example, choose the restaurant where their friends go, even if they judge it intrinsically less preferable. Let's imagine, for example, that the restaurant owner chooses, at the opening, prices that are fairly low in relation to quality. This will attract a large number of people, who will influence others, leading to comfortable profits, and probably to a waiting list. Suppose the chef now takes advantage of this strong demand to raise its prices a little. In classic models, this will drive away the least motivated customers, causing demand to drop a bit. However, since the price is higher, profits will rise. As long as the price increase remains low, this is probably what will happen, even if imitation is taken into account. But if the price increases further, more customers may find that the restaurant has become too expensive. They stop going there and may be imitated by enough friends to generate an avalanche, dramatically reducing demand. Imitation now works the other way around! And even if the owner returns to the lower price that used to guarantee his success, this would not be enough to recover the profits, again because of imitation. It would be necessary to start again with the very low prices charged at the opening, to attract customers again and regain profits.

Going Further

The original Schelling paper: "Models of Segregation" The American Economic Review, **59** 488–493 (1969)

Our technical paper, which proves mathematically that selfishness leads to collective frustration: Grauwin *et al,* "Competition between collective and individual dynamics", *PNAS,* vol. 106, 2009, 20622.

Hotelling's original article: "Stability in competition" *The Economic Journal,* Vol. 39, 1929, p. 41–57.

On the restaurant owner's dilemma in setting prices: M. B. Gordon *et al,* "Seller's dilemma due to social interactions between customers" *Physica A,* vol. 356, 2005, p. 628–640.

On the modelling of interactions by economists: S. Durlauf and Y. Ioannides, "Social interactions", *Annu. Rev. Econ.* Annu. Rev. 2, 2010, pp. 451–478 and S. Ioannides, 2010. Durlauf, "How can statistical mechanics contribute to social science?" *PNAS* 96 (1999): 10582

Chapter 3
What Can Be Learnt from Simple Models?

These three simple models are pedagogically interesting, as they try to answer important questions, but remain easy to understand. Note first that they are essentially conceptual as, rather than trying to explain the causes of specific real situations, they aim to answer a theoretical question, to challenge our ways of *thinking* about reality. And they manage to do so, by improving our understanding of collective effects of individual choices. But as soon as one tries to use these simple models to understand real events, one runs into serious trouble, as we'll explore further in the next chapter. The reason is as simple as these models: to be conceptually transparent, they rule out many features. For example, Schelling's model greatly simplifies the reality of housing choices, leaving aside dwellings prices, school or commercial facilities, the presence of friends… Neglecting so many ingredients makes any direct comparisons with real segregation very dangerous. For example, a colleague from computer science wondered, in light of the robustness of segregation to the details of the model, whether real segregation is a "fatality". The restaurant example does not take into account important elements, such as advertising or localization, which play an important role in the success of a business. Moreover, the effects of direct influences are highly correlated with other factors. If I choose the same restaurant as my friends, it may be by imitation, but also because I share their tastes.

The essential merit of simple models is that they help us to improve our ways of thinking. They show that intuitions may prove wrong, and they enrich the range of models used to interpret reality. By integrating the idea that the economy is not necessarily in equilibrium, and that direct interactions between agents make it intrinsically unstable, economists would become less confident in market mechanisms. They could conceive that it does not take an external shock to derail inherently efficient markets, and that misnamed "crises" or speculative bubbles are in the very nature of real markets, or that the stock market value of a company does not necessarily reveal its "fundamental" value, which represents all available information. In standard economics, each agent, individually, translates these information

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_3

into a price that he is willing to pay for his shares, leading, collectively, to an objective and optimal market price. On the contrary, the imitative model suggests that agents copy others to make as much money as possible. Price changes would then reflect these imitative dynamics, rather than changes in fundamental value due to new information. The analysis of real stock price movements seems to support the imitative model, as these fluctuations are much larger than those expected in efficient markets. Long ago, economist John Maynard Keynes had an intuition of the importance of these mimetic processes: "Investment based on genuine long-term expectation is so difficult to-day as to be scarcely practicable. He who attempts it must surely lead much more laborious days and run greater risks than he who tries to guess better than the crowd how the crowd will behave; and, given equal intelligence, he may make more disastrous mistakes".

As noted before, models are useful to show that our intuition can be wrong: the state of the whole city is *not* obtained by simply aggregating the preferences of its dwellers for diversity; demand may change abruptly when prices change. But can simple models lead to general principles? The point is that a single counterexample can invalidate a principle, but examples can never prove it. For example, Hotelling's principle of "minimal differentiation" breaks down as soon as Paul and Susan are allowed to change the prices of their (otherwise similar) ice creams. Hotelling had considered this option and concluded that it did not fundamentally affected his principle. Fifty years later, other economists have shown that he had made a mistake, and that introducing prices leads to another "principle", of… *maximum* differentiation! Why does the inclusion of prices change the outcome so much? Firstly, because price competition leads to zero profits if sellers are in the centre of the beach. Indeed, if Paul decides to sell at 1$, Susan can undercut to 50cts, winning all the market, since her offer is more attractive for everyone. Then Paul will undercut in turn and so on until both prices are zero. To make profits, Paul and Susan should position far from each other, to increase their "market power" and raise prices. Let us suppose that they must first determine their location and then, at this position, compete through prices. This would be the case, for example, if, instead of walking around easily with their carts, they had to build a permanent shop. Game theory provides a solution to this problem by supposing that they both make exactly the same calculation, perfectly anticipating the results obtained by the other, that all consumers buy regardless of the distance to be covered… Only by accepting those rigid (and somewhat unrealistic) rules we can compute the positions and prices that lead to a maximal profit for both sellers.

The solution stems from the tension between competition for customers and market power. Sellers want to be close to the centre of the market, to attract the maximum number of customers by minimizing distances. But on the other hand, they would like to be far away from rivals, to reduce price competition. When the products are identical, market power dominates, leading sellers to place themselves at the ends of the beach to minimize price competition and maximize profits. For example, if Paul's store is located in the middle of the beach, Susan should open hers at one end of the beach to capture consumers located on her side, who will accept to pay something to avoid walking. This transportation cost, which increases with the distance to the seller, gives market power to traders. But Paul follows the same reasoning and chooses the other end of the beach, thereby increasing his prices and profits (and Susan's). Ironically, adding a further ingredient to the model can bring back the principle of minimum differentiation! For example, if ice creams sold by Susan and Paul appear different to consumers, because of their brand, price competition is diminished. Paul can offer slightly higher prices than Susan while retaining customers who prefer his products and are therefore prepared to pay a premium. To optimize their profits, calculations show that Susan and Paul should gather again in the centre of the beach…

Going Further

Hotelling's error is exposed in d'Aspremont *et al.* on Hotelling's stability in competition, *Econometrica,* vol. 47, 1979

On the analysis of stock market fluctuations: J. Ph. Bouchaud, "The endogenous dynamics of markets: Price impact and feedback loops", arXiv, 2009

Keynes quotation is from chapter xii of his General Theory of Employment, Interest and Money, 1936

Chapter 4
Reality Check for Simple Models

Simple models are useful to improve our thinking. Can they also help us understand real situations? Let's analyse two recent simple models whose explicit aim is to understand in detail the social mechanisms leading to two specific social puzzles: tight election results and the shape of cities.

4.1 The Physics of Elections

In the 2000s, several elections in Western countries ended up with very close results. The best known episode is the infamous Bush-Gore recount dispute in Florida, leading to Bush victory by a margin of only 537 votes out of almost six million cast. Similarly, tight results were observed in elections in Germany, Italy and Mexico, for reasons that remained mysterious. At least until September 20, 2006, when the prestigious French newspaper *Le Monde* finally presented a "scientific" explanation, provided by a colleague working at the equally prestigious École Polytechnique. My first surprise when reading the journal was that the colleague did not come from political science or sociology: he was a physicist! And the second surprise was that the article was very annoying; the author indulged in lengthy descriptions, reducing the announced explanation to a few cryptic sentences: tied outcomes result from "contrarian" people, who systematically oppose the majority's choice. However, the author did not give any empirical evidence about the existence of contrarians nor any hint about his mathematical model, which was supposed to reproduce "exactly the phenomena observed in Italy, Mexico, Germany and the United States".

To understand this strange column, it is necessary to delve into its author's technical articles, a feat no average reader of the journal could achieve. At the beginning of the 2000s, Serge Galam had imagined a first so-called "opinion" model, without contrarians. The model analysed the dynamic of choice among two abstract options named A and B, for example, the two candidates of a presidential election. Each

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_4

day, all the voters randomly gather three by three, and systematically adopt the majority opinion within the trio. If two people prefer A and the other B, then the latter is inevitably convinced and becomes favourable to A. Clearly, the hypotheses of the model are not chosen for their realism, but to render them digestible by mathematics. In a way, one buys internal validity, i.e. mathematical rigour, with external validity. Thus, the odd number 3 is chosen to avoid an equality between the supporters of the two sides and the randomness of the grouping to avoid those social or geographical correlations which complicate the life of the modellers… As with many economic models, one chooses to stay under the simple mathematical streetlight, so much the worse if the keys are elsewhere.

In this model, the end state is simple to understand. Since any minority opinion is swept aside, the agents end up all adopting the same opinion, A or B. The final choice simply depends on which opinion was the majority at the outset. It is difficult with such a model to explain the close results. Never mind! Since nothing constrains the imagination of researchers in this field, Galam decided to introduce an additional mechanism that leads straight to the desired result, the "contrarian" behaviour, opposing the majority choice. The new model therefore postulates that, the day after this threesome meeting, a randomly selected fraction of the population systematically adopts the opinion contrary to that of the majority on that day. Note again the realism of the hypothesis: every day you can become randomly contrarian and deny any past electoral loyalty. These two steps (conversion of the minority opinion within each triangle, then contrarian change) are repeated until a state that no longer evolves is reached.

The model (single) interest is to quantify the minimum percentage of contrarians that must be introduced to achieve tight results, a number that is not easy to estimate intuitively. But the simplicity of the model so mutilates reality that it does not explain anything and does not give its author any legitimacy to write about "the future of the democratic system". For Galam was careful not to contaminate his beautiful equations with a laborious empirical justification, such as a field survey or a poll attesting to the increase (or even the existence!) of contrarians. Secondly, random grouping totally neglects geographical, family and social structures, which are known to contribute strongly to the formation of opinion and its persistence. The model will often group two socialist voters with a far right voter, who will then automatically be convinced to vote socialist. The next day, all three may decide to vote for the far right because the majority opinion becomes hostile to it… More generally, Galam never takes the trouble to methodically compare his results to the real world, for example, by studying the evolution of polls during election campaigns or by collecting rigorous statistics on the alleged increase in the proportion of close results. Because in the end, it is not even sure that there is anything to explain. To be scientific, one would have to start by making sure that four close elections are really indicative of a substantive trend, which recent history tends to disprove. For we do not know how to calculate the probability of observing results like 50.1–49.9 or closer. In fact, Hotelling's model (Chapter 2), which suggests a simple mechanism for a race to the centre, may be more relevant in explaining such results.

4.2 Mathematical Smoke Curtains

This cartoon example allows us to draw some more general lessons about simple social models. First, it seems abusive to use common language terms, such as "opinion", to name the simple variables used in the models. Second, this model illustrates very well one of the dangers of social models: the excessive freedom with which, in the absence of reliable data, we model human behaviour, in this case changes in opinion. Every practitioner knows that it is not difficult to tinker with models, by adding a variable here, a hypothesis there, to reproduce anything and everything. But without serious collaboration with specialists in the field, this activity remains a sterile exercise. Physicists are forced to attribute to their atoms behaviours that are compatible with all that is known about them from countless experiments. Social modelling is today in the same position as physics before the long empirical investigations that allowed to tame matter in laboratories and obtain information about atoms (see Chapter 11). Those who try to model social processes are free to arbitrarily choose the behaviour of their little robots to achieve the desired result. Models therefore become tautological, because any behaviour can be explained by arbitrarily adjusting the characteristics of individuals.

This is however the very logics followed by Galam. Since elections led to close results, he arbitrarily assumed that a significant fraction of the electorate was becoming "contrarian", automatically obtaining a model that predicts what he wanted. As it is more difficult to predict the future than the past, Galam announced in his 2006 column that this type of result "should happen more and more often, perhaps even in France" in 2007. In reality, Sarkozy won a clear victory, by more than six points! Like some of his colleagues attracted by social systems, Galam is concerned about scientific rigour as long as he remains on the side of mathematical equations, which he manipulates according to the rules of the art. But as soon as it comes to linking these equations to the real world, rigour becomes an option, and one is content with phrases such as "these results look like…", or "these results suggest that…". In the same way, the reality of contrarian agents is not ascertained by statistics or surveys, but by sentences like "it's increasingly frequent, who hasn't noticed it?". Alas, the strength of a logical chain is only as strong as that of its weakest link.

Is it worth spending time criticizing this bad science? Yes, because in this field, any scientific misconduct can have political consequences. One might object that we should not be too quick to condemn original approaches, which are not yet sufficiently advanced, and that we should give ourselves time, because all original research is born falsified. But then one should not publish columns in national newspapers stating that this science "explains" anything. All the more so as Serge Galam was not at his trial run. Hiding behind curtains of mathematical smoke, he talked about the breakthrough of the Front National (*Le Monde*, 30 May 1997) or about the impossibility of reforms in France (*Le Monde*, 28 March 2000). More discreetly, in "technical" articles, he gave advice to governments on how to fight terrorism, prostitution or drug traffic. His media career suffered from one opinion piece too many, again in *Le Monde*, where he claimed that "when Galileo concluded that the Earth was round, the consensus was against him, agreeing on the platitude of the Earth". In fact, the roundness of the Earth had been known since antiquity, and Magellan had made the first circumnavigation of the globe a century before Galileo. But we're no longer in the realm of equations…

4.3 The Shape of Cities

Trying to explain tight elections is fundamentally an empirical problem, which demands a robust link to reality. This link was completely missing in the previous work, so it is fair to assess the relevance of simple models by studying a more serious attempt. A good example is the elegant model proposed by my colleagues Rémi Louf and Marc Barthélémy, to explain the number of urban centres, thanks to a single "essential" mechanism. Cities are a fascinating object of research because of their complexity and are very important socially, since two-thirds of humans will soon be living there. Understanding their shape and their dynamics are major political issues. In their model, the two physicists try to explain a specific feature: the increase in the number of urban centres when the number of inhabitants increases. Urban centres refer to areas with a high density of jobs, typically the central business district. A small town has only one centre gathering most retail stores and administrations, but large cities have several centres, such as London's Square Mile, Southwark and many others.

My colleagues created a virtual city, containing several potential employment centres, each offering a different wage. Each dweller is only concerned with one thing: choosing the most financially interesting employment centre, the one that represents the best compromise between the salary offered and the cost of transportation from home. Clearly, when the number of inhabitants is low, traffic is low and everyone goes to the employment centre that offers the best salary. A small town therefore has only one active centre. But when the population increases, the traffic increases and so does congestion. As a result, potential centres offering slightly lower wages become attractive, because the lower cost of transport from home compensates for the loss of wages. The more populous the city is, the higher the number of active employment centres, each centre attracting nearby residents, as shown in Fig. 4.1a. Beyond its intuitive nature, the merit of this simple model is that it predicts quantitatively the observed increase in the number of centres with the number of inhabitants (Fig. 4.1b). This success led to its publication in the most prestigious physics journal, *Physical Review Letters*, in 2013.

I have chosen this example among many others because it seems to me representative of physicists' explanations of social phenomena. They use only a handful of explanatory elements (simplicity) and try to link real data and proposed causal mechanisms using mathematics. As explained above, our intuition suggests that congestion leads the number of centres to increase with population. The model is important to check this intuition, as Schelling's model shows that sometimes we

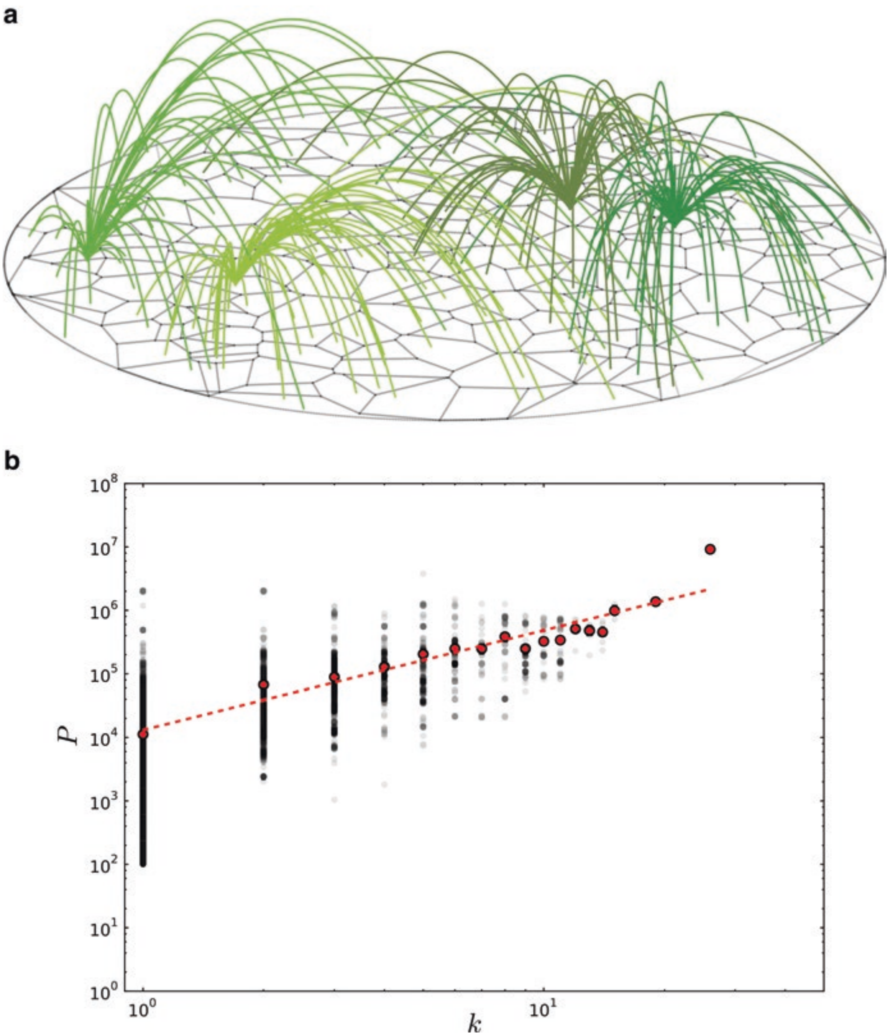


Fig. 4.1 (**a**) City dwellers work in the employment centre close to their homes (convergence points of the different trajectories); (**b**) The number of centres (k) increases as the population (P) increases. Each point represents a real datum (a city at a given date), the bright points the average population for a fixed number of centres and the dashed line the mathematical relationship predicted by the congestion model. (Based on R. Louf and M. Barthélémy, "Modeling the polycentric transition of cities", *Physical Review Letters*, vol. 111, 2013, 198702)

may be wrong. Moreover, a mathematical model can make a *quantitative* prediction, which is more selective than a simple qualitative comparison. Quantitative tests allow in principle to select rigorously an explanation, or discriminate between several alternative explanations. This is how physics has advanced over the last centuries. Couldn't we extend this approach to social systems?

The short answer is no, for reasons that we'll analyse in chapter 6, after understanding why simple mathematical models can do the job in physics. For the moment, let's say that the "thickness", the complexity of social reality, makes the quantitative comparison with real data less conclusive. First, because there exist many other reasons than congestion to explain the correlation between population and number of job centres. For example, as a city grows, rents in attractive areas become prohibitive, and competition for space pushes businesses to move to additional centres. How to be sure that congestion is a better explanation than high rents? In physics, one could quantitatively test both hypotheses and choose the one which better fitted the data. The problem is that, since data are very noisy, quantitative agreement is not very different from a qualitative one (the number of centres "increases" with city population). In Fig. 4.1b, some large cities have the same number of centres as some small ones. Therefore, it is too easy to come up with a model in agreement with the data. Much easier anyway than reproducing carefully measured physical variables, such as the fine-ray structure, which is known with a precision of better than a part in a billion.

Going Further

The opinion model: Christian Borghesi and Serge Galam: "Chaotic, staggered, and polarized dynamics in opinion forming: The contrarian effect", *Phys Rev E,* vol. 73, 2006, 066118. See also http://arxiv.org/abs/cond-mat/0404265.

For the city model, see Rémi Louf, Marc Barthelemy, (2013) Modeling the Polycentric Transition of Cities. Physical Review Letters 111 (19)

Chapter 5
A Physical Simple Model

To get a more complete picture and draw some conclusions about the interest of simple models for understanding reality, let's take a last example, this time from standard physics. I used to work in "nanoscience", i.e. the field that tries to understand the behaviour of tiny objects, containing a few hundred or thousand atoms and measuring a few nanometres (1 nm = a billionth of a metre). This field is technologically important, as original electronic and biological devices with these dimensions are being produced and challenge our knowledge.

At that time, we deposited nanosized clusters on a surface to observe them with microscopes. We were intrigued by a strange picture taken by a PhD student (Fig. 5.1, left). The "stars" (Fig. 5.1, left) represent the aggregation of many of these clusters on the surface. As the rain of clusters reached the surface at random locations, their grouping meant that clusters were able to move on the surface. We were intrigued because until then, the experts in the field knew that single atoms could move rapidly on a surface because of the unresting agitation of the surface, but considered that nano-clusters were too big to move. Upon landing on the surface, they would anchor strongly to it and remain immobile: after all, who has ever seen a drop of water moving on a table? Clearly, the very existence of these stars proved that clusters could move. But images alone could not tell us how fast they would do so. To quantify their speed, we needed a quantitative model that included the three essential ingredients at play: the "rain" of clusters, their motion on the surface and their sticking upon contact.

By coding these ingredients in a computer, I created a virtual deposition chamber, to observe step by step the progressive formation of the stars. As clusters land on the surface, they start to move around. When two of these wandering clusters bump into each other, they stick together and stop moving, creating the seed for a star that soon traps other stray clusters. The essential point is that I could change the clusters' diffusion speed at will and observe its influence on the shape of the stars. Intuitively, faster clusters lead to a smaller number of (larger) stars, as their higher speed increases their range of exploration of the surface and therefore the

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_5

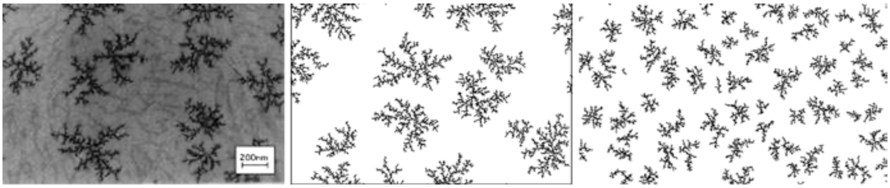


Fig. 5.1 Experiment (left) and numerical simulations from a mathematical model (centre, right). The resemblance of the two first images demonstrates a fact that was previously considered impossible: clusters containing thousands of atoms can move very quickly on surfaces. The two simulated images correspond to a high (center) and a low cluster diffusion speed (right)

probability of finding an already formed star before finding another wandering cluster and create a new star. This qualitative picture is confirmed by the simulations (Fig. 5.1 centre, right). But the advantage of using a mathematical model is that the relation between the speed of the clusters and the number of stars can be quantified. By tuning the speed in the virtual deposition chamber, in order to recover the number of stars found experimentally (Fig. 5.1 left), we could precisely deduce the speed of the real clusters. And what we found puzzled the whole community. Contrary to what was previously thought, nano-clusters are able to move very quickly on certain surfaces. We actually found a speed a billion times higher than people expected, which allowed us to publish our results in the best physics journal!

Let's pause and look back to analyse how we managed to understand something about a real physical system using a simple model. Why were we successful, while the examples taken from real social systems were failures? The essential point is that physicists are allowed to *transform* the real world, so that reality resembles a mathematical model. This could even represent a definition of physics. Those domains that can be connected to a rigorous mathematical model become the realm of physics. In our example, we started by depositing clusters on glass surfaces but failed to explain what happened, as glass surfaces seem to be full of "traps", which capture the clusters and prevent their motion. But these traps are poorly known, so we cannot integrate them into models. Then we managed to prepare very pure and flat surfaces, where clusters could move freely and the description became simpler. This took a long time, as we had to learn, by trial and error, that the surfaces had to be heated at 500 degrees for a few hours (to remove chemical impurities) and then peeled off using tape, to obtain a sufficiently smooth surface. Moreover, we had to carry out the deposition experiments in "ultrahigh vacuum", i.e. a purified environment without ambient air or anything that would make reality more complex than the simple model.

The extensive control over reality allows to deeply entwine model and reality, by multiplying the comparisons. For example, the stars do not all have the same size, and the model reproduces very well their size distribution and predicts that whatever the cluster speed (which changes the average star size, as explained above), the rescaled size distribution (i.e. after scaling by the average size) remains identical, which was borne out by experimental measurements. It is also possible to test what happens when two clusters meet and stick together. Do they really remain immobile, as assumed in the original model, or do they continue to move (maybe more slowly)? We introduced this mobility in an alternative model, and observed that it would change the size distribution of stars, deviating from the one observed experimentally. We could then rule out this hypothesis. By comparing with other alternative models, we could also rule out the possibility that a glued pair of clusters could detach after some time.

The model makes it possible to ask "what if" questions and compare its answers to those observed in real experiments. What if the rain of clusters becomes stronger? What if the size of the clusters changes? What if the temperature of the surface increases? What if the surface is tilted? By checking that, for many different modifications, both the model and the experiments react in the same way, we can be confident that what goes inside the computer is pretty similar to what is going on in the deposition chamber. And trust the model to deduce processes we cannot observe directly, as the speed of the clusters on the surface.

Going Further

L Bardotti et coll, "Experimental observation of fast diffusion of antimony clusters on graphite surfaces", *Phys. Rev. Lett.* **74** 4694 (1995)

Chapter 6
What Are Simple Models Worth For?

We can now draw some general lessons about the usefulness of simple models. Their most interesting use is to help us improve our thinking, by revealing the intertwined effects of several causes. Sometimes, it is easy to combine ingredients and guess the result, as for the introduction of contrarians in Galam's opinion model. In that case, using a mathematical model seems more useful to impress non-scientists than to understand anything. But as soon as there are complex causalities at play, our intuition may fail us. Think of Schelling's model, where interactions between agents that move to improve their satisfaction lead to a situation where… no one is satisfied! Moreover, creating a mathematical model forces us to transform our more or less vague ideas into clear, well-defined rules. And their simplicity makes them understandable, which is not the case for the more complex, black box models that we will explore later (chapters 8 and 13). So, simple models do help to think. And they can also be useful to understand real systems, but only those that can be simplified, as in physics.

My experience as a practicing researcher has taught me that social systems are too complex to be captured in this way. We can understand why this is so by comparing the cluster and the cities models. First, and most important, physical reality can be simplified in labs. In our experiments, we managed to create a simple world, with only a handful of ingredients, those taken into account in the model. Instead, cities develop by a complex network of countless and uncontrollable causes. Even to explore the effects of a single cause such as congestion, models have to take into account a large number of ingredients that are necessary for congestion even to exist: dwellers, transportation, jobs locations… And all these ingredients lead to many questions that have to be settled: Where do people live? How do they reach their offices? How do they choose their jobs? The quantitative predictions of social models then depend on the combination of many ingredients. Even more annoying, many causes, not taken into account in the model, may be actively contributing to the result we wish to explain. For example, new job centres may also emerge when rents in existing ones become prohibitive, or when stores decide to locate far from

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_6

each other in order to avoid competition on prices (chapter 2). Second, these few ingredients of physical models are quite simple and lend themselves to mathematical formalization: clusters reach the surface at random locations, the surfaces are flat… Instead, "simple" often means simplistic in social models. For example, in order to find a mathematical solution, one has to assume that all city dwellers and all companies are identical, that dwellers choose their residence at random… Assuming randomness has often turned out to be fruitful in physics, but is less relevant for social systems, which generally show complex structures. Dwellers try to live close to their workplace, cities are spatially segregated by income, voters do not meet randomly but following family and social links… Using random distributions then represent a way of hiding away real complexities for (bad) modelling reasons. Finally, social models and reality cannot be as deeply entwined as in physics. Remember how we gained trust in the adequacy of the model by multiplying the comparisons and checking that for many different modifications, both the model and the real system reacted in the same way. For social models, the comparison with reality is more superficial, as we cannot change society at will. Maybe, if people chose their residential location at random, this would change the number of centres, but how could we know?

To summarize, simple physical models can help in understanding reality, because they address simple situations, i.e. objects that have been simplified in labs. There are only a few ingredients, and these are under control because they are created by researchers and benefit from a long history of research: clusters are known to reach the surface randomly; the way they move on surfaces has been studied for decades by many groups, for many materials… In other words, simple physical models can be trusted because they tame a marginal increase in complexity, a controlled jump into the unknown from a well-known situation, a trustable bedrock. Instead, simple social models rest on quicksand: there are hidden ingredients, most of which are not well-known or need to be unreasonably simplified, assumptions are difficult to test, etc. Physicists have been brought up in the causal clarity of the laboratories, but this might well be a bad education for dealing with social systems.

Chapter 7
Complexity: Step by Step

The lesson is clear: simple models cannot account for social reality, because it cannot be easily simplified in controlled environments. Then, the strategy to follow seems obvious: let's make more realistic models, by taking into account additional mechanisms. I will now explore examples of this strategy, by climbing up progressively the complexity scale, and see how far it can lead. As a first step, let's model systems composed of relatively simple entities, as animals or humans in framed situations, where their actions are constrained by the environment and social rules (think of pedestrians walking in normal situations). This may allow us to master the complexities emerging from their interactions.

7.1 Ants' Nests

Some insects have an astonishing ability at collectively building very complex nests, exceeding their individual cognitive abilities. Some termite mounds measure 7 m, which is more than 600 times the size of each ant. These nests often have very elaborate shapes, with superimposed chambers connected by spiral staircases, with shortcuts and poor connectivity between the different areas, which make them easier to protect against external attacks. How can we explain the sophisticated collective coordination required by an organization at this scale? In one of the many wild analogies between science and society, it has long been assumed that the queen played a key role, centralizing all the information in the colony and then directing the activities of the workers as needed. Today, we know that no ant, even the queen, has explicit information about the whole. Zoologist Pierre-Paul Grassé introduced in 1959 the fundamental concept of *stigmergy*, which allows to understand the principle of the construction of such structures by insects with only local information, following a few simple rules. Stigmergy comes from the Greek *stigma* (sting) and *ergon* (work), and literally means "stimulating work", in the

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_7

sense that the coordination of tasks and the regulation of constructions do not depend directly on the workers, but on the constructions themselves. The worker does not direct his work, but he is guided by it. Individuals modify their environment and leave chemical traces (pheromones), which become sources of stimulation for others. The individual modifies his environment, but the environment, through a feedback loop, influences the other individuals, leading to a subtle collective organization.

While the concept of stigmergy allows the coordination to be conceived, its practical relevance had yet to be demonstrated. In a nice article published in 2016, a team led by biologist Guy Théraulaz managed to show, in all details, how ants of the species *Lasius niger* build their nests. To do this, they first studied experimentally how ants choose where to deposit the pellets they carry. By patient measurements, they showed that ants preferentially deposit a pellet next to other already deposited pellets, provided that these pillars have been made recently by other ants. In practice, the pillars are marked by a pheromone which evaporates quickly, and whose presence is necessary to stimulate other ants. An additional experiment, in which the researchers introduced "false" pillars without pheromone, showed that the ants did not continue their construction.

Stigmergy controls the amplification of pellet deposits in certain areas and thus the emergence of pillars that serve as a foundation for the nest. The second necessary mechanism is to stop the elevation of the pillars to start building capitals and then a first ceiling. Experiments suggest that ants use their bodies as a template to decide when to stop stacking pellets. Once the pillars reach their own height, ants begin to build a roof. Using these two simple and empirically validated mechanisms, the researchers created a model with 500 virtual ants enclosed in a box measuring about 50 cm on each side. At each iteration, supposed to last 1 s, ants move randomly, choosing whether or not to pick up a pellet of earth, carry it or put it down, following the rules measured empirically. They also mark the deposited pellets with a pheromone, which gradually evaporates, at a speed that is the only unknown parameter of the model. To fix it, the researchers built several virtual nests for different values of evaporation times, the best agreement corresponding to a time of the order of 15 min, which is similar to that measured for other species of ants. By bringing these mechanisms together, the researchers were able to build a virtual nest and compare it to real nests, using X-ray tomography, which allows the three-dimensional structure of the nest to be visualized without destroying it (Fig. 7.1). The agreement between model and reality validates the concept of stigmergy, and allows to dispense with the unrealistic hypothesis of the queen as a centralized planner.

Fig. 7.1 Three-dimensional image of an ant nest (top) and the first stages of its construction (bottom), observed in controlled experiments. (© CRCA, CNRS)

7.2 Framed Human Action: Car Traffic and Pedestrians

Models of similar complexity can be applied to simple human behaviours, such as pedestrian movements. A possible application of these models is to improve the fluidity of traffic in public spaces. To design such spaces, we generally use the know-how developed by previous constructions, which made it possible to find, through trial and error, suitable solutions. But how can new ideas be explored, especially if the objective is building a train station or a sports stadium, which should allow tens of thousands of people to enter and leave as quickly as possible? It would obviously be too expensive to experiment by building real facilities. But we can try to build virtual ones and study their performance through simulations, to explore innovative solutions while reducing the risks.

For this, we first need virtual pedestrians, who behave like real ones. Never mind! Many companies propose "ready to walk" virtual agents. Thus, the Massive company proposes a standard pedestrian, which "comes equipped with a full locomotion brain and a motion library for extras that run, walk, stand and sit". This agent can be rented for \$500 a month. But it is also possible to rent more specific – and more expensive – agents, such as the Mayhem agent, which can create public disorder of all kinds, including riots, political protests, mass panic and disaster scenarios.

How do you teach these little robots (in fact, computer routines) to behave like pedestrians? There are a host of different models developed by groups around the world. I will take here the example discussed by Mehdi Moussaïd in his thesis defended in Toulouse, in Guy Théraulaz's group. As for ants, the objective of these models is to provide these virtual pedestrians with simple rules, allowing quick calculations and reproducing their real behaviour. Initially, Mehdi had chosen one of the most popular models, inspired by the physics of granular materials. Pedestrians are like active pebbles reacting to "social forces" of three types: first, a force that moves them towards their destinations, at a speed of about one metre per second; then, a pedestrian interaction force, which essentially represents avoidance; and finally, a repulsive force exerted by walls. To find the precise expressions of these forces, Mehdi Moussaïd carried out several experiments. In one of these, about 40 people walked in a narrow corridor equipped with cameras. Three situations were studied: the person walks back and forth from one end to the other of the empty corridor; the same thing with a person standing in the middle of the corridor; finally, two people leaving at the same time from both ends of the corridor and avoiding each other. By accurately measuring movements in these controlled environments, Mehdi could compute the three forces. I present, for once, the real formula used by researchers to account for the forces of interaction between two walkers denoted by "*i*" and "*j*":

$$\vec{f}_{ij}\left(d,\theta\right) = -Ae^{-d/B}\left[e^{-(n'B\theta)2}\vec{t} + e^{-(nB\theta)2}\vec{n}\right]$$

θ represents the angle between the direction of travel of pedestrian "*i*" and the direction pointing from "*i*" to walker "*j*" and d the distance between the walkers. The other parameters are numbers adjusted to reproduce the experiments: $A = 4.5$, $n = 2$, $n' = 3$, and the value of B depends on the relative speed between the pedestrians. According to this model, we are supposed to measure the different characteristics of the situation (distance to the other pedestrian, angle θ, respective speed) about 100 times per second and then compute the force exerted on us using the above formula and change our speed accordingly. And of course, to be useful in more realistic environments, the model assumes that we repeat this calculation for all the people and obstacles in the vicinity, which gives a rather sophisticated calculation when we walk down a shopping street on a Saturday afternoon…

Mehdi suggested that we rather treat pedestrians as humans, with more realistic psychological and cognitive characteristics. We know, for example, that people walk by seeking a clear path through obstacles, rather than by computing complex equations. He proposed an algorithm based on heuristics, a term that refers to the quick decisions we make without thinking too much about them. He showed that two simple rules allow to calculate, at each instant, the direction and speed of the pedestrian and thus her trajectory. First, the person chooses the direction that is as close as possible to his destination, which is not obstructed by pedestrians or obstacles. He goes there at his comfort speed, unless an obstacle is too close, in which case he adopts a lower speed. These two simple rules, which seem compatible with our intuition, reproduce very well the measurements made on pedestrians in the corridor, but also the formation of lines of pedestrians walking in the same direction in the narrow corridors of the metro, or the time needed to evacuate a concert hall. Complemented by a term of friction when people come into contact in a very dense crowd, these rules seem to account for a large number of behaviours of interacting pedestrians.

7.3 What Are These Models For?

In these two models with simple entities, the elementary mechanisms are empirically validated and not arbitrarily postulated, as in the simplistic models of Chapter 4. The overall result (the nest, the movement of the crowd) is also characterized in detail and reproduced using the individual mechanisms. If understanding means being able to reconstruct a system by thinking, or by simulations, these models do the job. But besides reproducing known processes, are we really learning something? What is the point of simulating the transition from the individual to the collective, when we know all the ingredients and the result in advance?

For pedestrians, modelling is above all useful for applications. Thanks to the mechanisms validated by the model, reliable predictions can be made about what would happen in different spatial configurations, when building a new train station or football stadium. However, there is no guarantee that these mechanisms, calibrated in specific experimental situations, can be extrapolated to other contexts. In case of panic, pedestrians may behave differently! For ants, the consistency of the results makes it possible to empirically validate the idea of stigmergy and to specify its modalities. There is no need to assume that there is a “nest spirit”, or some other entity that would hold the plan for the whole. Stigmergy is therefore a kind of economic coding of the processes required to build such a complex structure. It regulates the behaviour of ants through the competition between the amplification of positive information via pheromones and negative feedback via its evaporation. As this evaporation time depends on atmospheric conditions (temperature, humidity), the distance between pillars will be modified by the latter and eventually lead to architectures adapted to the climate.

Going Further

The idea of stigmergy was introduced by Pierre-Paul Grassé in "La reconstruction du nid et les coordinations interindividuelles chez *Bellicositermes natalensis* et *Cubitermes sp.*", *Insectes sociaux,* vol. 6, 1959, pp. 41–80. The technical article reproducing nest construction is Anaïs Khuong *et al,* "Stigmergic construction and topochemical information shape ant nest architecture", *PNAS,* vol. 113, 2016, 1303–1308, which includes some nice videos.

An example of a pedestrian library: http://massivesoftware.com/massiveagents.html

The equation for pedestrian interactions can be found in Mehdi Moussaïd, Dirk Helbing, Simon Garnier, Anders Johansson, Maud Combe, Guy Theraulaz, (2009) Experimental study of the behavioural mechanisms underlying self-organization in human crowds *Proceedings of the Royal Society B: Biological Sciences* 276 (1668):2755–2762

Chapter 8
Complex Models to Understand Complex Social Situations

Now we know that the complexity scale can be climbed up to efficient models of pedestrian traffic. Can we proceed further, to more complex (and important!) situations, to address social issues such as unemployment, wage inequalities and economic growth? Again, the path forwards seems obvious: let's take into account all the important mechanisms to create realistic models. After all, computers are powerful enough to handle millions of entities, and, as our life has been digitalized, we have lots of data about human societies. Can we simply succeed by brute force? The following examples should convince you that it is not that simple! In short, making a model more realistic by adding poorly controlled mechanisms often leads to poor understanding and no predictive power. More is different, not better.

8.1 Predicting Economic Growth

Can economic growth be predicted? As many prestigious institutions (International Monetary Fund, World Bank, etc.) deliver forecasts every month, the answer seems obvious. Moreover, growth forecasts are not pure mathematical exercises without consequences, as they play an essential role in government budgeting. Growth determines tax revenue, so the reliability of these predictions seems crucial for economic life. Yet, an official White House document, released in 2016, leaves serious doubts about their accuracy (Fig. 8.1). It reproduces International Monetary Fund forecasts for global growth between 2010 and 2020. Let's take the highest curve, which corresponds to the growth predicted in September 2011 for the period 2012–2016. The forecast deviates from reality already in the first year: actual growth is just under 3.5%, instead of the 4% announced. After 5 years, the gap is considerable: the expected growth is 5%, almost twice as high as the 3% actually observed in 2016. One could imagine that, seeing their error as early as 2012, forecasters would then make less optimistic projections. This is not the case, since exactly the

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_8

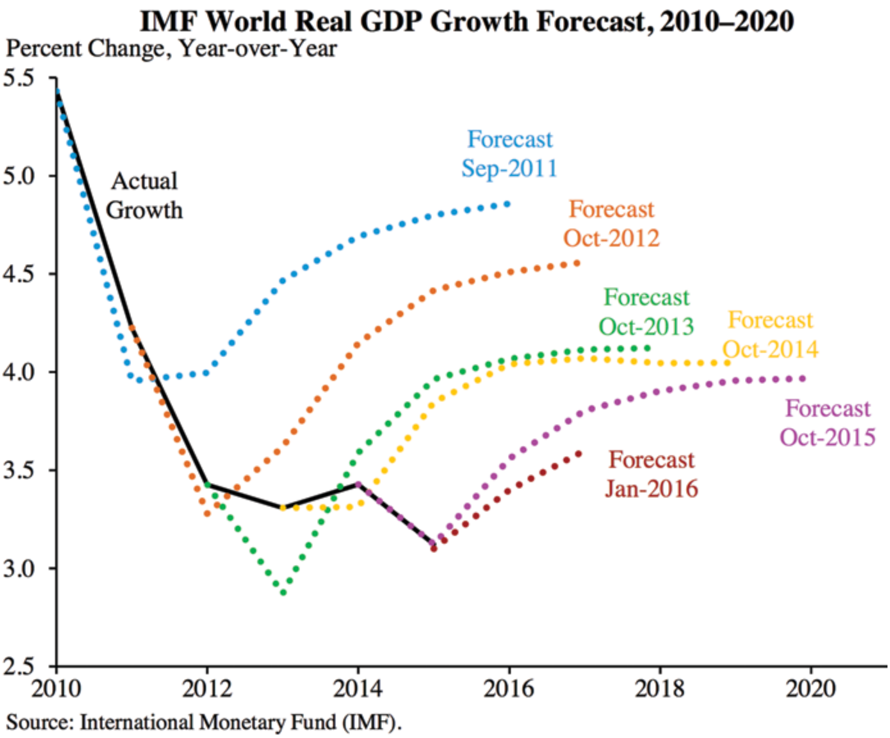


Fig. 8.1 International Monetary Fund forecast for world growth between 2010 and 2020. The black line corresponds to actual growth and the dotted lines show the growth observed in the previous year and the 5-year forecasts made at different dates. Each year, the IMF reproduces the same error, forecasting a strong rebound that is never observed

same upward jump is predicted every year. IMF economists do not seem to be able (or willing?) to learn from their past mistakes.

IMF prediction errors are not an unfortunate exception. In 2011, two economists systematically compared predictions and actual values for French growth since 1998. The results are edifying (Fig. 8.2). Even when it comes to short-term (1 year) predictions, sophisticated models do little better than the following simplistic “prediction”: next year’s growth will be equal to… this year’s! Another symptom of a serious problem is that the various assessments produced by public (International Monetary Fund, French government, etc.) or private (BNP, Total, etc.) bodies are generally optimistic, predicting too high a growth. If these deviations stemmed from inevitable inaccuracies in such complex areas, one would expect the forecasts to be sometimes too optimistic, sometimes too pessimistic, with an average error close to zero. The systematic positive bias underscores the importance of adjustable parameters, which allow researchers to modify the results to suit unscientific objectives, for example, the desire to induce a certain economic policy or to minimize future deficits. Forecasts therefore tell you a lot about those who make them, but not much

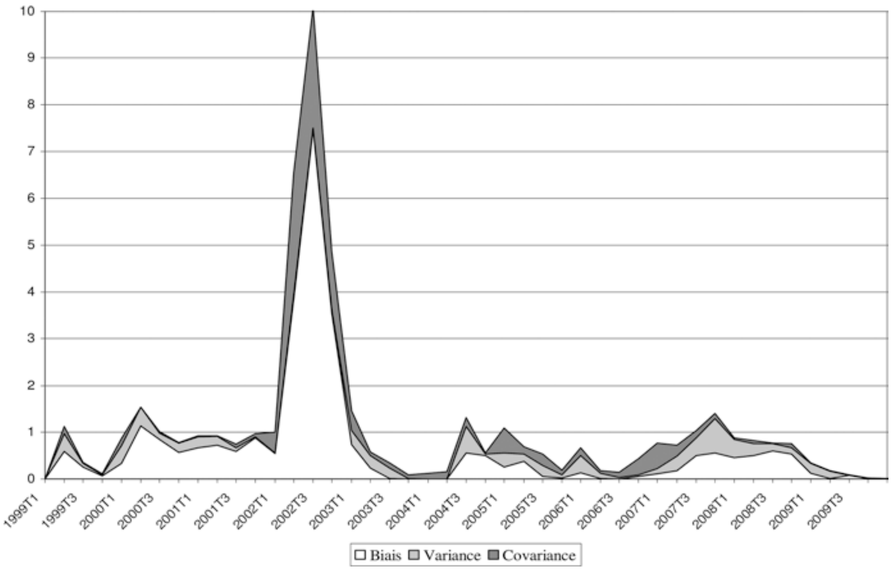


Fig. 8.2 Ratio of prediction errors between sophisticated models and the naive prediction that growth in the coming year will be the same as in the current year. A ratio greater than one means that the sophisticated models do less well than the naive prediction and vice versa. Weather predictions consistently remain below 1. (Taken from Jobert, Th and Persyn, L "Quelques constats sur les prévisions conjoncturelles de la croissance française", Revue d'économie politique, vol. 122, pp. 833–849 (2012))

about the future. If you add to these biases the well-known irrelevance of GDP in quantifying a country's true wealth, public debates on growth forecasting become a bit pathetic!

To better understand these repeated failures, it is necessary to study the two main types of approaches used for economic forecasting. The first consists in extrapolating past trends, using mathematical algorithms, and the second in building a virtual economy from more or less realistic ingredients.

8.2 The Future Is Like the Past

The first method attempts to predict the evolution of GDP based on its past values, to which is added a colourful inventory of data, chosen according to their predictive power. Predictions combine many indexes that were found to correlate with growth, such as investor confidence, the exchange rate of the euro against the dollar, the price of oil or the change in opinion about expected activity in the building trade. These variables need to be skilfully mixed, as the impact on growth of each may be more or less long term. For example, the building indicator does not seem to have an impact on growth until 9 months later. This type of forecast does not work too badly in the short term (a few quarters), but it does not do much better than simply extending the current trend, as we have seen.

Moreover, it seems conceptually uninteresting, for two reasons: first, because there is no guarantee that the future will be identical to the past. Economic crises have amply demonstrated this fact. And it is precisely when major changes occur, that we need forecasts the most – and that this type of prediction proves to be the least reliable! Moreover, the method cannot be improved by simply adding more economic data. Indeed, there is no guarantee that past correlations last forever. To take a simple example, every year between 1967 and 1997, growth in the United States was perfectly correlated with the geographical origin of the winner of the *Super Bowl*. A standard statistical test would have given a one in five million chance that this was a coincidence, which it was anyway. It is highly unlikely that a person would win the lottery, but since millions of people bet, a few still win. More seriously, there was a strong link between growth and employment up to 1999, but this link has been considerably reduced today, making forecasts of these two variables more uncertain.

What is at stake here is the "transferability" of economic correlations. This point has been well-known to economists since Robert Lucas' classic article in 1976, which distinguished two types of relationships: fundamental ones, which are supposed to be invariable in all contexts, and effective ones, which are observed only in certain circumstances. His objective was to criticize the macroeconomic models produced by Keynesians, which used observed regularities between aggregate quantities (government deficit, growth, etc.) to predict the effects of a certain policy, for example, a decrease in the central bank's credit rate. For Lucas, these predictions were not very rigorous, since the regularities observed for some rates have no structural reason to remain valid for other rates. And of course he was right. We must therefore look for "autonomous" relationships, those that depend as little as possible on the context. For example, if we study the relationship between the speed of a car and the distance of the accelerator to the floor, we will find a very strong correlation: the closer the accelerator is to the floor, the faster the car goes. But this relationship is not autonomous because it ceases to be valid in the event of a breakdown, or when driving uphill or downhill. On the other hand, the relations between pressure and temperature of the gases inside the engine, taken from the laws of thermodynamics, are autonomous relations, because they describe the functioning of the engine, independently of what is going on elsewhere. Standard economical theory assumes that the only stable relationships are the preferences of the actors. Relationships between macroeconomic variables result from the combination of myriads of individual decisions and can be greatly altered by changing circumstances.

8.3 Modelling the Job Market

The second method of predicting the state of the economy accepts these premises and is therefore carried out at the individual level. Take, for example, a typical model for the labour market, which simulates the behaviour of several thousands of companies and individuals. Each of these actors has specific attributes, determined once and for all, such as gender or preference for free time over work, or productivity for companies. They are also characterized by internal variables that change throughout the simulation, such as age, wage level or profit rate. Thanks to the availability of demographic and socio-economic data, these variables can be initialized in a relatively realistic way. In order to describe in detail the hiring mechanisms (fixed-term or permanent contract?), the choice to temporarily leave the working life or to look for another job, models use dozens of somewhat arbitrary parameters. For example, to calculate the evolution of "human capital" (worker's professional experience), its adaptation to the needs of each job and its productivity, the WorkSim model uses 11 somewhat obscure parameters. Overall, its 50 parameters are adjusted by reproducing about 50 empirical characteristics of the labour market, such as the unemployment and activity rates by age group, the proportions of fixed-term contracts compared to permanent contracts, income distributions, etc. Once these parameters have been set, the scale model of the labour market is perfectly defined, and the effects of various modifications can be observed. Researchers may find, for example, that making layoffs easier would lead to a 6% drop in youth unemployment after 4 years, to the detriment of older people, with unemployment remaining stable overall. The big winners would be companies, whose profits increase by 20% over the same period, thanks to the reduction in overstaffing, as economic layoffs become easier. Sometimes, effects stem from a combination of opposite trends, and the outcome is difficult to predict accurately, as it depends on the precise values of the parameters. For example, by abolishing fixed-term contracts, some people who feel unable to find a permanent contract are encouraged to leave the labour market, thus reducing unemployment figures. On the other hand, however, some job offers, previously made on fixed-term contracts, disappear, which tends to increase unemployment. The net result of these two opposite effects is very difficult to ascertain with some confidence.

These realistic models need a large number of parameters, which are almost impossible to adjust in a robust way. Two quite different sets of parameters can lead to an equivalent replication of the empirical data. But there is no guarantee that the two virtual job markets built in this way react similarly to changes. This approach suffers from two other major flaws. First, the individual preferences of the actors are not as stable as neoclassical theory claims (see Chapter 11). Second, behavioural economics has shown that the way actors make choices (leaving the labour force, taking a particular job) cannot be easily summarized by simple formulas. As a result, the predictions of individual-based models are generally less reliable than those extrapolating past data!

Going Further

The White House Report on forecasts for global growth: https://www.govinfo.gov/content/pkg/ERP-2016/pdf/ERP-2016-chapter3.pdf

On systematic errors in economic forecasts: Jobert, Th and Persyn, L "Quelques constats sur les prévisions conjoncturelles de la croissance française", Revue d'économie politique, vol. 122, pp. 833–849 (2012)

Robert Lucas' paper on the fragility of macroeconomic forecasts: "Econometric policy evaluation: A critique", *Carnegie-Rochester Conference Series on Public Policy,* vol. 1, 1976, p. 19–46.

An example of job market model: http://worksim.lip6.fr/publis/AE2015_GoudetKantBallot.pdf.

Chapter 9
Modelling Epidemics

We tried to climb too high the complexity ladder. In this chapter and the following, we'll understand how far models can go and why. We'll first explore three different ways of forecasting an epidemic outbreak, to show how complexity can sometimes be tamed. The next chapter explains why meteorological models have succeeded in forecasting the complex evolutions of the atmosphere.

9.1 Individual Models

Let's start with the propagation of an epidemic. Can we create a virtual society, introduce a virus and watch it propagate in different ways as we explore different mitigation measures? What if we increase mask-wearing? What if we lockdown only elderly houses? To study the effects of these measures, the first idea that comes to mind after reading the preceding chapters is to develop models at the individual level, following how persons become sick and infect others. A study published in July in *Nature Medicine* followed this strategy, by using half a million agents, representing the demographic characteristics and household structure of the whole French population. The activity of each agent is described in excruciating detail: meeting friends and family for 180 min at a 1-m distance once a week; working in either a small company (leading to contacts with two colleagues) or a large one (ten colleagues); visits to the closest grocery store (1.2 times per week, meeting five people)… Finally, a disease model translates these social contacts into infection probabilities. Overall, the model included 194 parameters. One hundred forty were needed to create a synthetic French population from demographic data, 33 for determining social contacts and 21 for virus characteristics. Two of them (contamination risk and proportion of undiagnosed cases) were calibrated by fitting the number of cases, the number of deaths and hospital beds occupied. What are the conclusions reached through such a huge effort? The authors summarize them as "Both physical

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_9

distancing and mask-wearing […] would be ineffective in ultimately preventing […] a second lockdown. However, these measures coupled with the shielding of vulnerable people would be associated with better outcomes, including lower mortality […]. Benefits would nonetheless be markedly reduced if most people do not adhere to these measures".

Qualitatively, most of these conclusions are obvious: "shielding of vulnerable people" of course "leads to better outcomes", and if people "do not adhere to physical distancing and mask-wearing, then the benefits would be reduced". No need of such a complex model to predict these effects. The interest of a mathematical model consists in predicting the right numbers. This leads to a second remark: can we trust the quantitative predictions? The authors made a substantial study of the robustness of their predictions, testing the effects of varying the values of model parameters. They concluded that this "had little impact on outcome estimates", but a close inspection of their results shows that, for example, the predicted peak of infections (with physical distancing, mask-wearing but no shielding of individuals at risk) is not very precise. The model predicts July 1, but there is 50% probability that it happens before June 1 or after July 20. And as I'm writing this (September 11, 2020), the peak has not been reached yet…

A preceding individual-based model had focused on Ebola epidemics in Liberia, to estimate the force of spread through three main channels: within families, in the hospital and at funerals. Scientists from Boston created a virtual country, with families distributed in the different cities, calibrated according to the real demography. By running several virtual epidemics for many values of the three forces and comparing the results with real data, the percentages of the three main transmission routes can be estimated. But the result is rather imprecise: for example, the proportion of hospital-transmitted infections is estimated at 38%, but the possible values range from 10% to 76%! As in the COVID case, the huge uncertainty results from the combination of two main sources. First is the lack of reliable data, both for the number of cases recorded – which often corresponds to a fraction of the actual cases – and for household characteristics. The other source of uncertainty is the enormous amount of different individual histories compatible with the available data (the total number of infected people), much as a value for the pressure of a gas can result from a multitude of different atomic trajectories. For Liberia, some stories compatible with the data contain only 10% hospital transmissions – when, for example, infection begins in overcrowded homes – while others include 76% hospital infections, when many families decide to hospitalize their patients very quickly. Put another way, the data are too aggregate, the tip of the complex iceberg of underlying processes at work at the individual level.

9.2 Brute-Force Big Data

This suggests to improve predictions by using the avalanche of social data. Such an idea is not new. Between 2008 and 2013, Google was able to reliably report on the prevalence of influenza in many countries, based on searches made by Internet users. The idea is simple: the more people make queries that appear to be related to flu ("fever", "muscle aches", etc.), the more likely they are to be sick. Another example of Google's frightening power? Not quite… In February 2013, *Nature* magazine announced that the predictions had become incorrect, and GoogleFlu was discontinued… What had happened?

Essentially two things: naive confidence in the power of big data and lack of structurally sound modelling. Google's engineers probably had too much confidence in the power of their datasets. As a result, they made elementary errors in the data modelling, which could have been avoided by specialists. But above all, the methodology used was fragile. It basically consisted of finding matches between the 50 million search terms used by Internet users and the thousand or so available data on flu cases. In this situation, it is inevitable to find fortuitous matches, which exist in past data, but are structurally unrelated to influenza, and therefore do not help predict the future. For example, queries about the results of school basketball competitions are highly correlated with the flu epidemic, simply because both take place in winter! In 2010, specialists had already noted that Google's accuracy was no better than that obtained by conventional methods using medical data from the previous 2 weeks.

9.3 Controlling Complexity

From this failure, one might be tempted to draw the easy lesson: *big data* is overrated, even dangerous. This is all the more tempting since this data is often kept confidential, because Google's logic is: "I give you a service and in exchange I appropriate your data". But it seems more reasonable, even inevitable in the long term, to try to take advantage of these data, which is valuable for public health. How to do this in a more thoughtful way? The secret lies in taming complexity by finding regularities and stable relations that allow to build a robust framework that can then be informed with relevant data.

Contrary to economics, epidemiology can take advantage from the stability of contagion processes, which are grounded in biology. Compared to economic choices (investing, buying, saving, etc.) which involve subtle cognitive processes and would depend on a lot of parameters, being susceptible, becoming infected and recovering are essentially driven by biological processes beyond our control. A great part of the dynamics of an epidemic outbreak can be captured in a robust way by the simple equations of basic models such as SIR ("Susceptible, Infected, Recovered"). They describe how people that are susceptible (S) to the virus become infected (I) and then either recover (R) or die. When these equations are combined with socially imposed regularities (work, school, transportation, etc.), complexity can be controlled, leading to reasonably robust predictions. The key point consists in properly calibrating the complexity of the model, to take advantage of the available regularities, but resisting the temptation of introducing too many details (and parameters), which would then be difficult to adjust reliably. This is a matter of expertise in the field, of playing with data, with models and learning from past experience. For example, only two figures are needed to obtain a rough estimate of the number of deaths that would result from the herd immunity strategy. First, the percentage of the population needed to reach collective immunity. For COVID-19, a reasonable optimistic estimate gives 50%, in line with the result of a “natural experiment”, the outbreak on the Charles de Gaulle aircraft carrier, where 70% of the young adult sailors became infected before the epidemic came to a halt. The second figure is the mortality rate of those infected, which has to be adapted to the population relevant characteristics (age, obesity, etc.). For COVID-19, it is estimated at 0.3–1.3% for countries such as France and the United States. This would translate into 100,000–450,000 and 500,000–2,100,000 deaths, respectively. These may seem crude estimates, but they are certainly sufficient to forbid that strategy. What government could survive such a deliberate choice made in front of its citizens?

Going back to forecasts, once the basic equations are right, one could even try to feed in some Google-like data. A team at Columbia University combined their expertise in epidemiology with the very accurate real-time data provided by Google, winning the $75,000 prize offered by the Atlanta Center for Disease Control for the best prediction of the flu epidemic in 2014. The team first created several hundred virtual epidemics, basing their dynamics on robust epidemiological equations. Each epidemic differs only in the value of the parameters that govern the dynamics. Researchers therefore have a reservoir of possible trajectories for the epidemic, which they then feed with real data. In addition to Google “flu” -type queries city by city, authors included confirmed cases of flu and data on humidity, which is known to foster the epidemic. These data favour certain trajectories, those that are closest to the real epidemiological data. An important point is that by observing the degree of convergence of the different forecasts, one is able to estimate, in real time, the reliability of the forecasts. By combining the expertise of epidemiologists with these new data, it is possible to predict epidemic peaks 2 months in advance in a fairly robust way. This allows health managers to be more proactive in their prevention and mitigation actions and to plan in advance, for each city, prevention, vaccination or drug distribution. Upon receiving his award, Jeffrey Shaman painted an optimistic picture for the future: “Weather forecasting has been around for almost 60 years, and it wasn’t very good, but it has been steadily improving. If we make the choice – at the level of the scientific community and the country as a whole – to invest in the prediction of infectious diseases, we will make progress too”. Understanding the way weather forecasts improved, and to what extent they represent a sensible analogy, is the scope of the next chapter.

Going Further

The individual model for Covid-19 in France: Hoertel, N., Blachier, M., Blanco, C. et al. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. *Nat Med* **26**, 1417–1421 (2020) and for Liberia: Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: A computational modelling analysis, *Lancet Infect. Dis.*, **15**, 2015.

Arguments for formal modeling in epidemiology: "Mathematical models: A key tool for outbreak response", PNAS, **111** 2014, 18095. Unfortunately, epidemic forecasts made during an outbreak are rarely investigated during or after the event for their accuracy. A recent exception is Funk, S. et al. PLoS Comput. Biol. 15, e1006785 (2019) who noted that forecasts made in a 2014–15 Ebola outbreak in Sierra Leone reliably predicted the epidemic's course one or two weeks ahead of time, but no longer.

A critique of the "black box" flu models developed by Google: http://gking.harvard.edu/files/gking/files/0314policyforumff.pdf

On the competition for flu prediction in the United States: http://www.cdc.gov/flu/news/predict-flu-challenge-winner.htm

Chapter 10
Why Weather/Climate Forecasts Can Be Trusted

Epidemic models provide an example of the importance of robust relations, based on biology, to tame complexity. This chapter analyzes the success of weather and climate forecasts, thanks to a virtual scale model of Earth's atmosphere. Why are the simulations about such a complex system reliable, when forecasting the economy fails? In a nutshell, atmospheric simulations take advantage of robust physical conservation laws (energy, momentum) to create a backbone which, when fed with relevant data, can lead to robust predictions.

10.1 Predicting the Weather

Scientific weather forecasts were first developed for the military. On November 14, 1854, a violent storm devastated the Anglo-French-Turkish fleet on the Black Sea. Urbain Le Verrier, director of the Paris Observatory, gathered scattered data from nearby observatories and showed Napoleon III that the storm could have been anticipated. He took the opportunity to present his project for a permanent meteorological network. By the end of the nineteenth century, several countries had set up a vast system of production and accumulation of meteorological observation data, taking advantage of the brand new telegraph network. A new stage was reached with First World War air battles. It became important to know the state of the atmosphere not only at ground level but also at altitude, leading to new developments in meteorological science. This warlike context is still present in the terms used in this discipline, for example in the expression "collision" between cold and hot "fronts"!

In the beginning of the twentieth century, some meteorologists proposed to forecast the weather based on the laws of physics. Indeed, hydrodynamics and thermodynamics were well-established sciences that described the behaviour of the atmosphere under the effect of heat and pressure. What are the relevant ingredients for weather forecasting? Firstly, the light and heat input by the sun, absorbed in the

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_10

atmosphere and the oceans. The amount of heat received is greatest at the equator, creating a heat flow towards the poles, via the oceans and the atmosphere. This flow is influenced by the Earth's rotation, relief or ocean currents, making it difficult to model. The first daring attempt was made in 1922 by an English mathematician, Lewis Fry Richardson. He divided Europe into 18 rectangles, 200 km long, each serving as a base for five slabs stacked 12 km high. For each of these 90 blocks, he tried to calculate the seven fundamental quantities of the atmosphere: temperature, pressure, humidity, air density and the three dimensions of air speed (in the east, north and vertical directions). To define the initial state, he used the values obtained during a rare event, a simultaneous balloon release throughout Europe in 1910, which provided an unusual harvest of meteorological data. His model predicted a sharp rise in pressure in the following days, which did not occur at all!

For many years, the idea of calculating the evolution of the atmosphere by the equations of physics proved ineffective. Progress was based on the collection of ever more observations, gathered in maps. To make forecasts, the map of the day was drawn and the archives were searched until a previous map was found that was as similar as possible. By looking at the maps of the following days, one had a good idea of the weather to come. This method gave the Allies a significant advantage for the landing in 1944. Their meteorological service managed to predict that June 6 would be the only day of good weather in the middle of bad weather days. This lull allowed to surprise the Germans, who had not been able to anticipate it. It was only gradually that the physical approach to weather forecasting prevailed over more empirical methods. It took the enormous power of computers, coupled with the establishment of a global network of observation stations. As late as 1995, humans were still able to do better than numerical predictions, especially during extreme events, the early stages of which they were able to detect through their experience. But today, that's over. Thanks to computers capable of performing millions of billions of operations per second, numerical methods do better, solving hundreds of equations on the millions of cubes a few kilometres wide that carve up the globe.

However, difficulties remain, firstly, because some important processes occur at scales that are too small in relation to the atmospheric partitioning required for the calculation. The formation of clouds by condensation of water vapour into droplets is a complex event that depends of course on the percentage of humidity in the air, but also on local wind values, the difference in temperature between the ground and the altitude or the concentration of impurities, such as aerosols. To model cloud formation in detail, these processes would have to be described on scales ranging from millimetres (the size of droplets) to a hundred metres (the size of clouds). This is impossible for models that divide the Earth into blocks of a few kilometres. Therefore, approximate formulas are used, which give the percentage of cloud cover as a function of the average values of these variables over the entire cube. These formulas can be calculated by physical models that seek to describe the underlying processes. Alternatively, machine learning algorithms can learn them from real data linking the cloud cover to the percentage of water vapour. But is it really important to know, for each cube, the precise distribution of water between clouds and vapour? Unfortunately, the answer is yes, because atmospheric water plays a very different role in the circulation of heat depending on whether it forms liquid droplets (in clouds) or remains dispersed as a vapour. In clouds, it strongly reflects sunlight, reducing the amount of solar heat that reaches the Earth's surface. But these clouds also form a blanket that prevents the infrared radiation emitted by the Earth from escaping into space, which tends to increase the temperature, especially at night. The role of clouds in global warming is therefore ambiguous. In addition, clouds promote updrafts that carry the air from the Earth's surface to very high altitudes, carrying energy and moisture, which has a profound effect on meteorological phenomena. As a result, despite its importance, the overall effect of clouds is not yet well-known, leading to forecast uncertainties. These inaccuracies are reinforced by patchy observations over the oceans. And the trouble is that the slightest difference between the actual state of the atmosphere and the state entered into the computer can dramatically change the forecasts, due to the chaotic nature of the atmospheric equations, known as the "butterfly effect". Mathematical analyses of the equations show that it is impossible to go beyond the 2-week predictability horizon, as this would require a perfect knowledge of the initial state of the atmosphere, as well as of the inevitable disturbances (the famous flight of the butterfly). As a result, forecasts are always probabilistic: in 4 days, there is a 60% chance that it will rain between 9 and 11 am. These percentages are obtained by running several simulations in parallel, each starting from a slightly different state of the atmosphere, to account for uncertainties. The percentages are obtained by counting simulations that result in rain and those that give the opposite result. A striking example of the difficulty of forecasting is given by the error of several degrees in the 6-day temperature forecast for Europe, made on February 15, 2014. To understand the cause of this error, meteorologists rewinded their weather simulators. They realized that the air over Europe on February 21 was, when the forecast was made (6 days earlier), off the coast of the United States, in the middle of the Pacific Ocean. In this isolated area, upper-level wind observations are scarce and inaccurate, leading to a poor initial state in the predictions. The initial error was then amplified as this air zone passed over the United States and then the Atlantic Ocean. By slightly modifying, in later simulations, the strength of the winds in this area of the Pacific, the observations of February 21 were well reproduced, confirming the need to improve weather observations over the oceans to arrive at reliable 1-week predictions.

10.2 A Tamed Virtual Earth

Meteorologists have managed to develop a virtual Earth locked up in computers, set up to reproduce as much as possible the real atmosphere, but more manipulable, more docile, to allow us to speed up time and make predictions, or to go back in time to better understand the gaps between predictions and reality. This virtual Earth must be regularly "fed" by observations to behave properly. And it has an enormous appetite, since it needs values for each of its cells, each of the cubes that make up the model, and which make the predictions all the more accurate as the

Earth is finely cut. Ideally, observations should be available for each of these cells. But in the 2000s, global models included five million cubes, when only 50,000 observations were available. To solve this problem, the first idea was to feed cubes with no data by computing average values from the nearest stations. Gradually, it was realized that it was smarter to do the opposite. Run several simulations with different starting conditions and, for each, compute the expected values of the observations at each station. By comparing expected and real values, one can privilege the simulations with the smallest deviations. The key point is that, in this way, the internal consistency of the model (mass or energy conservation) is ensured, which was not the case with the previous method.

These scientific and technological developments have greatly improved predictions, by about 1 day per decade. To properly assess the quality of forecasts, they must be compared to two "easy" predictions:

– Tomorrow, weather will be the same as today (weather persistence).
– Tomorrow, weather will be the average weather for this city on this day of the year (climatology).

Today, weather forecast surpasses these two simple models, for up to a week. At 2 days, the temperature error is less than two degrees, compared to three for average weather and four for persistence. Ability to surpass these simple models may seem trivial, but there are not many other examples throughout the book! Models are now so accurate in the short term that they may even be useful to correct real observations. For example, a large discrepancy between the wind direction predicted by the model and the one observed was noted at a station in South Africa. When analysing the reasons for this unusual discrepancy, it was found that the probe had been incorrectly installed, using the geographic instead of the magnetic pole as reference.

10.3 Predicting the Climate!

Confidence in numerical models of the atmosphere led to an even bolder challenge: extending the simulations over time to predict the climate of the coming decades. At first glance, this seems contradictory with the chaotic 15-day barrier discussed above. However, climate represents the *average* state of the atmosphere over long periods (30 years by convention). And, even if we are not able to predict the temperatures of each day, we know that next summer will be warmer than this winter. To predict climate, it doesn't matter that the models don't respect the real succession of states of the atmosphere: it's enough that they know how to reproduce its regularities, its average behaviour.

Modelling the climate, however, requires the introduction of additional ingredients, because on a scale of decades, we have to take into account processes that are too slow to affect the evolution of the atmosphere over the few days of the weather predictions. For example, the CO_2 level can be taken as constant for weather forecasts, but over decades, it changes as a result of human emissions and carbon exchanges between vegetation, oceans and the atmosphere. It is therefore necessary to integrate the carbon cycle into climate modelling, which implies modelling the evolution of vegetation as well. The dozens of parameters needed to describe these processes are based on ancient climatological data. Since the mid-nineteenth century, data collected through the network set up by astronomical observatories can be used. For older times, meteorological data can be derived from dependent phenomena, such as tree growth or the concentration of gas in Arctic ice. This involves the obscure – but essential – task of standardizing these scattered data to make them comparable. Indeed, each measuring instrument is a priori in its own world, and these different contexts must be coordinated by transforming the data. For example, at the beginning of the nineteenth century, meteorologists measured rain on roofs and temperature at windows. To avoid the effects induced by dwellings, they then moved the devices to ground level and open spaces. As a result, the measured temperatures decreased and rainfall systematically increased. Previous measurements must therefore be corrected in order to obtain homogeneous data over the entire period. In the same way, the Pula station, located today in Croatia, was managed successively by the Austrians, by the Italians between 1918 and 1930, then by Yugoslavians until the occupation by the Germans in 1941… Each nation having its own practices or units of measurement, a meticulous verification work is necessary before integrating the data into the world bases.

Despite all these pitfalls, climate predictions are becoming increasingly accurate and reliable. And the politically important point is that they all predict strong global warming. This was not obvious, because one could have imagined, for example, that the increase in evaporation caused by an initial warming would cover the Earth with clouds, which would reflect sunlight, producing a cooling effect that could counteract the warming. It is only through accurate simulations that we can better understand the effects of some factor when so many variables come into play. Despite the diversity of approaches attempted by the climate science community, variations in poorly known parameters or the integration of different types of data, no model is able to reproduce current climate changes, such as the rise in temperatures between 1980 and 2010, without including the effect of carbon dioxide emitted by human activities. There is no other way to reproduce the present warming. We can therefore have some confidence in the models to guide our economic and political decisions. But reducing our carbon emissions is also justified on the grounds of reducing our environmental footprint and consumption or social justice… Important political issues always have dimensions other than scientific ones.

10.4 Why Did Climatologists Succeed?

This story began by an improbable bet: relying on the equations of physics to forecast weather. It should be remembered that at the beginning of the nineteenth century, there were only a few scattered observations of the state of the atmosphere, little computing capacity and no understanding of the degree of predictability of the weather. Today, modelling routinely produces reliable predictions over several days, sometimes up to 2 weeks. Why have meteorologists succeeded, where economists have failed? First, the fundamental relationships for atmospheric prediction are well-known and remain valid at all important scales, both in time and space. The basic equations for the evolution of air masses, which describe their dynamics or heat exchanges, are well established. More generally, the equations that express the conservation of certain physical quantities (air mass, quantity of motion, energy, etc.) represent a "skeleton" that structures all the calculations and makes them robust. In addition to these equations, there are terms that reflect the forces experienced by air masses – such as gravity or the force due to the Earth's rotation – which are also very well-known. We have seen that the models also include less understood processes, such as cloud formation. These phenomena create uncertainties in the predictions, but without calling into question the validity of the previous fundamental equations. This is because these equations have proved their worth in other fields, totally independent of meteorology, such as river flow or aircraft aerodynamics. Second, meteorologists do not try to predict the precise evolution of the atmosphere, which would be impossible because of the chaotic nature of its evolution. They forecast only its average state, and that this is sufficient to support political decisions. The third ingredient of success is more institutional, the painful work of creating global bodies capable of gathering data and homogenizing them. Without this work, yet to be done in epidemiology, numerical models could not have been calibrated effectively.

Going Further

A great review of meteorologists' modeling practices: Paul Edwards, *A Vast Machine*, MIT Press, 2010.

The example of the six-day prediction error is taken from a recent synthesis: "The quiet revolution of numerical weather prediction", *Nature*, September 3, 2015.

I drew the comparison between climate predictions and "easy" predictions from Nate Silver's book, *The Signal and the Noise*, Penguin, 2012, which contains many other examples of predictions.

Answers to frequently asked questions in climatology by the IPCC: https://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-faqs-fr.pdf

Chapter 11
We Are Not Social Atoms

The lesson to be learned from the previous chapters is that most important social issues cannot be tackled by the “virtual society” approach, as social scale models lack the robust foundations that underlie physics simulations. As explained in the preface, such an approach was enticing for physicists, enthralled by Newton’s success in understanding the solar system. Creating a virtual scale model of society would have ensured understanding society “workings” and allowing to predict its future. This ambitious idea had feet of clay, as it rested on a superficial analogy between physics and sociology: the social atom.

In 2007, Mark Buchanan, former assistant editor of the prestigious magazine *Nature*, proclaimed that physics can now predict whether a city’s neighbourhoods will segregate, stock markets collapse or a new wave of crime hit your city. The title of his book, *The Social Atom*, nicely sums up the project of calculating society as physicists are supposed to do with matter, by relying on atoms. This image of a society composed of a priori independent individuals with fixed characteristics serves as a theoretical basis for contemporary “social physics”. It dates back at least to the dawn of liberalism and philosopher JS Mill famous statement: “The laws of the phenomena of society are, and can be, nothing but the laws of the actions and passions of human beings united together in the social state. Men, however, in a state of society, are still men; their actions and passions are obedient to the laws of individual human nature”. Almost two centuries later, an influential review of the field concurs: “In social phenomena, the basic constituents are not particles but humans [… social physics] attempts to understand regularities at large scale as collective effects of the interaction among single individuals, considered as relatively simple entities”. At the heart of the atomic vision lies a simple idea: each of us can be characterized by a set of “internal” characteristics, which remain stable in different situations. Our actions could then be explained from the combination of these characteristics with some relevant features of the context. For example, Schelling agents (Chapter 2) are provided with a utility function, fixed from the outset, which determines their actions given the relevant context, namely, the proportion of same

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_11

color neighbours. More realistic simulations, such as those of the job market (Chapter 8), create a myriad virtual agents interacting like robots endowed with some predefined preferences.

To understand the problem, let's go back to the (pre)history of physical atoms. Many "explanations" put forwards by atomists make us smile today. Thus, Lucretia interprets the flavours of different bodies by the shape of their respective atoms: "And note, besides, that liquor of honey or milk; Yields in the mouth agreeable taste to tongue; Whilst nauseous wormwood, pungent centaury; With their foul flavour set the lips awry; Thus simple 'tis to see that whatsoever; Can touch the senses pleasingly are made; Of smooth and rounded elements, whilst those; Which seem the bitter and the sharp, are held; Entwined by elements more crook'd, and so; Are wont to tear their ways into our senses; And rend our body as they enter in". Nearly 20 centuries later, Descartes' disciples were to argue that lemons are acidic because their atoms have prickles (Figure 11.1)… As I explain in Annex 1, atomism would become scientifically fertile only centuries later, when intense experimental work would show the stability of some chemical species and allow to estimate atoms' characteristics, limiting the freedom to postulate them ad hoc. The following paragraph explains why nothing remotely similar is likely to be achieved for the so-called social atoms.

The social atom approach has three main problems: how to measure the "internal" characteristics, how to be sure about their constancy and, most importantly, how to ensure their stability in different situations and different contexts. To illustrate these challenges, I'll use a simple example: the Intelligent Driver Model (IDM), which describes how a driver reacts (by accelerating or breaking) to the surrounding traffic, to maintain its desired speed and a safe distance to the closest (leading) vehicle. The model assumes a relatively simple formula to predict how the driver changes her speed, using three main driver characteristics. First, the "desired speed", the maximum speed the driver will reach under unobstructed driving conditions. A natural value for this parameter would be the typical (highest) speed on the considered road element. However, some (timid) drivers may prefer to drive a bit below this limit, while an aggressive driver may exceed

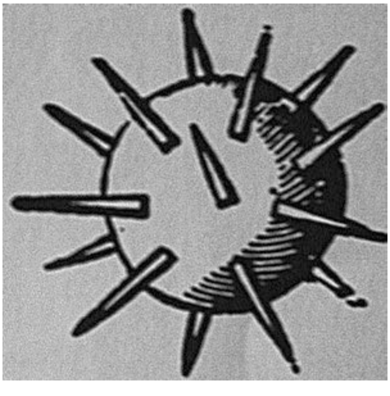


Fig. 11.1 Descartes' disciples argued that lemons are acidic because their atoms have prickles

it. Second, each driver wants to avoid bumping into the car in front. This is quantified by a “time gap”, defined as the time available for a driver to brake and avoid collision. A typical value in dense traffic is about 1.4 s. This time gap can be transformed into a gap distance, by multiplying the gap time by the vehicle speed. The third parameter describes the “desired acceleration”, i.e. the way a driver accepts to change its speed. An aggressive driver prefers to accelerate and decelerate fast, while a timid driver will perform gentler changes, represented by a lower value of the parameter.

Let’s go back to the three problems raised by the social atom approach. First, a key point is that social atoms’ features are never directly observable. They are an *interpretation* of the measured data using one model among many possible models. And, crucially, the predictions depend on the framework chosen. In the driver model, one studies how the person changes speed according to his environment to infer the values of the three parameters that best account for the observed behaviour. However, the IDM is but one among hundreds of driving models, each of which has different parameters with different interpretations. Some account better for usual braking/acceleration decisions, others for emergency braking, others for the observed dependency of the time gap on the traffic density… There is no universally accepted model. The point is that interpreting the same behaviour using different models leads to different values of their parameters and, more importantly, to different predictions. By the way, inferring the characteristics of physical atoms raises similar problems. For example, there are many models for attributing characteristics to water molecules. Some are able to reproduce well the freezing behaviour of water, others its evaporation, others the maximal density observed at 4 °C… The problem gets much worse when social atom characteristics have to be inferred from the observation of aggregate, statistical data. This is the case for large-scale models such as WorkSim (chapter 8). One infers the values of agents’ features (preference for leisure, productivity, etc.) indirectly from socio-economic data of the present situation, such as unemployment rate. It is well-known that inferring individual characteristics from collective data is tricky, as Schelling model shows: observing segregated cities cannot tell us anything about preferences, since individual preferences for both mixed and segregated neighbourhoods may lead to segregated cities.

Second, once values for social atoms features have been computed, a further problem arises, their temporal stability. For otherwise, the whole approach becomes useless, as it cannot be used to predict future behaviour. Empirically, the driving parameters fluctuate, for the same person, from day to day, depending on whether she is more or less concentrated, hurried… More profoundly, our parameters can be permanently altered by life events. For example, a serious accident may transform our driving behaviour from normal to timid, which changes the supposedly stable values of the parameters. Actually, our social experience teaches us that people do change in time, as they learn or meet people that change their mind. This is a key difference with physical atoms.

But the main weakness of the social atom perspective is the third one, the intractable “frame” problem. When we humans take a decision, how do we determine the relevant variables, those inside the frame and those that can be ignored? This selection process is crucial for modelling, as those variables not taken into account literally disappear from reality. For example, the traffic model assumes that the relevant variables are the actual speed, the desired speed and the time gap. From these, the model can compute how the driver adapts its speed. However, further experiments showed that traffic density is also relevant, as most drivers increase their preferred time gap after being stuck in congested traffic for some time. Of course, it is possible to include this into a more sophisticated model, adding more variables and more parameters. However, research on the "frame problem" has shown that this is a hopeless task. The reason is a bit technical, but the key insight is simple: we know more than we can explain. An important part of human decision processes consists in defining the appropriate context, which gives meaning to the situation and helps finding the appropriate actions. In a restaurant, we expect to pay for food, but not at a friend's house. The relevant characteristics of a situation are filled in automatically and unconsciously by the brain, thanks to our social experience (see also the Annex 2 for a more realistic vision of human characteristics, beyond the social atom). However, for modelling, we need to specify all the relevant characteristics, which is impossible as we do not know them explicitly.

Let's look at some examples of the difficulty. What variables determine our choice of a variety of wine? Certainly food tastes in general, income… and music! Indeed, shoppers in wine stores disproportionately buy German wine when German music is playing in the background. More seriously, Jean Piaget wanted to find out at what age children acquire a clear notion of numbers. He presented them with two configurations of tokens: either four widely spaced or five more closely spaced, so that the total length of the four was greater than that of the five. He then asked them: which row has more tokens? Children under 7 chose the longest line, the one with only four objects. Piaget concluded that they hadn't yet acquired the notion of numbers. Later, another researcher came up with the idea of repeating the experiment using candies instead of tokens and asking another question: which line do you want to take? And now, kids systematically chose the line with five candies! The type of objects used therefore changes the meaning that the children give to the question and their action. Here is a final example taken from economics. A waiter proposes for dinner steak or salmon. The customer thinks a little bit and chooses the salmon. The waiter then remembers that today's special is frog's legs and adds this choice. The customer thinks again and finally decides… to have the steak! It is impossible to understand this change within the framework of fixed preferences. If the customer preferred steak to salmon (and frog's legs), as his final choice reveals, he should have chosen it from the very start. To understand his decision, we have to assume that his choice depends on a wider context, including the estimated quality of the restaurant. At first, he thinks that, given the few options available, the restaurant is not worth much, and that it is better to choose salmon, as it is hard to miss, rather than a steak that may be overcooked. When he learns that the chef also offers a more sophisticated dish, his opinion about the restaurant rises and he chooses the steak.

To sum up, individual characteristics are generally difficult to find, depend on the framework used to interpret the data, and they are seldom constant. Despite this, their values sometimes represent a convenient summary of the data for predicting further observations, at least when individual characteristics are directly observed, as in the traffic model, and in the short term, where one expects only small drifts. However, the frame problem often leads to meaningless predictions, as we can never be sure in advance what the relevant variables are, unless we strictly control the environment, as in physics' laboratories. In real traffic, people change lane and approach traffic lights, and the model does not account for these situations. How could we expect social atoms with fixed characteristics to accurately predict the complex and creative behaviours of economic actors?

From a more political point of view, the social atom vision is only relevant when individuals are conceived as simple entities whose collective action one hopes to compute from an external point of view. By construction, one therefore assumes by that agents are incapable to understand and control the events at the global level. As Elinor Ostrom notes, most simulations assume that only "human beings external to those involved – scholars and public officials – are able to analyze the situation, ascertain why counterproductive outcomes are reached, and posit what changes in the rules-in-use will enable participants to improve outcomes […] It is assumed that the momentum for change must come from outside the situation rather than from the self-reflection and creativity of those within a situation to restructure their own patterns of interaction". Instead, the "complex" vision of individuals proposed in the Annex 2 acknowledges that human collective action is generally too complex to be simply "computed" and respects people's efforts to build and impose their own categories for defining the meaning of their collective actions.

Going Further

On the use of atoms to understand matter, see my previous book: Des atomes dans mon café crème, Seuil, "Points Sciences", 2001.

On the frame problem, see HL Dreyfus, Why Heideggerian AI failed and how fixing it would require making it more Heideggerian, Artificial Intelligence 171 (2007) 1137–1160

JS Mill's quotation is taken from "*A System of Logic, Book 6*: The Logic of the Moral Sciences" (1843)

Traffic model: Kesting, A., Treiber, M., & Helbing, D. (2008). Agents for traffic simulation. Multi-agent systems: Simulation and applications, 11, 325, available at https://arxiv.org/pdf/0805.0300.pdf

The quotation from Ellinor Ostrom can be found in page 648 of her article: "Beyond Markets and States: Polycentric Governance of Complex Economic Systems" *American Economic Review*, vol 100 (2010)

Chapter 12
Social Data Are Soaked by Social Complexity

The failure of the "virtual society" approach does not mean that numbers are irrelevant for social matters. I will present in the next chapters two alternative approaches that use mathematical tools to extract information from real data, instead of starting from an imagined bunch of mechanisms. If appropriately obtained, these data are intimately "soaked" by social reality and therefore contain relevant information about it. Of course, in addition to numerical data, social sciences have developed many qualitative approaches to understand complex social situations. Narratives, in-depth qualitative descriptions are often the best way to account for the complexity of specific situations. Mathematical approaches become important when the scope is to generalize findings across different cases. In chapter 14, I'll present statistics, which uses relatively simple mathematical tools to disentangle the contributions of different possible causes, to create a shared vision about the problems and to help controlling them. But we'll first study a more recent approach, called machine "learning", that uses complex mathematical tools which are able to "learn" when fed with enough data. It is essentially an engineering approach, whose priority is to predict human behaviour in areas of commercial interest rather than to increase human understanding. Both approaches critically depend on the relevance of the analysed data, so it is important to first discuss what social data can tell us about social reality.

12.1 Careful with Data!

Data literally means "something given", as it originates from the Latin verb "dare", to give. However, data are rarely "given". They are generally "obtained" through complex processes that *transform* the studied situation and therefore can never be entirely trusted to faithfully represent it. Consider the data used to calibrate COVID-19 epidemiological models. Even basic quantities such as the number of

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_12

infected people or deaths are collected and categorized in different ways by different countries. This is partly due to political reasons, as governments want to claim superiority of their health strategies by showing low number of deaths. But there are technical reasons too: When someone already close to death gets infected, should her death be attributed to COVID or not? Should one include those who succumb to disruptions to regular health care, which can delay treatments? These decisions have an important impact on the numbers. For example, in April 2020, the recorded deaths from diabetes were 20–45% higher than the 5-year average in the United States. You may think that counting the excess mortality, by comparing year 2020 to the average number of deaths of preceding years, could help avoiding such difficult decisions. The trouble is that the lockdown decreased many deaths, especially from road accidents and crime. In South Africa, the number of non-natural deaths had dropped to half their usual number during the lockdown! Finally, in many countries, the number of deaths is not even recorded. A programme specialist in Ottawa notes that "across the world, about 50% of the deaths occurring in a given year are registered ... The other 50% do not exist at all. They are invisible." The situation is even worse for the number of infected people, as most of the "asymptomatic" (roughly half of the infected) have not been tested at all. This is but one example of the important lesson drawn from French statistician Alain Desrosières. To be able to quantify anything, to transform a phenomenon into a number, one needs first to "agree" and then to "measure". It is necessary first to agree on the way to measure and to categorize fuzzy and complex events into clear cut boxes. For example, how to differentiate between people who died of COVID-19 and those who were infected but died from a serious illness they had before? This requires a death-classification system that accounts for the underlying conditions that make COVID-19 more likely to kill. Once specialists agree on how to distribute cases in different categories, one still needs a reliable and costly infrastructure to "measure", to apply those conventions and centralize the "data".

12.2 Data on Social Regularities

What do we learn from social data? Sometimes, not much, as we already know a lot about our social sphere. For example, in 2010, a group of physicists stated in the prestigious journal *Science* that human behaviour is "93% potentially predictable". This figure was impressive: thanks to physicists capable of exploiting the information provided by *big data*, predictability of human behaviour seemed within reach of computers… The research team had access, under somewhat strange (and secret) conditions, to an anonymized database, from an undisclosed country, containing the locations of the telephone calls of ten million people over a 3-month period. These data enabled them to reconstruct, hour by hour, the trajectory followed by each person in the 60 or so places they visited during this period. However, once we put aside the exaggerations often necessary to publish in prestigious journals, it turns out that this study only revealed social features we all know about. For this figure of "93% predictability" is not that impressive after all. Firstly, because the locations are only known according to the closest relay tower, which defines areas of about 3 square kilometres, corresponding to the size of a district of a large city. Predicting that you are somewhere in such a wide area is not so difficult. Second, contrary to what the authors state, our "common sense" does not tell us that our movements are "erratic and unpredictable". On the contrary, we know that a large proportion of our journeys are driven by commuting to and from work, and that we tend to sleep at home, giving us 90% predictability at night. It is easy to understand that by knowing, thanks to 3 months of data, our most frequent locations, such as home, work and a few additional privileged places at other times of the day (shopping, meals, school, etc.), we can quickly obtain a very reliable forecast of our location in the coming hour. But this strong predictability is largely trivial: in an hour, we have a good chance of being in the same place! A Danish group has shown that if we try to predict our next (different) location, predictability drops to 68%. As often, "found data" merely allow to quantify a well-known social regularity, bringing in no important new knowledge. Annex 3 discusses how controlled experiments allow to gather small but *relevant* data.

12.3 Rich Data, Needs Only (Conceptually) Simple Modelling

However, it is fair to acknowledge that massive amounts of data, soaked by social information, can be useful for commercial purposes. Let's look at Google's main business, predicting the most appropriate ads for Internet users. In practice, as soon as you open a page, advertising algorithms compute on the fly the probabilities that you click on a large number of different ads and display the ones deemed most favourable. The estimation is done using all the characteristics algorithms can retrieve about the user and his recent browsing. In other words, you "buy" Gmail by allowing Google to track you. Even a modest increase in advertising effectiveness represents billions of euros for advertisers and is therefore actively sought after. A colleague working at the University of Pennsylvania told me an anecdote that helps understanding the real source of the (poor) predictability of our behaviour. His team had developed a sophisticated algorithm to predict whether an Internet user would click on an ad. To test their algorithm, the team used a database containing 236 million ads presented during search queries such as "bicycle seller in Lyon". For each query, the file contains a dozen characteristics: the age and sex of the Internet user, the position of the ad on the screen, the words used in the query, the keywords of the ad, an ad identifier (there are 100,000 distinct ones) and obviously the result: was the ad clicked or not? Most of these characteristics can take a large number of values, just like the words in the query. At the end of the day, each situation is described with 170,000 distinct features! The game consists in predicting whether the user will click on the ad or not according to all of them. How does one proceed to estimate this probability? As in most of these applications, the researchers used a machine learning technique (see next chapter), to estimate the 170,000 coefficients on a few million queries and trying to predict the rest. The results of these sophisticated algorithms are incomprehensible and can only correctly predict a very small percentage of the ads clicked.

The fact is that when my colleague submitted the article to a computer science journal, the publisher, who worked for a large web company, told him that his algorithm was certainly very sophisticated mathematically, and very interesting, but that web giants try to avoid such algorithms, because they are too complex and unreliable. In fact, instead of trying to deduce our actions from our intrinsic characteristics, they seek predictability in the large amount of data they accumulate, thanks to the famous *cookies*. When the database is large enough, there is always another Internet user whose behaviour is similar enough to ours, to take him as a model of our future behaviour with reasonable efficiency. Thus, when I surf on Amazon, the site looks for other people who browsed the same books, to predict that other books that have interested them in the past may also be of interest to me. The information is hidden in the small differences between people with very similar backgrounds. No need for a sophisticated algorithm, no need to understand why I'm interested in this book. The mass of data makes it possible to operationalize commonplace sayings such as "birds of a feather flock together". Nothing conceptually exciting, but very efficient when enough data are available.

Going Further

Examples about unreliable data on Covid-19 have been taken from "The true toll of the pandemic", *Nature* (3 September 2020, p 22)

On human predictability: Chaoming Song et al, "Limits of predictability in human mobility", Science, vol. 327 p 1018 (2010).

On advertising prediction, Aryan Mokhtari and Alejandro Ribeiro, "Global Convergence of online limited memory BFGS", Journal of Machine Learning Research, vol. 16, 2014.

A must read on the effectiveness of digital advertising: The new dot com bubble is here: it's called online advertising, thecorrespondent.com, Nov 6th, 2019

Chapter 13
Machines that Learn How to Model

– What do you think, Maigret?
– I never think
Someone had retorted one day "He soaks it all in".
It was true in a way.
Words were too precise for him,
which was why he preferred to keep quiet.

Maigret and the tramp, chapter 5

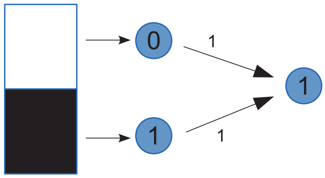
"Machine learning" has become the fashionable approach to derive information from social data. I'll first focus on "neural networks", a type of algorithms inspired from the human brain, to give an idea of how these complex mathematical tools work. Neural nets are only one example of machine learning algorithms, which all follow the same logic: use lots of data to fix the value of the large number of internal parameters. I'll then study several recent examples to show that machine learning approaches can be trusted only in "framed" situations, when social variability has been somewhat reduced.

13.1 Neural Networks

The originality of machine learning has been nicely captured by Pedro Domingos. For him, standard programming, "like all engineering, is a lot of work: we have to build everything from scratch. Learning is more like farming, which lets nature do most of the work. Farmers combine seeds with nutrients to grow crops. Learners combine knowledge with data to grow programs." Let's look at a simple example, the programming of a black square detector by a neural network. Figure 13.1 shows a simple setup, illustrating the activation of an output neuron (right) when at least one of the squares in the left image is black (logical function "OR"). It is assumed that the neurons on the left each observe one square and become active when that

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_13

Fig. 13.1 Simple example of an artificial brain containing only three neurons



square is black. They then send a signal of value 1 to the output neuron. The output neuron weights each signal by the influence of each neuron (the value of the corresponding link), adds them, and becomes active when the signal exceeds 1. If both influences are set to 1, only one of the two left-hand neurons needs to be active for the output neuron to become active, as it will receive a signal equal to 1, which is obtained by multiplying the value of the active neuron (1) by that of the link (1). And the method is flexible, since by modifying the influences, one can achieve the logical function "AND." Check that with links' values 2/3 and 2/3, the output neuron becomes active only when the two left-hand neurons light up, i.e., when the two squares in the image are black.

Machine learning proposes a different approach. Instead of requiring the user to enter link values, the algorithm learns them from examples. Let's go back to the black square detector. How to teach this device the logical "OR" function? Concretely, one needs to find the values of the influences that reproduce this function and that activate the output neuron when at least one of the two squares is black. Let's assume that both influence values are zero at the beginning. The learning process is the following: the machine is given an image, e.g., black square at the top and white square at the bottom, which activates the first neuron (value 1) and not the second (value 0). The machine then injects 1*0 + 0*0 = 0 in the output neuron, which remains inactive. Now we tell the machine that it was wrong, because a black square should be enough to activate the output neuron. What will the artificial brain do with this information? The trick is to "learn" by changing the values of the influences, to get closer to a correct answer. The formula is:

$$\text{change\_influence} = -\text{error}\,\text{x}\,\text{neuron\_value}$$

To understand it, let's take the previous example, starting with the top neuron. We know all the values of the right side of the equation: neuron_value_top = 1; error = observed_response-expected_response = 0–1 = −1. As a result, the change in influence_top is 1, obtained by multiplying the opposite of the error (+1) by the value of neuron_top, and the value of this link is now 1 instead of 0. The same calculation shows that link_bottom does not change, as the equation leads to "change_influence" = (1–0) * 0 = 0. The machine has "learnt" from its error, since its structure (the values of the influences) has changed, and this change alters its future behavior. Note that when the answer is correct, the formula does not induce any change, since the error is 0, which leads to a change_influence also null. This makes sense, since the structure of the artificial brain already finds the right answer. Today, the explosion of computer power makes it possible to feed these algorithms with the resources – data and computing power – they need to learn incredibly complex tasks. To reach or exceed human performance, a learning algorithm has to contain not two but thousands of neurons connected by millions of links. To adjust the values of these millions of parameters, it needs to be fed with several million examples. This has been made possible by the creation of huge databases, such as *ImageNet*, which contains 14 million images manually categorized by "click workers" into more than 20,000 classes, such as "dog," "car," etc.

13.2 World Chess Champion

In order to understand the revolutionary logic of neural networks, let's take a closer look at an emblematic example: the world chess champion *AlphaZero*. In 2018, it defeated *Stockfish*, the world's best program, heir of *Deep Blue*, the computer that had won in 1997 against Kasparov, the then world champion.

Let's start with classic programs like *Deep Blue* or *Stockfish*. To surpass humans, they rely on two ingredients: a position evaluation function, summarizing the intuition of the game, and their enormous computing power, which allows to explore combinations several moves in advance. These programs evaluate the position using a function introduced explicitly by humans, which summarizes the knowledge accumulated by generations of champions. At the simplest level, one counts the value of pieces, knowing that a pawn is worth 1, a horse or bishop 3, etc. The scoring function used by *Deep Blue* was the most complex ever employed: it included "passed" pawns (which can become queens), the presence of rooks on the seventh row, as well as complicated calculations related to the structure of the pawns or the safety of the king. The latter considers the security of a castling on the queen side, on the king side or of keeping the king in the center, the security depending on the pieces likely to attack, the structure of the pawns in front of the king, the open columns, etc.

Instead, *AlphaZero* was only given the rules of the game and had to *learn* its evaluation function by playing against itself. For each position, the network learnt to compute two quantities: the probability of winning and the probability of playing each move, with the idea that the most probable moves should be those that are closer to victory. At the beginning, as for the *perceptron*, the values of the influences between neurons are chosen randomly, and the evaluation of a situation is very poor. The computer then plays a few hundred thousand games, mostly at random, and discovers that some positions have often led to victory and others to defeat. The neural network then integrates this knowledge, changing the influences between neurons to attribute a higher probability to the positions and moves that led to victory. The computer then uses this new network to select moves for a few hundred thousand more games and updates the neural network with the results. By repeating this loop hundreds of thousands of times, the researchers came up with a neural network that evaluates the position well enough to play better than anyone. And thanks to the current power of computers, the whole process took only…3 days!

What does this example teach us about neural networks? The essential novelty is that *AlphaZero* has developed its own evaluation function, which one would be tempted to call superhuman, because it is qualitatively different from that used by humans. It has rediscovered many of the winning tactics and strategies accumulated by experts over centuries of practice. But it has added other strategies, which are still being studied by professional players. As Kasparov says, with his usual modesty: “I was happy to see that *AlphaZero* had a dynamic and open style like mine.” But he is right to note that “*AlphaZero* gives priority to the activity of the parts, preferring positions that seem risky and aggressive.” An example among others is given by the move “b4” chosen in the position of the Fig. 13.2: instead of taking the pawn in c4, *AlphaZero* prefers to give *Stockfish* a poisoned gift, since during the whole match the black bishop will get stuck behind these pawns, leaving the hands free for white to win the match. Kasparov continues his analysis: “Chess programs generally reflect the priorities and prejudices of [either cautious or aggressive] programmers, but since *AlphaZero* is self-programmed, I would say that its style reflects the *truth*.” Has AlphaZero really managed to find the intrinsic style of chess, which arises straight from its rules? There is some truth in this: no one put by hand a predefined strategy: his style “grew” spontaneously during the learning process, from the rules but also from the precise architecture of the neural network used. So, we could test Kasparov’s idea by building many *AlphaZero*, with different architectures and learning strategies, to see if they all share the same style! Finally, it is symptomatic of the engineering of neural networks that the structure of *AlphaZero* is not reproducible, probably because of some “tinkering” or homemade recipes that the Google team did not disclose at the time of publication. As a result, an open competing machine was developed by Facebook, with the aim of making reinforcement learning “reproducible and available to researchers around the world.”

13.3 Predicting the Success of Tweets

These results are indeed impressive. But how well does the approach fare for more complex tasks? To illustrate the difficulties of machine learning for taming the real world, let’s take two recent examples. Note that there are only few works that allow a critical study, as most algorithms and data remain inaccessible.

In 2016, a team lead by Duncan Watts at Microsoft tried to predict the number of retweets generated by a particular tweet. In general, our messages remain within the circle of our followers, but some of them cause cascades exceeding 100,000

Fig. 13.2 Position of the December 4, 2017, game between AlphaZero and Stockfish

retweets. Is it possible to predict the success of a tweet, based on its characteristics and those of its sender? The paper also discusses the general limits of predictability of complex social processes, by presenting two visions, optimistic or pessimistic. The first attributes social variability to that of an intrinsic characteristic that has not (yet) been identified. For example, the strong variability of box-office success of movies would be determined by something like their "quality." In this view, variability is only apparent, and someone who could measure quality would be able to predict success or failure. The second vision sees the social world as a roulette wheel. There are no hidden variables to predict success, which results from phenomena too complex to be predictable. Retweets represent a relatively simple social feature that can be quantified without ambiguity and on which there are lots of data. For the authors, it looked like a promising field for finding social regularities and discriminating between the two visions.

Using Microsoft power, the authors managed to gather the one and a half billion tweets sent out in February 2015. By filtering usable English content, they built a database of 852 million messages from 51 million distinct users, which lead to almost 2 billion retweets. Each of these messages was described by 13 characteristics, such as: What time was the message posted? On what subject area? How many followers did the sender have? How many tweets has he already sent? The algorithm was provided with all messages sent during the first three weeks of February and their success (learning phase). It then predicted the success of the last week's messages based on their characteristics. Something like: "if the user has sent more than 53 tweets, has between 109 and 242 subscribers, and the subject of the tweet is sports," then it will be retweeted 4 times.

Their result is clear: success remains largely unpredictable. Technically, only 20% of the variability in the success of different messages is explained by their model, which is very complex and incomprehensible, as often in machine learning. It is interesting to note that the prediction accuracy can be doubled by adding a single additional variable: user's past success, his average number of retweets so far. To complete the analysis, the authors created a virtual twitter, in which they could control the respective contributions of message quality and randomness of diffusion in the network. In this model, inspired by epidemiology, tweets spread like a disease, with a reproduction number R corresponding to their quality. Authors show that a small error in estimating the quality of the message leads to a very high unpredictability. Assuming a perfectly known value of R for each tweet, the predictability of the cascades is about 80%. The deviation from the perfect predictability (100%) that could be expected in a virtual world comes from the partially random nature of the retweets. A higher quality tweet has a higher probability of being retweeted by each subscriber, but two sources of variability remain. The first is the precise number of retweets, due to the small number of neighbors likely to be "infected". The second uncertainty stems from the highly variable number of subscribers of each person spreading the message and therefore from the variability in the number of second-generation retweeters. And the unpredictability increases very sharply as soon as realistic values are introduced, assuming, for example, that there is a 40% margin of error in measuring tweet quality. This leads to a predictability similar to that observed with real data, around 40%.

My interpretation of Watts' work is that social life is intrinsically unpredictable because of the strong interactions between people. To predict with some accuracy the number of retweets of a message, one would have to know not only all its characteristics and those of its sender but also those of all its subscribers likely to pass it. And even then, we would have obtained an opaque prediction algorithm with no guarantee of transferability, i.e., with no guarantee that it could make good predictions for other months and even less for other countries. Note also that the complex algorithm obtained using all the data cannot predict retweets much better than a mere extrapolation of user's past success. Two lessons can be drawn from this work. First, that the best predictors of social systems evolutions are thick features, "soaked" with social interactions, such as this empirical measure of past success. Second, that the trivial prediction "the future success is similar to the past success" is very difficult to beat for social systems.

13.4 Predicting Life Trajectories

A last example concurs in suggesting social life unpredictability. A large scientific collaboration tried to predict life trajectories using data from the *Fragile Families and Child Wellbeing Study*. This cohort study followed 4242 families after the birth of a kid and gathered 12,942 variables about each family such as income, parental discipline, sibling relationships, vocabulary tests, etc. Despite using such a rich dataset and applying many different machine-learning methods, even the best predictions were not very accurate and were only slightly better than those from a simple benchmark model using only four predictor variables selected by a domain expert. Let's analyze in more detail the best prediction, which was achieved for the child grade point average (GPA) at age 15. Teams had access to all data (including GPA) from birth to age 9 and also to training data that included the GPA at age 15 for half of the families. Therefore, the prediction task was easier than forecasting outcomes at age 15 using only data from birth to age 9.

Even then, only 19% of the variability in GPA at age 15 was accounted for by the best predictions using complex machine learning methods with thousands of predictor variables. This score should be compared to the 11% obtained by a simple benchmark model using only four variables selected by a domain expert: three variables about the mother (race/ethnicity, marital status, and education level) and GPA at age 9. And half of the teams' predictions fared worse than this simple benchmark. Funnily enough, the team that achieved the best prediction was essentially…lucky! If their approach is used to predict the GPA of a different set of children, the results are poor, suggesting that the winning team may have been just lucky for this particular set of families. You may (by chance) predict one roll of the dice, but this does not guarantee that you will succeed in a second roll.

13.5 Computer Vision

To better understand why machine learning is a brilliant student for chess but a dunce for the social world, it is helpful to study a last example, the automatic identification of images. The whole field was revolutionized by neural network algorithms. They managed to reduce the classification error rate, that had been stable for years, at around 25%, to a few percent in years 2012–2016. These recognition algorithms then started to be used for automatic driving. But in May 2016, a Tesla car on autopilot was driving on a Florida highway, when a truck turned left, crossing its path. The autopilot failed to brake in time, and the impact killed the driver. Apparently, the neural network piloting the car did not recognize the white truck, as it appeared against a backdrop of very clear skies. Computer vision systems must, in order to drive in real time, quickly classify an object while using low-resolution cameras so as not to be overwhelmed by the amount of data. The company behind the car argued that its automatic emergency braking systems were not yet designed

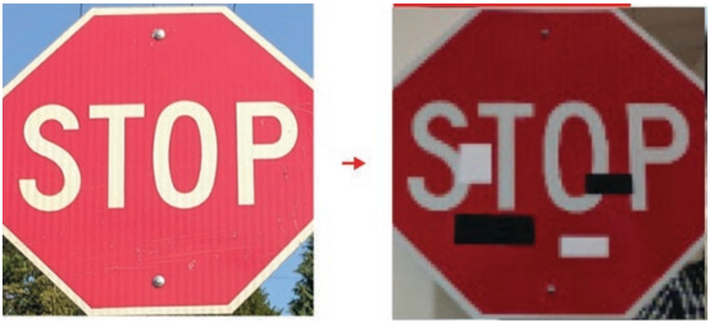


Fig. 13.3 Four stickers are enough to make a neural network interpret a “stop” sign as a “speed limit 45” one

for “side-crossing” vehicles... Autonomous cars also have to solve problems that seem trivial to us, such as assessing whether two people around the corner are chatting or preparing to cross the street. The important point for our discussion, however, is that image recognition systems are fundamentally brittle: brilliant at what they do until, taken into unfamiliar territory, they break in unpredictable ways. Indeed, the right image of Fig. 13.3, which looks similar to the left one to human eyes, is interpreted as a “speed limit 45” sign by many image recognition systems... This brittleness is not a specific flaw that could be easily overcome by more research. As François Chollet, AI engineer at Google says in *Nature*: “There are no fixes for the fundamental brittleness of deep neural networks.”

A simple example can help to better understand this brittleness. A group at the Virginia Institute of Technology taught a machine to guess whether there are curtains on a window. After training the computer with thousands of photos, for which answers were provided, the computer learned to recognize this feature just as reliably as humans, but much faster. While trying to get a sense on how the neural network decided whether or not curtains were present, the team was surprised to find that the machine was only looking at the bottom of the picture. If it identified footboards, it would stop and, without looking at the windows, say that they were well covered with curtains. The reason is simple: in all the images used to train the machine, the bedrooms contained both a bed and curtains. The presence of one was enough to infer the presence of the other. We tend to forget the obvious fact that computers learn, strictly speaking, nothing about the world. They only learn about the database provided to them. The rules they infer are so exquisitely adapted to the training data that they can fail completely outside. The problem is that it is impossible to know these limits until one has a good understanding of how the machine works, which is almost impossible when dealing with the large neural networks necessary to digest *big data*. AlexNet, the neural network that revolutionized image identification in 2012, had 650,000 neurons and 60 million parameters, making its workings (hidden in parameter values) impossible to interpret for humans. So far, there is not much theoretical understanding behind deep learning. When something doesn't work, it's difficult to figure out why. People keep trying things, until some miracle happens.

In sum, machine learning has successfully automated some cognitive tasks, with fascinating results for some framed problems such as chess. However, it remains brittle, as it is difficult to determine the limits of their efficiency, an essential point in areas where our safety is at stake. We should not be fooled by the name "intelligence," which was chosen for propaganda purposes. A more suitable name would be "cognitive automation." But the deep variability of social systems requires more than cognitive automation. It demands real intelligence, that is, as François Chollet puts it, the "capacity to adapt to unknown unknowns, the ability to solve problems across an unknown range of tasks and domains." Machines can only handle "known unknowns over known tasks." One may wonder how human intelligence manages to cope with such a variability. The precise answer is unknown, but the key point is that we are not condemned to *computation*. Somehow, thanks to our experience of the social world, we manage to identify the relevant context and apply simple rules adapted to it. Thus, to choose between two alternatives, in the absence of any other criteria, we take the most familiar one: to deal with uncertainties, we do not put "all our eggs in the same basket," or we assume that "tomorrow will be like today." These *heuristics* are effective strategies when we need to act quickly and face ill-defined problems or need to reconcile contradictory objectives. Each rule has its limits, but it seems that we often guess the one that is adapted to each situation. To understand the usefulness of these strategies, we need an "ecological" – and not simply logical – vision of human rationality. Herbert Simon compared the couple intelligence/environment to the two blades of the scissors. You can look at one blade alone as much as you like, but you won't understand why scissors cut so well. The ecological vision allows us to understand the adaptation of our reasoning to the usual contexts they encounter, which are highly variable. From a Darwinian perspective, the purpose of the organism is to survive, to find food, to reproduce, etc. Logic can sometimes help, but the rationality of our intellectual toolbox is to be sought in its adaptation to these goals and not in an abstract logical coherence.

Going Further

Pedro Domingos provides an overview of the tricks of the trade in "A Few Useful Things to Know about Machine Learning", Communications of the ACM, Vol 55, p 78 (2012). A technical introduction: Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press, http://www.deeplearningbook.org, 2016.

To understand how AlphaZero works, see the blog https://hackernoon.com/the-3-tricks-that-made-alphago-zero-work-f3d47b6686ef. The original paper is D Silver et al, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, *Science* 7 December 2018, p 1140 (Kasparov's comments are on page 1087). *Deep Blue*'s approach is summarized by its designer in Feng Hsu, IBM's Deep Blue Chess grandmaster chips, IEEE Micro, volume 19 (1999).

On deep network brittleness, see *Deep trouble for deep learning: the flaws of neural networks,* Nature, Douglas Heaven, October 10, 2019.

The example of curtains on windows: Das, A., Agrawal, H., Zitnick, C.L., Parikh, D., & Batra, D. Human attention in visual question answering: Do humans and deep networks look at the same regions? *Conference on Empirical Methods in Natural Language Processing* (2016).

On the unpredictability of the social, the article by Duncan Watts' team, Martin, Travis, et al. "Exploring limits to prediction in complex social systems." *Proceedings of the 25th International Conference on World Wide Web*, 2016. See also Duncan Watts' wonderful book, *Everything Is Obvious: *Once You Know the Answer* (Crown, 2011).

François Chollet discussion on intelligence and cognitive automation can be found in his preprint "On the Measure of Intelligence", https://arxiv.org/abs/1911.01547v2

For heuristics, see Daniel Goldstein et Gerd Gigerenzer, «Models of ecological rationality : The recognition heuristic », Psychological Review, vol. 109, 2002.

Chapter 14
Starting from Data to Hunt Causes

After studying in detail the machine learning strategy, let's now turn to simpler mathematical tools that extract information from social data, to improve our understanding, to disentangle the contributions of different possible causes, and to create a shared vision about the problems and help controlling them.

14.1 Why Is This Happening?

To disentangle the different possible causes of a phenomenon, statisticians have invented several mathematical techniques. To understand how they work, let's take the well-known case of wage inequality between men and women. In most countries, men earn more than women. In France, the pay gap is 27% and in the UK 17%. Can we attribute this gap to straight gender discrimination, that is, to a person being paid less for the same job because she is a woman? Not sure, because these are average differences, and we must make sure that there are no other factors that could explain part of the gap. For example, suppose that all women worked part-time and all men full-time: the 27% gap would in fact signal a much higher *hourly wage* for women! If we look closely at the differences in employment between sexes in France, we find that 30% of women work part-time, compared to 7% of men. If we add the effects linked to professions, some of which are overwhelmingly feminized (nurses, secretaries, etc.) and relatively low paid, and those linked to their position in the hierarchy (women occupy less often senior positions), we can explain 17% difference in pay. Clearly, these differences represent other sources of gender discrimination. However, even "other things being equal," i.e. for the same job, position, and working hours, there is still an unexplained 10% wage gap. It is tempting to see this as the effect of "pure" discrimination, even if, strictly speaking, this gap could be explained by other differences, which are difficult to ascertain from available data, such as differences in level of education or seniority in the company. To

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_14

identify rigorously gender discrimination, we should find identical people, sharing the same education, the same job, the same skill, etc. and differing only by gender. The difference in salary would then be attributable solely to their only difference, namely, their gender. Varying only one cause at a time is often possible in physics, through controlled laboratory experiments. In society, many factors vary at the same time: how can we attribute to each one its share of the effect?

The standard approach for isolating the effect of a cause was introduced by the Scotsman George Yule in 1899. It was originally designed to isolate the effect of social assistance on poverty by eliminating the effects of other factors. This type of approach allows health officials to state that "about 400 000 people die prematurely due to excessive air pollutants," even if none of these premature deaths is easily assignable. To link in a convincing way pollution and premature deaths, the pioneering study gathered data by following the mortality of more than 8000 adults in six US cities for 15 years. As these regions also differ in other aspects that have an impact on mortality (average age, smoking rates, etc.), the effects of these different factors must be "controlled" using these mathematical techniques. For example, cities with higher percentage of smokers tend to have higher mortality rates. The technique allows to correct for this, so that the specific effect of each cause can be isolated and its correlation with mortality checked. It turns out that the concentration of (fine) particles correlates extremely well with the *controlled* mortality rate, suggesting a strong causal link. In order to ascertain the link, it is however necessary to understand by which plausible biological mechanism fine particles could affect our health. Any statistical study of this type is fragile in relation to a variable that is not controlled for but could affect the result. For example, if authors had not taken into account the effects of smoking, a quite different result would have emerged.

14.2 Complex Causality: Whimsical Causes and Effects

This standard method assumes that each cause always has the same effect, whatever the context. Fine particles do increase mortality, no matter whether a person is a smoker or not, a man or a woman, old or young. The method then finds the average effect of the cause in different circumstances. However, when a cause has opposite effects in different circumstances, this approach leads to incorrect interpretations, as I will show on a real example. In the early 2000s, I had access to a great deal of data on the 10,000 researchers of the largest French research organization, the CNRS. This allowed me to study the causes that explain the promotion to senior positions. The standard statistical approach showed that the number of articles published, their impact on the community, and the age of the candidates strongly influenced success, which is in line with what we know about promotion determinants. On the other hand, the analyses showed no effect of gender: all things being equal, a woman has the same chance of obtaining a promotion as a man. This seems to contradict the low proportion of women at the highest levels of the hierarchy. To understand this paradox, we need to complicate the analysis a bit and admit that the effect of gender is not necessarily the same for all ages or all levels of academic achievement.

When you think about it, to assume that social events such as a racial conflict, a strike, or the merger of two firms can be the outcome of a complex combination of causes is not surprising. Thus, one can imagine that a strike breaks out when either a new technology is introduced *and* wages are stagnant in times of high inflation, or when there is a reduction in overtime hours *and* a contemporaneous outsourcing of production. In this case, stagnant wages sometimes leads to a strike (when a new technology is introduced), but not otherwise. Stagnant wages are neither a sufficient cause for a strike (they do not always lead to a strike) nor a necessary one (since a strike can break out through the second path). The so-called comparative method, by ordering data in "profiles," i.e., groups of individuals that share the same characteristics, allows to study complex causalities.

Table 14.1 shows that three main paths favor promotion to senior positions at CNRS. The first concerns middle-aged researchers ("Age=1", i.e. between 40 and 48 years; profiles 3, 4, 9, 10, 15, and 16). The standard approach is dominated by this population, which shows no gender effect, and represents 37% of promotions (note however that men tend to be promoted at lower impacts; see the difference between profiles 3 and 4). The complex analysis reveals two additional paths to promotion. The first concerns researchers with high impact publications ("Impact=2"), who are promoted in their 30s ("Age=0"; profiles 13 and 14). This path represents only 9% of all promotions but strongly favors men, since 14 of them were promoted, as opposed to 1 woman! A closer look at the data reveals that young women seldom apply: one for ten men, against one for three on average. The second profile leading to promotion corresponds to older scientists (over 48, "Age=2"; profiles 5, 6, 11 and 12), with an academic impact close to the average or lower. In this case, women have a higher chance of being promoted, since 22% of them take advantage of this path, compared to 14% of men. This more complex analysis allows to better understand the logic of discrimination at work. For many reasons worth investigating, young women seldom apply and only manage to be promoted much older, which limits their chances of promotion to the highest hierarchical positions. If we choose to look at the situation in a simple way, with a uniform gender effect for all individuals, we end up with an average zero effect resulting from the compensation between the biases in favor of young men and older women.

This description has been confirmed by colleagues who are familiar with the workings of promotion committees. Indeed, the standard age for promotion is around 40, but committees often promote a young candidate with high impact work. Conversely, in order not to despair the over-50s, who still have 15 years of career left, a candidate of that age is promoted, often a woman. As each committee can promote less than ten people, these two atypical promotions end up accounting for a quarter of all promotions.

Table 14.1 Different profiles observed for the promotion of CNRS researchers characterized by their age, gender and impact of their publications. Each line corresponds to a set of researchers with the same characteristics. For each profile, I also show the total number of candidates, the promotion ratio, and the corresponding number of promotions

Profile	Impact	Age	Gender	# Candidates	Ratio promoted	# Promoted
1	0	0	M	17	0.176	3
2	0	0	F	2	0	0
3	0	1	M	58	0.31	18
4	0	1	F	21	0.048	1
5	0	2	M	63	0.127	8
6	0	2	F	14	0.357	5
7	1	0	M	36	0.25	9
8	1	0	F	6	0.167	1
9	1	1	M	142	0.296	42
10	1	1	F	61	0.41	25
11	1	2	M	31	0.355	11
12	1	2	F	7	0.714	5
13	2	0	M	51	0.275	14
14	2	0	F	5	0.2	1
15	2	1	M	53	0.509	27
16	2	1	F	15	0.533	8
17	2	2	M	1	1	1
18	2	2	F	1	0	0

14.3 Farewell to the Analytical Method?

The dependence of the gender effect on age is but one example of a serious problem for mathematical methods in social sciences. Finding constitutive entities that behave in the same way whatever the context is essential to build a formal science. For then we can proceed by analysis and reconstruction. Situations are broken down into separate factors, and the effect of each factor is analyzed separately. This "analytical" method is very convenient, as it is much easier to study the specific effect of each factor in controlled environments. Once each effect is understood separately, all can be combined back to understand the initial system. For example, to find out what happens when a leaf falls through the air, one looks at the effect of gravity alone by reducing air friction, then at the effect of air resistance, and then adds the two forces to find the overall effect. This strategy is possible only because we can trust gravity to push the leaf down in exactly the same way when it acts alone and when the air resists. This is seldom the case for social systems, where complex entanglements of causes seem to be the rule. For example, to the question: "What is the effect of age on long-term unemployment?", the only possible answer is: "It depends." French data show that, for graduates, age tends to increase the unemployment rate, while the reverse is true for non-graduates. This can be interpreted quite simply, by the progressive de-skilling of diplomas: knowledge acquired during higher education, 30 years ago, is no longer necessarily as relevant. For non-graduates, the inverse effect of age is explained by the role of accumulated experience, which has enabled them to obtain the qualification they did not acquire at school. They have found their place in the labor market over the years and suffer less from unemployment than young people without a diploma. A final example: what is the effect on the number of bicycle accidents of increasing by 20% bicycle traffic? Again, it depends on how many bikes there are to begin with. If not many, the number of accidents will increase in roughly the same proportion, because each bicycle sees traffic independently of the others, and all incur the same risk. On the other hand, if there are already a large number of them, this increase may alter motorists' perception of the presence of bicycles, pushing them to reduce their speed. As a result, a 20% increase in the number of bicycles can even reduce the number of accidents!

Prabhakar Raghavan, Senior Vice President at Google, sees this dependence of effects on the context as the major challenge for the social sciences in the twenty-first century. For him, what we need to find are stable bricks, robust rules of behavior that we could then combine to make predictions, such as the combinable physical laws "that allowed us to send men to the moon," simple principles such as "friends are always the strongest influence on purchases," which would therefore be transferable between gadget or holiday purchases. Unfortunately for Google, the instability of the effects of the same cause in different contexts seems to be the rule in social systems. One reason is that social systems are designed to channel our actions, to generate causal relations that we like and rule out causal relations that we do not like. For example, the road network, its laws, and its signage are designed to help drivers avoid accidents. This dense web of mechanisms creates specific effects for each cause, as in the amusing Rube Goldberg machines. To use a trivial example from a clear review on randomized trials, we use a lever to toast our bread, but levers only operate to toast bread in a toaster. Human social experience brings informal knowledge of this causal structure and allows to anticipate the local effects of our actions, but formal methods aim at general statements that cannot benefit from such intuitive learnings.

Going Further

The importance of the stability of the effects of the same cause in different contexts is well explained by Julian Reiss in the chapter "Social capacities" in *Nancy Cartwright's Philosophy of Science*, Routledge, 2008, available at http://www.jreiss.org/papers/Social%20Capacities.pdf. For the specificity of social causality, see also Angus Deaton and Nancy Cartwright, in "Understanding and misunderstanding randomized controlled trials", *Social Science & Medicine* **210**, pp 2–21 (2018)

For comparative qualitative analysis, see the article by Charles Ragin *et al,* "Assessing discrimination: A Boolean approach", *Am. Soc. Rev.* Vol. 49, 1984.

The future of social sciences by Google: Prabhakar Raghavan, "It's time to scale the science in the social sciences", *Big Data & Society,* April 2014.

"The correlation between mortality rates and concentration of fine particles was studied in" Douglas W. Dockery, C. Arden Pope, Xiping Xu, John D. Spengler, James H. Ware, Martha E. Fay, Benjamin G. Ferris, Frank E. Speizer, (1993) An Association between Air Pollution and Mortality in Six U.S. Cities. *New England Journal of Medicine* 329 (24):1753–1759.

Chapter 15
A Moral Thermometer?

So far, I have analyzed the use of numbers in social models, showing how difficult it is to capture society workings. It is time now to turn to numbers themselves. How do we transform a complex situation into a number that can be inserted into a model or a decision? I'll focus on indicators, whose aim is to quantify a situation, someone's activities, the quality of a university, to try to describe these situations in an "objective" way. Before studying several examples in next chapter, I compare here a physical indicator (the temperature) to a social one inspired by it: the happiness index. At the end of the eighteenth century, Jeremy Bentham, the English philosopher famous for his model prison, the *panopticon*, proposed to build a "moral" thermometer, to objectify happiness. He wrote in his principles of the civil code: "Legislation, which has hitherto been founded principally upon the quicksands of instinct and prejudice, ought at length to be placed upon the immoveable base of feelings and experience: a moral thermometer is required, which should exhibit every degree of happiness and suffering… The feelings of men are sufficiently regular to become the object of a science or an art; and till this is done, we can only grope our way by making irregular and undirected efforts." This moral thermometer is the ancestor of the current gross domestic product (GDP), which is supposed to capture the happiness, or at least the wealth, of societies. Bentham also foresaw the central role that this indicator would play in the "objective" orientation of public policies toward increasing happiness. Yet, unlike the physical thermometer, which is hardly challenged as an indicator of heat, GDP remains highly contested. By comparing these two ways of transforming complex situations into numbers, we will better understand the difference between the objectivations proposed by the natural and social sciences.

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_15

15.1 How Hot Is It Today?

At first glance, heat phenomena seem straightforward. Heating water in a pan increases its temperature, as you can check by soaking your fingers. In winter, wearing a "warm" sweater protects you from the cold. This simplicity is only apparent. Thus, we feel that a tiled floor is colder than parquet, even when they are in adjacent rooms, at the same temperature. A thermometer would show that continuing to heat boiling water does not increase its temperature. A "warm" sweater also helps to keep…cold, for example, a drink taken out of the fridge on a summer afternoon. Finally, some people are more skittish than others and need more heating to feel comfortable. All things considered, heat may well be more complex than anticipated…. It took scientists three centuries to tidy up these phenomena, by separating three main components: heat, temperature, and thermal conduction. For example, scientists would now say that the tile "feels" cold because stone conducts heat faster than wood and it absorbs our body heat more easily, giving us a feeling of coolness. In this chapter, we will follow the slow construction of thermometers, which manage to quantify, in an "objective" way, one of these three ingredients, temperature.

The modern history of temperature begins in the seventeenth century. Santorio, a Venetian doctor and Galileo's friend, invented a "thermoscope," a simple graduated tube filled with a liquid. The dilatation of the liquid, measured by the graduations, allowed to quantify differences in fever, making diagnosis more reliable. However, the thermoscope soon proved to be insufficient. Because each instrument was different from the others, it recorded the patient's condition in its own way. Two people with the same fever could be characterized by "3 graduations" by one thermoscope and "2 and a half" by another. The budding scholarly community also wanted to be able to compare heat levels at distant places or times. In France, Réaumur was "absolutely curious" to know how hot the tropics are, as they seem to be burned by the sun's rays falling straight on them. For if it has been a long time since "we know that the torrid zone is not uninhabitable, as the Ancients believed, we have no idea how much heat we have to suffer there." Scientists also wanted to compare present and past climates. They wondered whether the chilly Paris 1776 winter had been colder than the famous 1709 one. Beyond these specific questions, the growing importance of steam engines made the precise measurement of temperatures and heat quantities of paramount importance.

As early as 1665, the English scientist Boyle had proposed to standardize thermometers, in order to arrive at a notion of degree as easily exchangeable as those of distance or weight. The first idea was to define a standard model of thermometer, which would be manufactured everywhere in the same way, like the standard meter. This simple solution proved impossible in practice. A single manufacturer could obtain thermometers that were fairly comparable with each other, but it was impossible to extend this uniformity to several manufacturers. The problem is that the dilatation of liquids is small, leading to tiny differences in volume. Mercury increases its volume by only 1.8% between 0 and 100 °C. And, by an elementary geometric effect, identical differences in volume are only transformed into identical differences in height if the tube is perfectly straight, if the liquid is exactly the same, which was never the case in a time when matter had not been standardized. A common joke at that time was: if you have a thermometer, you know the temperature, if you have two, you no longer do!

15.2 The Construction of the Scale

The variability of each instrument was therefore acknowledged, and a more complex approach was attempted. The idea was to calibrate thermometers by arbitrarily defining two stable and easily reproducible physical phenomena, to which degrees 0 and 100 would arbitrarily be assigned. This would guarantee that all thermometers give the same indication at these two points. What about intermediate values? Jean-André de Luc came up with a clever idea: use two containers of water at degrees 0 and 100, and mix them in the right proportions. By convention, a mixture of one liter of each would lead to two liters at degree 50. Unfortunately, experiments showed that the amount of heat needed to bring the water to 100 was not twice as much as that needed to reach degree 50. In other words, the "heat capacity" of water, the amount of heat needed to raise its temperature by one degree, seemed to change with temperature. This made the de Luc scale not very intuitive, all the more because these variations are difficult to study as long as the temperature scale is not defined independently. To get out of this vicious circle, instruments were simply calibrated at two fixed points, and it was assumed that their variations between these would be similar. Many possibilities were proposed for the 0 and 100 points of the temperature scale. For the lowest temperature, some favored a deep cellar, others the freezing point of aniseed oil, melting butter or melting snow. After many tests, a mixture of water and ice was selected, as it was easy to produce and gave a stable mark. For the highest temperature point, the vicissitudes were just as numerous. In 1699, Edmond Halley found that the indication of a thermometer immersed in boiling water was constant, even when heat was continuously added. He proposed to take advantage of this stability and use it as a fixed point. Long controversies ensued, with Newton and many others, who believed that the thermometer's indication did depend on the boiling time. This controversy was deemed insoluble by a report of the Royal Society of London in 1777. Others interpreted the differences as resulting from impurities and tried to resolve them by using purified water. This only further confused the issue, as the discrepancies between different thermometer readings increased! Today, we explain the controversy by the essential role played by impurities in starting the boiling process. This causes the measured boiling temperature to vary, depending on the circumstances, between about 98 and 105 degrees. However, a consensus was eventually reached on "moderate" boiling, which reduced the variations.

But the problems weren't over yet. As the use of thermometers grew, especially for precise monitoring of industrial processes, inconsistencies were discovered, even after calibration. For example, three thermometers calibrated at 0 °C and 100 °C gave different intermediate measurements depending on the liquid used. When the mercury thermometer indicated 50 °C, the alcohol thermometer measured 44 and the water thermometer only 26! The expansions of the three liquids were clearly not similar. Worse still, systematic measurements taken by Henri Victor Regnault in 1842 showed that even thermometers using a single liquid were unreliable. The issue was not purely scientific, because steam engines needed reproducible measurements. Indeed, Regnault's memoir was entitled: "Relationship of the experiments undertaken by order of the Minister of Public Works, and on the proposal of the Central Commission of Steam Engines, to determine the main laws and numerical data used in the calculation of steam engines." The problem arose from a hitherto neglected detail: when the heat increases, the liquid expands but so does the glass of the thermometer! This expansion changes the diameter of the capillary and thus the volume in which the mercury expands, affecting the reading on the scale. And the problem gets worse for temperatures above 100 °C, a territory that had become important for manufacturers. Regnault tried to stick to glasses of the same nature, having undergone exactly the same manufacturing process. After a series of systematic measurements, deviations of several degrees remained. Given the precision required, glass variability made mercury thermometers unrecoverable. Fortunately, Regnault's work showed that air thermometers could be trusted, because strong air expansion makes that of glass negligible. Another key advantage was that their indications did not depend on the density of the air enclosed in the thermometer, as long as it remained fairly low. As a result, air thermometers made by different manufacturers, with different glasses or air densities, gave the same indication. By finding a sturdy material anchor, Regnault had put an end to the long quest for objectification of temperature measurement.

15.3 An Absolute Temperature

Well, this was not the opinion of one of the most influential scientists of the time, Lord Kelvin. In an article published in 1848, he reviewed previous research and agreed that, from a practical point of view, Regnault's work was conclusive. However, he believed that "the theory of thermometry is still far from satisfactory." Two points seemed problematic to him. The first concerns the very principle of the construction of a thermometric scale and the second the fact that the temperature scale proposed by Regnault is intrinsically linked to air. But first of all, it should be noted that Kelvin does not mention, probably because it seems obvious to him, the primary condition for success: the existence of a thermal equilibrium to define the very idea of temperature. If the thermometer can give an accurate indication of heat, it is because equilibrium ensures that "its" temperature is the same as that of the system it is in contact with. This empirical equality, which is valid in a large number of cases, is essential because it makes it possible to link many situations and thus to control them better. Today, this point constitutes one of the four founding principles of thermodynamics, the science of heat processes.

Let's go back to Kelvin's first point, the thermometric scale. For him, a perfect thermometer should "indicate that equal additions of heat correspond to equal increases in temperature, estimated by the numbered divisions of its scale." He acknowledges, however, that this elegant scale is impossible to achieve, as "it is now recognized as an experimentally proven fact that thermometry under this condition is impossible." We already met this problem with de Luc's water scale. Experiments show that, for a unit of heat injected into a body, the rise in its temperature depends on the body chosen and also on its initial temperature. But Kelvin had a solution to the second problem, by proposing a thermometric scale that does not refer to a particular body, now known as Kelvin's "absolute" temperature. His proposal, related to theoretical work carried out by Carnot on ideal machines, is too technical to be detailed here. It is also impractical to turn into a working device, but it can be approximated by a thermometer using a gas under very low pressure. And the important point is that, whatever the nature of the gas used, all such thermometers would indicate the same temperature when the gas pressure is close to zero. Kelvin's work completed the slow and painful task of objectifying the temperature. Thermal equilibrium ensures that the thermometer does not invent an arbitrary quantity but draws a characteristic from the observed system. And Kelvin scale produces a number that depends neither on the researcher measuring it nor on an arbitrary choice of the substance used to build the thermometer.

15.4 Temperature as an Institution

This quest for objectivity can be summed up by taking a perspective that may seem odd at first glance: absolute temperature as a social "institution." French sociologist Luc Boltanski has suggested that the role of institutions is to gather social actors, by setting a common reference point that goes beyond their particular points of view. Since "human beings possess a body and are therefore necessarily situated," each individual can only have one point of view on the world. It is therefore necessary to "delegate the task of telling what is [real] to a bodiless body," namely, an institution. The example of temperature highlights the essential role played by scientific objects in the construction of many social institutions. As human beings, we often disagree about how hot it is. But thermal equilibrium makes it possible to create a dimension of heat that is shared by all bodies, at least inert bodies, as we shall see. To the vague question: "is it hot?", each body gives a different answer, depending, for example, on whether it is close to its melting point or not. For a plastic, 80 °C is very hot; for gold it is very cold. But to the question "what temperature is it?", all bodies answer the same thing. And the thermometer based on a (nearly) perfect gas gives us the example of a being literally "without a body," capable of escaping the constraint of the point of view and fixing "objectively" the value of the temperature. Thanks to ingenious assemblies, the collective of experimenters managed to stabilize the response of this thermometer, which became a reliable spokesman for the degrees of heat. In this sense, the temperature is an institution that allows to go beyond particular points of view and give an "objective" meaning to one of the dimensions of heat, as it has withstood all relevant objections.

15.5 Building the Moral Thermometer

I have dwelt at length with the history of temperature to show how difficult it was to transform heat into an objective number. Of course, the challenge to objectify wealth or happiness using a "moral thermometer" is far more important. The first step toward it was achieved a century and a half later by American economist of Belarusian origin Simon Kuznets. He was looking for a simple indicator to characterize the state of the economy after the 1929 crisis and more broadly to compare different countries and years. His quest was similar to Réaumur's, who wanted to compare the "degrees of heat" in France and in the tropics. Kuznets' goal was to calculate "the total activity of a national economy." In order to add various activities without falling into the problem of adding cabbages and carrots, the totality of market exchanges had to be made homogeneous. To do this, monetary valuation based on market prices is used, since the latter are supposed to reveal the relative values attributed to different products by those who buy them. If we are prepared to pay ten times more for a kilo of coffee than for a kilo of potatoes, it is because we have a tenfold preference for coffee. Of course, the reality is more complex, because many goods or services are priceless (domestic childcare, household activities, etc.), because some activities included in the GDP are harmful (activity generated by pollution or road accidents) or do not enrich the country's households (profits of companies domiciled for tax reasons, but which repatriate their profits, as in Ireland). Moreover, using an average GDP per capita does not take into account the inequalities of distribution that play a major role in the well-being of the inhabitants.

These points are well known, and I will only take one example, the exclusion of unpaid household activities, which would account for about a third of total GDP. Kuznets justified the exclusion by the difficulty of estimating their value. There were two or three methods to quantify them, and they lead to very different results. Should domestic activities of a person be assigned a monetary value based on the hourly cost of a housekeeper or rather on the hourly cost of his non-worked overtime hours? Should cooking time be counted as a chore or considered as an enjoyable leisure activity? As an economist friend of mine said, everyone knows that GDP is not a good measure of wealth, but at least its current definition, governed by international standards, allows comparisons between countries and years. In other words, the priority is not fidelity to reality but stability in its calculation. This could be called "weak objectivity," in the sense that the way GDP is calculated has met the objections of some experts, making it difficult to manipulate. But GDP is not as objective as temperature, which has been able to meet the objections of the entire scientific community, but also those of all the materials studied, to which researchers have lent their voice through their experiments. The objections of the materials did not allow the researchers to found a notion of temperature that would have suited them well, in which, as Kelvin said, "equal additions of heat correspond to equal increases in temperature." In attempting to achieve temperature objectivity, GDP should take into account the definitions of (economic) well-being of all those concerned. And this is of course a political issue, not least because this indicator is used to guide public policy. Translating it into quantified data should not give the illusion of technical neutrality, because as Alain Desrosières has shown, quantifying is first "agreeing" and then "measuring" (see chapter 12). As in the example of household activities, we first have to agree on values and political priorities that create codings and replicable procedures and only then proceed to measurements, as a regulated implementation of these conventions.

15.6 The Perceived Temperature

The history of temperature has taught us that a long experimental work, based on stabilities such as thermal equilibrium, is needed to quantify, in an objective way, a complex phenomenon. No objectivity without stability! And this is true not only for GDP, as we shall see by returning, for the last time, to Réaumur's question: how much do we suffer from heat in the "torrid zone"? Is temperature the right indicator to answer this question? Not sure, because we humans never (as long as we are alive) reach thermal equilibrium with the environment.

Air temperature measured by thermometers is only one of the many parameters that determine how we feel. Our thermal sensation depends on the details of the heat exchanges between the body and the atmosphere, as our body maintains an internal temperature close to 37 °C. Basically, if we lose heat, we feel cold, and if on the contrary we have difficulty evacuating it, we feel hot. These exchanges are affected by meteorological and physiological factors. Among the former are wind speed, humidity, solar radiation, heat emitted by the ground, etc. In cold weather, a strong wind quickly carries away the heat emitted by the body, leading to an additional feeling of cold. Physiological factors include the thermal conductance of the body, which depends on the percentage of fat (a good insulator!) or the type of activity, which changes the internal temperature. Is it possible to "objectively" quantify these sensations, which reflect such complex processes? How can we best represent the thermal exchanges between the body and the environment?

The first method, developed as early as 1941 by two American explorers in Antarctica, consisted in measuring the freezing time of the water contained in a small bottle, according to the temperature and the wind. In strong winds, the water froze much faster, and they derived an index quantifying the effect of the wind. But, in use, we realized that this index greatly overestimates our feeling of cold, because it does not take into account the reactions of the body, which produces more heat to compensate for the effect of the wind. The indicator should summarize the effects of air temperature, wind, humidity, sunshine, etc. leading to a single number for the same sensation, even if the values of the variables are different. If a temperature of 5 °C and a wind speed of 40 km/hour lead to an index value of 1, and this index value is also obtained for a temperature of 1 °C without wind, the body must feel the same in both these environments. The simplest way to condense the complex effect of the atmosphere on our body is to use as an index the air temperature which, in the absence of wind, sunshine, etc., provokes the same physiological response. The question that remains to be answered is: whose physiological response? And here we come back to the same type of problems as for GDP, because there is no "objective" answer as there is for temperature. Is it necessary to calculate the exchanges of the whole body, or just the face, which is generally more exposed? What value should be taken for the conductance of the body, which depends strongly on the percentage of fat? An influential group has chosen to build a standard spokesperson in the person of Michael, a 35-year-old man, measuring 1.75 meter and weighing 75 kg, wearing a tie in winter and sandals in summer. Too close to the average German male to be an "objective" spokesperson for this institution, Michael is in fact a virtual being, a physiological model that allows the response of his "organism" to be calculated in a reproducible way. But of course, this virtual body is only an average, while physiology varies enormously with age, stature, or gender. The thermal conductance of the face, which determines how easily our face loses its internal heat, varies by a factor of 5 from one person to another. In short, depending on the precise hypotheses chosen, perceived temperatures can vary by several degrees and in any case will not represent the range of personal sensations well. As with GDP, one may question the usefulness of announcing a single figure, which gives a false impression of rigor. By knowing the temperature and other characteristics of the weather, each of us is able to estimate our own thermal sensation, taking into account our physiology, our state of fatigue, the type of activity we are going to carry out, and exposure to the sun, all of which are impossible to deal with in a standardized manner. To know how much he would suffer from the heat and humidity the tropics, Réaumur cannot count on temperature, even the perceived one: he will have to make the trip!

Going Further

On the history of thermometry: Thomas D. McGee, *Principles and Methods of Temperature Measurement,* Wiley, 1988; W. E. K. Middleton, *A History of the Thermometer and its Use in Meteorology,* Johns Hopkins University Press, 1966; Réaumur, "Observations du thermomètre faites par M. Cossigny", *Mémoires de l'Académie royale des sciences,* 1733.

The moral thermometer was proposed by J. Bentham in his *Principles of the Civil Code,* Part I, Chapter VI, 1838.

On the history of temperature, see Hasok Chang, *Inventing Temperature,* Oxford University Press, 2004, and his experiments on variations in the boiling temperature of water: http://www.sites.hps.cam.ac.uk/boiling/.

J. Regnault, *Mémoires de l'Académie royale des sciences de l'Institut de France,* **v 21**, 1847

On GDP, its construction and its limits: Simon Kuznets, *National Income, 1929-32*, Washington, D.C., Government Printing (1934); Stiglitz, J. E., A. Sen and J.-P. Fitoussi, *Report by the Commission on the Measurement of Economic Performance and Social Progress* (2009).

On Michael, the standard spokesman for heat sensations, Jendritzky and Nübler, A model analysing the urban thermal environment in physiologically significant terms, *Arch. Meteor. Geophys. Bioclimatol* 1981.

Chapter 16
Where Do Indicators Lead Us?

What modern Pythagoras, what Einstein of our own age
can determine with unquestioned accuracy
the proportionate share of the benefits
to be derived from the construction
of reservoirs in distant lands?

TH. BILBO, Senator from Mississippi, 1936, cited by Th Porter

The previous chapter showed that social indexes lack the "objective" foundation of physical ones, based on stable relations such as thermal equilibrium. By looking at several concrete examples, I would like to suggest that, despite their fragility, some social numbers can, if appropriately used, play a positive role in modern societies.

16.1 Compute the Profitability of Public Projects

Cost-benefit methods were invented in the beginning of twentieth century to assess public investments in an "objective" way, by computing both the benefits expected and the costs incurred. Today, these methods are used to evaluate most public initiatives, as governments seek to avoid the accusation of clientelism, of favoring some companies or political allies. For example, for the project of airport near Nantes in France, costs encompass construction works, profits of private operators, and also the damage to the environment or the loss of farmland and forests. Benefits included time saved by users and deaths avoided by improving the infrastructure. In principle, there is no limit to these two lists. But in practice, if a single indicator is to be achieved, it is necessary to homogenize all these ingredients, by assigning to each one a monetary equivalent, which is difficult to obtain objectively. For example, a recreational value of 126 euros per hectare per year is assigned to the forest, but estimates vary enormously depending on the methods used. The same applies to the value of a death avoided, estimated at around one million euros.

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_16

This amount is computed by taking into account the investment made by the community for the "production" of the individual (education, etc.) or the moral prejudice of a death, judged on the basis of the compensation paid by insurers. These different approaches lead to a range of 20–1500 thousand euros per death avoided, the official value being one million euros.

Let us return to the case of Nantes airport. JLR Conseils design office estimated that its construction would bring a benefit of 911 million euros to the community, to be compared to the cost, estimated at 560 million. This large positive difference justified the construction of the new airport. But opponents soon started checking out these figures. The calculation was based on the assumption that if the new airport is not built, the saturation of the current airport will force many passengers to transfer to other airports in the region. This causes a double loss of time for users, who will experience longer road journeys to reach these airports, and also additional "virtual air time." The latter comes from the reduced destinations offered by these regional terminals, which forces users to make additional stopovers, generally via Paris airports. This lost time is then transformed into money, using a standard "value of time" of around €10 per hour. The overall benefit is obtained by multiplying the estimated number of passengers in the coming decades (until 2042) by the amount of time (and money) saved. Opponents replied that extrapolating passenger traffic up to 2042 from users' surveys carried out around year 2000, before major changes that occured in airline business, was meaningless. Moreover, the result depended on questionable additional assumptions. For example, oil or carbon prices were likely to rise, making air fares less attractive and reducing the number of passengers. Similarly, the saturation of the current airport was not clearly established, as the (hypothetical) additional demand could anyway be absorbed by using larger aircrafts. The counter-expertise carried out by the opponents argued that the initial analysis was "flawed" and that, in the end, "the construction of the new airport represented an overall cost to the community." This specific example shows a general feature of cost-benefit analysis: the figures produced by the experts are fairly easy to handle. But at the same time, the legal requirement to make them public allows both open criticism and counter-expertise, and it seems more transparent than a discussion between experts not based on explicit figures.

16.2 Ranking Universities

Shanghai's ranking of the world's universities has become an obsession for many public officials. In 2011, French Minister of Higher Education was happy to announce that the new Shanghai ranking would reward "the efforts that have been made […] The results are extraordinary. Four groups [of French universities] could directly enter the top 50." Does this satisfaction reflect an improvement in the quality of research and training at French universities? Not really, because the ranking improvement hides a purely mechanical effect. The point is that Shanghai index does not measure the productivity of an institution's researchers or the quality of teaching but rather a mass effect. Indeed, it is obtained by summing six weighted scores: number of students who have won a Nobel Prize (weight of 10%), number of researchers who have won a Nobel Prize (20%), number of highly cited articles (20%), or published in *Nature* or *Science* (20%) or in international journals (20%), and finally the productivity obtained by dividing the sum of these five scores by the number of permanent researchers (weight of 10%). Two criticisms of the alleged objectivity of the ranking come immediately to mind. First, what would be the objective justification for the weights of each element? Second, this simple accumulation mechanically favors large institutions. If only the score divided by researcher (the last criterion) were used, the Ecole Normale Supérieure from Paris would climb from 69th place to the…4th! Wouldn't this weighted score be more relevant for a student looking for stimulating teaching?

16.3 Evaluating Researchers

A final example is the quantification of the quality of research. The management of the largest French research organization, the CNRS, is periodically tempted to evaluate its researchers automatically, to make promotions quicker, more "objective", and above all to avoid reviews by colleagues, who are suspected of promoting buddies. In practice, there are several possibilities for estimating research quality. The most reliable is a thorough reading of the work by competent colleagues, followed by a direct discussion with the scientist about the meaning she attributes to the research. However, this assessment takes time and is subject to the interpretation, necessarily subjective, of the importance of some works compared to others. In order to make the evaluation quick and objective, the first step was to measure the production of researchers by counting the number of articles they publish. However, this indicator soon showed its limitations: first, because it is easy to slice an article into several pieces. But above all because not all articles are equal: some are published in obscure journals that will accept anything as long as you pay for it, while other papers become classics, opening up new fields of knowledge. It therefore seems more relevant to evaluate researchers by the impact of their work on the scientific community. This can be done by counting the number of "citations," i.e., the number of times their work has been used by other articles, thus demonstrating its importance. To classify researchers, it is sufficient to compare their number of citations, normalized by years of career, so as not to systematically give an advantage to the oldest researchers. This index is indeed more reasonable than the simple number of articles and less easy to handle. But it is far from perfect. First of all, one may cite an article to criticize it, or just because it has already been cited by others, or because the researcher knows how to sell his papers at conferences. Moreover, citations only measure the take-up of research by other colleagues, but scientists also teach, register patents, write popularization books, manage research teams, collaborate with associations or companies, etc. All these activities are important, but they do not directly advance the citation counter.

To judge an indicator, it is essential to follow its concrete use. In some cases, the citation index leads to hasty evaluations, but it can also allow to query claims made by colleagues with strong hierarchical authority. One of them attempted to promote one of his friends, arguing that he had a huge international influence. But committee members used the indicator to show that he was never cited. Moreover, experts can put indicators into perspective. For example, a researcher who had accumulated a large number of articles and citations was evaluated negatively, because his work was not considered creative enough. He ran a large scientific instrument used by a huge number of colleagues, and systematically coauthored their articles, even when his scientific contribution was insignificant.

16.4 The Origin of the Indexes

Before discussing the interests and limitations of indexes, it is useful to briefly review their history, which dates back to the twentieth century. One could imagine that the quantification of society is the logical continuation of the quantification of nature promoted by physicists. History shows that this is not the case: indexes have arisen from the public's mistrust of the subjective judgement of experts. It is then understandable that their use did not begin in France, a country with a strong tradition in natural sciences and mathematics, but in the United States, which had a much weaker scientific culture but a stronger mistrust of technocracy. French experts, protected by public confidence in state expertise, had not attempted to base their decisions on explicit quantification. They believed that informal discussions, based on mutual trust, were more likely to produce reasonable compromises. It was not until after the Second World War that quantitative analysis of the profitability of public works would be imposed in France and in the rest of the world, under American influence. From the 1940s onward, the public engineering corps in the United States had been involved in strong public controversies over the construction of dams, particularly on the Mississippi River. To defend the objectivity of its expertise, it was then forced to comply with standardized rules for calculating costs and benefits. Objectivity had become a rhetorical argument, a sign of minimal human intervention, for in our modern societies, expertise based on personal experience seems obscure and potentially undemocratic. It is defenseless in the face of suspicion, corruption or favoritism because, for nonexperts, it is difficult to distinguish a judgment validated by experience from a subjective bias.

16.5 Indexes Are Not Perfect…

It is understandable that the supposed objectivity of indexes is attractive to public officials: their use allows decisions to be made without appearing to decide. Measuring their evolution also makes it possible to give the image of concrete and (possibly) effective action. As a mayor said about air quality: “If you don’t measure it, it doesn’t get any better!” There are many arguments that call into question the objectivity of the indexes: how to choose the ingredients deemed relevant and discard all the others? How to weight them? For example, to quantify the number of “poor” people, one has to decide whether to fix a threshold using absolute purchasing power or relative to the minimum wage; consider either individuals or households; include or not social and family benefits. All these choices are justifiable but lead to different numbers. We have seen that these questions have no “objective” answer. And most of the indexes are fragile: apparently minor changes in definitions may dramatically change the results, undermining claims of rigor, as for the Shanghai ranking of ENS Paris.

16.6 But Can Be Useful!

The rhetoric about the objectivity of indexes should not be taken too seriously. Assessing a situation by following predetermined rules, like a computer program, avoids individual bias, but it is not necessarily a good strategy for seeking the truth. However, a well-conducted quantification allows to represent and compare the complex situations we are living through, which is indispensable for orienting our actions. Trustable unemployment figures assess the real economic situation and prevent the government from arbitrarily claiming success; inequality records for educational attainment according to social class make it possible to go beyond the causes of individual failure, by pointing out the overall logic of social reproduction. International comparisons make it possible to put each country’s results into context, to show for example that some school systems manage to be more egalitarian than others. Finally, recent work by the economist Thomas Piketty shows the usefulness of numerical indexes of wealth and its distribution. By allowing a comparison between the present and past centuries, they show that after a few decades of unregulated capitalism, we are moving toward inequalities close to those of the early nineteenth century. Of course, none of the above results is totally objective or indisputable. But the publication of the calculation rules for each indicator makes the results verifiable and open to criticism. True, criticism is restricted to experts, as only they can judge the robustness of the hypotheses, their weak points, or the ingredients that are easy to manipulate. Use of indexes must be well supervised, by organizing a culture of counter-expertise and making it financially accessible to associations. This would make it possible to create a counterweight to the state technocracy, as we saw in the example of Nantes' airport.

16.7 Where Does This Indicator Lead Us?

The lesson to be learned from this chapter is that criticizing the imperfections of social indexes is necessary but insufficient. The interesting question is rather: toward what future, what concrete actions does the use of this particular indicator lead us? In the case of growth, we know that gross domestic product (GDP) has been constructed with a definition of wealth that does not take into account the depletion of resources or income inequalities. The construction of alternative indexes is therefore essential to steer our actions in directions that are more respectful of the environment and equality. More generally, *benchmarking* – the production and publication of ranking lists that make it possible to establish a hierarchy of organizations (companies, schools, administrations) – creates a reflexive looping effect according to the logic of self-fulfilling prophecy. For example, the Shanghai ranking list builds a homogeneous global space for ranking universities, locking us into a logic of globalized competition for international visibility. This competition reinforces existing inequalities, because the most visible universities attract the best students, making them even more visible and so on. But does this international visibility point toward socially interesting research? The example of economics, dominated by the Anglo-Saxon school, which has proved totally incapable of understanding the financial crisis, leads us to doubt it. Would it be possible to produce a better indicator that would point us toward an academic world that is more interesting for society?

The question of guiding our actions by indexes becomes even more acute when they assess individual performance. If the CNRS decided to use citation indexes to award promotions, there is no doubt that researchers would implement appropriate strategies to improve their scores artificially. A team could decide to include all its members in each of its publications, as the indicator does not take into account the number of authors of an article. And even if it were changed to prevent this type of manipulation, researchers could adapt by exchanging citations. Similarly, when the work of police officers is evaluated by numerical indexes, the possibilities of "tampering" with them are infinite. As one of them told Le *Monde* on December 5, 2003: "You don't cheat with numbers, but you get smart." For example, "three young people arrested with a joint in their beak by a night school class, that's three facts, three police custody and three elucidations. As a result, the bosses prefer this to a painful case of trafficking" which only counts for one. It is a well-known rule: when a measure becomes an objective, it stops being a good measure. More importantly, this type of individual evaluation, by numerical objectives, perverts the action, by imposing an external meaning on it, out of step with the deep meaning attributed to it by the actors. The aim of researchers is to improve knowledge, not to accumulate citations. The generalization of such evaluations would destroy the very meaning of our actions, turning us into pure machines for improving indices, leading us to look for meaning elsewhere, in a religion, or a national affiliation.

Going Further

Essential, on the quantification of society: Theodore Porter, *Trust in Numbers,* Princeton, 1995, from which the epigraph is taken.

Chapter 17
Which Numbers for the Future?

I've so far presented the many ways in which our life is transformed into numbers, to feed mathematical models or create shared knowledge. I'll now put these examples into a wider perspective, by showing that numbers are tightly linked to the rise of modernity, the emergence of human societies connected at a global scale. Numbers are invaluable to transport information through long distances. Therefore, they are subject to an inherent tension: they allow powerful centers to control populations, and at the same time they are indispensable for bottom-up coordination of people on a large scale. The history of social numbers makes clear that they have so far mostly benefitted power centralization. But the digital revolution has democratized numbers, which could now be used for planning a common future, to master the ecological crisis.

17.1 Social Numbers Were Invented by Centralized Powers

Why have social numbers emerged? A key insight is provided by the etymology of the first mathematical method invented to deal with them, namely, "statistics." This word originates from the Italian *stato*, state. Initially, it referred to the body of knowledge useful for governing a country and did not include mathematics. The decisive impetus for the combination of statistics and mathematics was given by the avalanche of social data generated in the nineteenth century by modern centralized states. States counted populations and wealth produced, to better collect taxes or enlist soldiers, and generalized the use of those supervisory tools that we now take for granted, such as maps, land registry, homogenization of units of measurement, stabilization of surnames, etc. This control required the establishment of a legal and material infrastructure, an investment similar to that of road or rail networks. In short, states transformed their territories and their inhabitants to make them governable from a center, thanks to mathematical tools invented by the scientific elite. For

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_17

example, Pierre-Simon de Laplace, the great astronomer and mathematician, who was Napoleon's minister in 1799, developed different approaches to estimate the French population based on limited data, because it was difficult to carry out an exhaustive census. He assumed that the number of births per inhabitant was more or less constant throughout the country, an assumption he tested in about 30 regions carefully chosen to be representative of the entire territory. Then, by accumulating data on the number of births, which was well known from parish registers, the central government could obtain an estimate of the total population.

Circulation of numbers on large scales was also driven by the needs of the emerging industrial revolution. The invention of the steam machine, the train, and the telegraph dramatically lowered transportation costs. These discoveries, together with the unused land provided by colonies and the energy provided by coal, lead to the creation of a global technological network. This network allowed to extract materials and energy and transfer them to huge factories, thus benefitting from economies of scale to manufacture large quantities of goods at low cost. Monitoring this global network necessitated the little-known "control revolution," to regulate the flow of information. At the beginning of the train era, it was easy to avoid accidents on short lines by using single cars, back and forth. With increasing traffic and line interconnection, one needed the telegraph, to circulate information about delays and incidents faster than trains. Moreover, gathering information on daily and monthly traffic according to prices avoided half-empty trains, optimizing flows and profits. Modernity had an enormous need for numbers to extend its grip on the world.

17.2 Lost Individuals

> Reality has long ceased to be what it was for my grandmother, who lived in a Moravian village and still knew everything through her own experience: how bread is baked, how a house is built, how a pig is slaughtered and the meat smoked, what quilts are made of, what the priest and the schoolteacher think about the world […] She had, so to speak, personal control over reality.
>
> Milan Kundera, Immortality

The transition from small communities to an era of bigness, big industry, big market, big government... meant moving from intimate knowledge and face to face trust, to impersonal relations and trust in abstract numbers. Consumption that used to be local, supplemented with some exotic imported foodstuffs (tobacco, spices, etc.), now became interconnected and homogenized. Unit measures used to be tied to the local context and to the particular qualities of the thing being measured. To the question "How far is it to the next village?", people in Asian villages could reply: "Three rice-cookings," assuming that the questioner is interested in how much time it will take to get there, not how many miles away it is. In varied terrain, distance in miles is an utterly unreliable guide to travel time, as if you measured the contents of books by their weight. But these locally meaningful measures made global markets or regulation by central authorities impossible. The creation of global markets demanded abstract measures, indifferent to the specific qualities of the objects measured, and technical standards, allowing to build interchangeable parts everywhere. The coordination of large human communities was possible through the standardization of reality, the contraction of the world to the communicable.

As entire sectors of activity crossed national borders and escaped traditional state controls, individuals lost control of their existences, becoming vulnerable to mechanisms of which they were unaware and in the face of which they found themselves powerless. These lost individuals lead to the great upheavals of the time. Fascism managed to capture popular discontent by tying back disoriented people to an "identity" constituted by their land, to their nation, as for peasant communities before modernity. Socialism, instead, embraced modernity and aimed at transforming it to socialize its benefits. Centralized planning by the state was one of the privileged tools for doing so. For example, China's communist revolution in 1948 managed to industrialize the country, thanks to a high rate of investment, of the order of 40% of GDP, which would have been difficult to reach with private capital. Combining this centralized industrial policy with health coverage and alphabetization, the country managed to lift hundreds of millions out of poverty, a unique achievement in human history. Communist planning had the ambition to achieve a full centralized control of the activity of enterprises by the administrations in charge of the elaboration, implementation, and control of the plan. The central organisms decided not only the allocation of resources and investments but also the production and distribution of all goods and services, in their smallest details.

A more flexible planning, combined with markets, was also implemented in capitalist countries. From the end of the nineteenth century, driven by social struggles, states had begun to collect data on wages, working hours, or workers' budgets, in order to organize protection against unemployment, illness, or old age deprivation. For reasons of wartime planning, relations between statisticians, academics, and policy makers were strengthened as early as 1914. This cooperation would inspire French planning after the Liberation from Nazis. Claude Gruson, head of French statistical office, wanted to use planning to "program hope." The data and calculations were to make this opaque world "intelligible" and lead to a great democratic debate on the different options. French planning proceeded by forecasting GDP trends and growth rates for the various sectors of the economy by consultations between the state, trade unions, and employers. The aim was to agree on common objectives to reduce uncertainty and direct investment by public and private enterprises toward the priority sectors. In the United States, the economic debacle of the 1930s led to a profound disorganization of society, legitimizing federal systems of social regulation. The centralized administration relied on social science experts, leading to the development of modern statistics, with sample surveys, various measures of inequality, econometrics, and the use of computers from the 1940s onward. Thus, the idea of the "representative sample," the basis of today's surveys, was developed by George Gallup's company. In 1936, it contradicted the forecast of Republican victory, obtained through a survey of two million readers of the magazine The Literary Digest. Using a survey involving 20 times fewer people, but more representative of the whole population than readers of a pro-Republican magazine, Gallup was able to predict Roosevelt's victory.

17.3 Markets to Replace Planning

The first two decades after WWII lead to an improvement of social conditions in many countries. But at the end of the 1960s, several signs of a modernity crisis emerged. First, centralized planning in Communist countries proved unable to move beyond basic industrialization. Authoritarian planning even led to disasters, when governments carried out forced collectivization to bring the "benefits of modern life" to "backward" peasants. Between 1973 and 1976, Tanzania forced five million peasants to settle in large standardized farms inspired by Soviet kolkhozes. This organization was totally unsuited to local agricultural conditions (soil type, family situations, etc.) and turned experienced peasants into unskilled workers, leading to a catastrophic drop in production. These centralized politics assumed that only government officers were able to analyze the situation, determine the reasons for poor results, and find the changes needed to improve them. To borrow an image from James Scott, people became, as in the Taylor factory, the molecules of an organism whose brain is elsewhere, in the state center. Second, at the end of 1960s, as people in rich countries became more educated and had their basic needs satisfied, a widespread request for higher social benefits developed. People demanded autonomy and argued that there is more to life than earning a living. Cooperatives bloomed, betting that capitalist competition of individuals could be replaced by respected individualities within cooperative structures. Finally, people started to realize that modernity's "great acceleration" was polluting the planet and demanded tighter environmental regulations.

In Western countries, these protests lead to a counter-attack from the business community. They piggybacked on the request for autonomy and built on the anti bureaucratic atmosphere to promote the market as society's main regulative body. They argued that society had become too complex to be governed by planning and rigid regulations, which restricted individual freedom and creativity. Instead, the market would allow each person to follow her own objectives, and prices would achieve coordination by disseminating useful information. This is clearly summarized by one of the main thinkers of neoliberalism, Friedrich von Hayek: "Among the members of a Great Society who mostly do not know each other, there will exist no agreement on the relative importance of their respective ends. There would exist not harmony but open conflict of interests if agreement were necessary as to which particular interests should be given preference over others. What makes agreement and peace in such a society possible is that the individuals are not required to agree on ends but only on means which are capable of serving a great variety of purposes and which each hopes will assist him in the pursuit of his own purposes." These "means" were impersonal laws and, above all, markets. For Hayek, market coordination enables businesses and consumers to find the best products and cheapest production methods. Centralized planning attempts to achieve this by calculating a priori the optimal quantities, qualities, and prices for different goods, taking into account the (supposedly known) needs or preferences of individuals. Instead, market rules create a bubble that puts actors at the service of one another, which leads everyone "by the visible gain to himself, to serve needs which to him are invisible." Market rules, especially competition, induce actors to reveal their preferences and private estimates of objects' worth, by transforming these into a number, a purchase or sale price. Prices then disseminate this knowledge, which can be used by other economic actors to make well-founded decisions, in particular thanks to their aggregation in stock markets. Hayek takes the example of a manufacturer who "will release resources for additional production by others by substituting, say, aluminium for magnesium in the production of his output, not because he knows of all the changes in demand and supply which on balance have made aluminium less scarce and magnesium more scarce, but because he learns the one simple fact that the price at which aluminium is offered to him has fallen relatively to the price of magnesium."

From the 1970s onward, markets gradually replaced planning as a means of social coordination. In addition, many states were dispossessed of their powers by world corporations and international organizations such as the OECD or the IMF. This trend has become even more pronounced these last decades, as global private digital platforms are profoundly changing our economy and our social organization. They have replaced retail and oil companies in the top market capitalizations and involve billions of users around the world. Thanks to the data collected in exchange for their services, they already have more information on many countries than their respective governments. Using machine learning algorithms, they are now appropriating some of our skills (detecting friends in photos, driving cars, etc.), which had previously proved too complex to centralize. These are transformed into numbers (the values of the links in neural networks) that can be detached from the context, and be transported in time and space, to be combined in computer centers to trace our behavior.

17.4 Care for Public Numbers

The problem is that this new world governance, based on markets and competition, instead of solving the crisis, has aggravated it. Work is not more fulfilling than before, and the deep inequalities created by unregulated competition have led to a worldwide social and economical crisis. Not to mention the ecological catastrophe that has only deepened and will now profoundly affect human presence on Earth. New institutions are needed. This is of course a complex and deeply political challenge, and I'll only focus here on the role that social numbers could play in two directions: public knowledge and planning.

First, a fundamental role of democratic governments is to take care of procedures that improve public knowledge, including robust social numbers, to take appropriate decisions or assign responsibilities. Today, many feel that grounding politics in official statistical numbers is elitist, just one more way that privileged people in state capitals try to manipulate reality to impose their worldview. But what is the alternative? Leave tabloids and politicians each choose the facts that suit their purposes? For example, a recent study from India's National Statistics Commission reported unemployment at a 45-year high. President Modi, who was elected on a promise to create millions of new jobs, decided to suppress it. In the United Kingdom, the statistics regulator recently complained to the government for repeated inaccuracies in its COVID-19 testing data, which seem to be aimed at showing "the largest possible number of tests." In the United States, the gun lobby makes it very difficult to carry out research or even to collect data on the number of civilians killed by weapons. Yet we know that there are more civilians killed with guns in the last half-century than American soldiers killed in all US wars since the nation's independence! The construction of robust numbers must constantly be defended in the face of pressure from the powerful. Through statistics, these figures can help decision centers to combine many situations and identify relevant causes. This is how we manage to make visible the excess mortality due to pollution, which is so scattered and slow that it is impossible to detect it in each death taken separately. Note that these numbers prevent us from simplifying the discussion but do not dictate the policies to be followed. Moreover, their reliability can only be guaranteed by a pluralist scientific community, capable of questioning the assumptions and calculations made by the "dear colleagues." When predictions are attempted, for example, for epidemics, it is necessary to foster a plurality of models that can test and complement each other. Finally, the people that are transformed into numbers must be allowed to object, to criticize the ways centers want to describe them.

17.5 Planning as a Common Destiny

The absolutely central challenge of this century
seems to be to find forms of articulation
between local ways of living the world
and their integration into institutional systems
that would federate them
without destroying their particularities

Philippe Descola, AOC, 2019

The second role of numbers, collective planning, is more controversial. It seemed that markets had definitely won the battle against planning. However, the ecological mutation highlights two well-known market failures: price reliability and long-term blindness. Whether Hayek likes it or not, the capacity of markets to generate prices that carry reliable information and allow to achieve long-term objectives is doubtful. Prices represent an essential weapon in the economic battle and do not spontaneously take into account essential long-term factors that actors have no interest in "internalizing," such as the depletion of resources or pollution. More generally, the digital revolution has revealed that prices are a very poor characterization of reality, as they compresses complex, multidimensional individual preferences into a single number. Market players are increasingly relying on richer data to coordinate their activities and achieve a better match of production and consumption in fast-changing environments. Thus, Amazon or Walmart adjust their business processes in real time to changing market conditions by processing thousands of times more data on products than the Soviet Gosplan. Digital markets allow both sides of a transaction to specify numerous matchmaking criteria that go far beyond price. BlaBlaCar, for example, the French ride-sharing company, allows passengers to specify the chattiness of their ride, and machine-learning systems can automatically infer our basic preferences through observation and feed matching algorithms using finely detailed criteria.

Public authorities should seize these capacities and put them at the service of common goals through planning. Hayek was right in pointing out that, in the Great Society, it is difficult to achieve agreement on ends. But respecting the strict emission limits fixed by planetary boundaries demands planning and international regulations. For example, China has recently pledged carbon neutrality by 2060, which would represent the largest favorable shock that models have ever produced, lowering the projected temperature increase by 0.2 degree Celsius. Contrary to communist planning, new planning should continuously adapt to priorities that have been decided democratically and include strict limitations on digital technologies because of their enormous ecological impacts, such as energy and rare earths consumption.

Of course, the concrete way to implement this new planning is open for invention, but a wide range of possibilities were already put forward half a century ago. For example, a computer amateur imagined in 1975 a software program centered on the needs and intentions of individuals. It allowed them to manage a personal food database, automatically integrating the purchases made and suggesting others according to planned recipes ("what's missing for the lemon pie?"), or possible recipes according to the food available. This was pure science fiction at the age of punch cards and computers weighing tons, but today such personal data managers could allow people to record their data, organize them, and communicate them to companies or administrations under conditions defined by individuals. These secure data managers would allow us to regain control of our data and, at the same time, multiply their uses: simplified web-based procedures, thanks to a unique identifier, medical records preventing deaths, etc. Stafford Beer, in his prophetical "Designing Freedom" published in 1974, called for renewed planning, which would give to individuals as much autonomy and power as possible. For him, "societary plans should continuously abort, and be recast, before they give birth to a monster. If this is true, there is no need to base them on the predictions that no-one can correctly make in any case, but only on the analysis of an unfolding situation in which every decision constrains [the] future… [Planning] means that the future is something we use our freedom to determine, rather than something that is lurking out there, and will happen to us, unless we are mighty smart. We can make, rather than prophesy, the future." To those that fear control by numbers, (rightly) arguing that digital technology can also favor centralized control, I'd ask not throw out the baby with the bathwater. It's not with numbers that we have a problem, but with politics, with trust in common goals. And here we are, back to planning as common destiny.

Going Further

On the emergence and abuses of centralized States and control : Seeing Like a State by James Scott, Yale University Press, 1998; James Beniger The Control Revolution: Technological and Economic Origins of the Information Society, Harvard University Press (1986)

On gun's statistics : Fight the silencing of gun research, David Hemenway, Nature **546** 345 (2017)

On the counter-attack by business in the 1960–70's, see Grégoire Chamayou, *The Ungovernable Society: A Genealogy of Authoritarian Liberalism*, Polity Press (2021)

On markets and information carried by prices, Joseph Stiglitz, The Contributions of the Economics of Information to Twentieth Century Economics, The Quarterly Journal of Economics (2000) and Richard Bronk, *Hayek on the wisdom of prices: a reassessment* Erasmus Journal for Philosophy and Economics (2013)

On new planning, see Evgeny Morozov – Digital socialism? The Calculation Debate in the Age of Big Data, New Left Review March 2019

On decentralized data management systems, Doc Searls, The Intention Economy, Harvard Business Review Press, 2012

On the vast movement of standardization and "contraction from nature to the communicable", see the classic work by Lorraine Daston: "Objectivity and the escape from perspective", Social Studies of Science, vol. 22, 1992. Also William Cronon's superb book on the history of Chicago, Nature's Metropolis, Norton, 1991.

Chapter 18
Conclusion

Let's look back at this book's main lessons. We've learned that for human affairs, one should not be as ambitious as for physics. Newton's equations allowed to build a realistic virtual scale model of the solar system. The model's ability to predict with high accuracy eclipses and planetary motions was seen as a proof of his deep connection with heavens' workings. Today, physicists manage to connect their simple models to reality in their laboratories, as I showed for cluster deposition (Chap. 5). A useful metaphor for describing the physicists' experimental approach is the taming of a wild tiger, which jumps freely in the jungle, into a tiger jumping reliably through a ring of fire in a circus show. Mathematical models are only relevant for a tamed world. Taming is a better metaphor than "discovering," as it highlights the huge amount of careful, creative work carried out in laboratories, and two further important points. First, not all animals can be tamed, so physicists' approach may not succeed; second, it is important to remember that the tiger often dreams of jumping again freely in the jungle. We should never trust too blindly our technological achievements, as the nuclear accident in Fukushima shows.

The taming metaphor means that, for social systems, the physicist dream could only be achieved by taming humans, to make them reproducible. Only the simplest social situations, where our actions are channeled by social constraints, can get close to such an ambition. For example, epidemics can be modeled with reasonable confidence because their dynamics rely on stabilities at both the biological (contamination processes) and social (work, school, transportation, etc.) levels. When no underlying regularity is at work, models become fragile, as shown by the failure of complex models containing many parameters or relying on brute-force analysis of *big data*. Whatever the model used, social forecast is never much better than the trivial "tomorrow will be as today!" To sum up, we must avoid any mysticism about the power of mathematics. They play a great logistical role, allowing us to go from premises to conclusions by rigorous deduction. But they do not add certainty along the way: results will never be more certain than hypotheses.

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9_18

These technical arguments against the capacity of mathematical models to capture social complexities and evolutions must be completed by political arguments, because addressing social issues "scientifically" does not mean using "rigorous" mathematical models. It means first of all establishing the specificity of social facts, namely, that models deal with humans who are concerned by what models assume and what their results are. Models that pretend to dictate the right policies or understand social reality better than people by using complex mathematics are similar to those authoritarian centralized governments performing social engineering, in which people become, as we saw in the previous chapter, the molecules of an organism whose brain is in the center. Democratic societies need a different use of social numbers, essentially to transfer information and allow coordination and planning. But these social numbers should be controlled by those who are supposedly described by them. For this, one could draw inspiration from the actual practices of physicists in laboratories. Thanks to ingenious experiments, they give the objects they aim to tame the capacity to "object," the ability to disprove their hypotheses and their descriptions. Similarly, one should lend to people the capacity to build their own numbers, to be tamed in the directions they would choose.

Acknowledgments

Trying to bring together expertise from disciplines as varied as physics, sociology, history of science, statistics, or philosophy is the best way to be superficial or even incorrect here and there. Thanks to the interdisciplinarity of ENS de Lyon, I could enjoy many discussions with colleagues from these fields and avoid some naiveties. I would especially like to thank Julien Barrier, Pierre Borgnat, Freddy Bouchet, Sabine Collardey, Philippe Corcuff, Sara Franceschelli, Stéphane Grumbach, Bernard Lahire, Frédéric Le Marcis, Gianluca Manzo, Tommaso Venturini, and Alessandro Vespignani for their availability. I am a bit less ignorant in philosophy, so thanks to exchanges with Didier Debaise, François Jullien, and Isabelle Stengers during the AIME project (ERC Advanced Grant, 2011-2015), led by Bruno Latour.

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*,
https://doi.org/10.1007/978-3-030-65103-9

Annexes

Annex 1: Are Physical Atoms Really Atomic?

As early as antiquity, atoms were given the role of explaining the texture of the world. Lucretia interpreted the flavors of different bodies by the shape of their respective atoms (chapter 11). In most popularization books, atoms are tiny balls with a unique characteristic: their diameter. The density of some materials can indeed be understood by a compact staking of such hard balls. However, depending on the way this length is measured, its value can vary by a factor of five! For example, hydrogen atomic radius varies from 0.025 to 0.120 nanometers. Moreover, this simple image cannot account for the cohesion of matter, which requires attraction between atoms. Newton already knew that additional properties should be attributed to atoms (from the *Principia*'s preface, 1687):

> "From forces of gravity [...], we deduce the motions of the planets, the comets, the moon, and the sea. I wish we could derive the rest of the phenomena of nature by the same kind of reasoning from mechanical principles; for I am induced by many reasons to suspect that they may all depend upon certain forces by which the [atoms], by some causes hitherto unknown, are either mutually impelled towards each other, and cohere in regular figures, or are repelled and recede from each other; which forces being unknown, philosophers have hitherto attempted the search of nature in vain."

Today, atomic forces are derived from experimental data on materials properties. However, these forces are almost as unreliable as atomic radius or the drivers' characteristics inferred in chapter 11. For example, one can infer carbon atomic features from graphite or diamond, which both contain only this element. But as the properties of these two materials are in stark contrast to each other, the two forces turn out to be very different. Atoms' characteristics do depend on the context. Physicists then looked at atomic components, nuclei and electrons, to test whether their characteristics were more stable. And fortunately, their charge and mass do not depend on the context, at least outside nuclear power plants or CERN's particle

P. Jensen, *Your Life in Numbers: Modeling Society Through Data*, https://doi.org/10.1007/978-3-030-65103-9

accelerators. All the problems faced when using social and physical atoms vanish…to give way to other difficulties! For electrons are weird little creatures obeying the complex rules of quantum mechanics, which makes calculations very cumbersome. Even for small molecules containing only ten electrons, computing atomic forces would take several billion times longer than the age of the universe for a modern computer.

So, what are physical atoms used for, *in practice*? Their main interest is their chemical stability that was empirically observed two centuries ago. To understand how this was carried out in practice, let's analyze the example of water. This word already represents a generalization that allows us to link many a priori unique experiences, which show us an object that is sometimes cold, sometimes hot, calm or undulating. For example, common language does not make an explicit link between liquid, vapor, and ice. From this point of view, "water" is an isolated concept and not a transitive one. And this is okay, because in our ordinary practices, the unification brought about by the entity "water" is sufficient. But if we begin to study closely the transformations between liquid water, steam, and ice, we observe that these three substances can transform into each other without residue, which is hardly obvious in the case of steam, as water seems to disappear into the air. Then, we can link these three substances through a single chemical concept, water. Science extends these generalizations by abstraction as far as possible, by connecting, in laboratories, objects perceived by common sense to other objects that are easier to control. For example, the chemical formula H_2O (Fig. 1) connects water to the entire system of chemical elements. Indeed, further experiments showed that when an electric current was passed through it, water really disappeared. It decomposed into two gases, which could be recombined to give back the same amount of water. These gases, named oxygen and hydrogen, were found to be unalterable and were given chemical symbols H and O, leading to the well-known water chemical formula H_2O, after decades of heated controversies. A century of empirical work on many materials revealed a few dozen such stable "elements" which, by construction, can be transferred without alteration during all chemical reactions. For example, the gas obtained by heating mercury oxide to 500 °C happens to have the same properties as one of the two elements (oxygen) obtained by electrolysis of water. The transferability of this empirical invariant makes it possible to connect many compounds which at first sight are very different, such as water and mercury oxide, which is an orange powder. Another crucial point is that the elements obtained at the end of decomposition are identical in nature and mass, whatever the transformations used to obtain them. This allows to attribute an unambiguous chemical formula, a reliable identity card to each substance, and to order the teeming diversity of materials.

The atomic composition of materials can be obtained by analyzing the light they emit, as each atomic species radiates a specific color combination. However, as atoms' forces or characteristics do depend on the context, so does the light emitted, fortunately to a very small extent. Thus, by a very precise measurement of the light emitted by the atoms of a solid, physicists can both recognize the atomic species present (by the color) and also the environment of each species, which slightly shifts

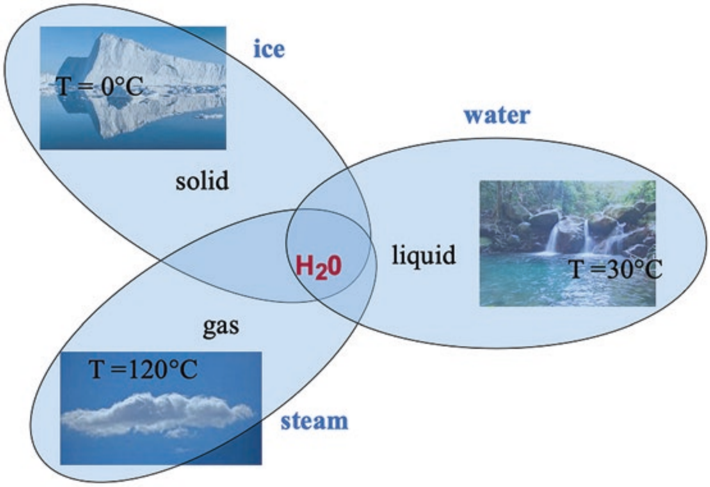


Fig. 1 The substance *H_2O* connects water, ice, and steam by means of a stable, invariant entity hat enriches the description of these three objects and allows for better control

the colors. A bit like recognizing the nationality of a stranger from his language, and then his region of origin, by his specific accent. This allows to precisely study the composition of natural substances, for example, the amount of copper present in an alcoholic beverage distilled in a copper still, and to check whether or not it exceeds the standards. Pure elements are also crucial to allow the production of controlled materials in laboratories. Physicists' atoms are therefore not "atomic," as their characteristics do change with the context, but their chemical stability make them useful for linking many experiments and controlling materials.

Going Further

For the importance of connecting situations through abstractions, see John Dewey, *The quest for certainty* Allen & Unwin Limited (1930)

I discuss in more detail physical atoms in my book (in French) *Des atomes dans mon café crème* (Seuil, 2001).

On the controversies about water chemical formula, see Hasok Chang *Is Water H_2O?* (Springer, 2012)

Annex 2: If We're Not Social Atoms, What Are We?

The basic idea of s*ocial atoms* is that humans have some *fixed* characteristics, from which one can compute their collective behavior, as physicists do with their atoms. As the saying goes, the group, the whole, is more complex than the parts. In chapter 11, I explained why this vision has serious methodological problems, as it works only at the price of simplifying the properties of individuals, the rules of interaction, and the nature of the whole so that they conveniently fit each other. Here, I give an example of a richer conceptualization of humans, considered as *complex* and *evolving* entities. These complex entities are continuously changing through their endless social relationships and may sometimes agree to simplify some of their features for collective action.

French philosopher Gilbert Simondon developed the idea of "individuation" to conceptualize our continuous evolutions. Unlike atoms, which are formed before any encounter, living beings are "like time condensed into bodies." We actively produce our most intimate structures – mental patterns, habits, and dispositions – to solve problems. These structures may prove unsuitable for new situations, which makes necessary to invent new ones, a rare but conceptually dominant event, because it constitutes the basic mechanism of our construction. These problems should not be considered as stimuli imposing an effect but as signals calling for an invention, actively selected by the individual, only if they are compatible with his present structure. However, "activity" does not mean "freedom," a mysterious concept that is not functional, one of those "detestable words which have more value than meaning; which sing more than they speak" as Paul Valéry wittily wrote. For freedom is never freedom to act outside one's own scheme of action, but freedom of inventing new schemes of action which will also have a limited freedom, because the individual structure circumscribes the horizon of possible and impossible encounters. In sum, each of us contains two parts, one already "crystallized," individuated by his biological, psychic, and collective history, and another part that is still "amorphous," plastic, waiting for further crystallizations. The crystallized fraction carries the memory of past individuations and defines the present personality, which we can try to capture, for a while, by social atom features.

This dynamical picture can be combined to the idea of humans as complex entities agreeing to simplify to form a group. Each individual can only extend his structuring by starting from his past crystallizations, which are always particular. As a result, it is always difficult to achieve a coherent crystallization of the amorphous parts of all individuals, which is however the only way to bridge the gap between them all and form a group that will further individualize its members. A simple example shows that this complex vision gives a more interesting description of our collective life. I sing in a small vocal ensemble, Ginga, which has no conductor. In the social atom vision, each of the 12 singers would have some pre-determined characteristics, and their interactions would lead to the whole, the vocal group. Instead, I suggest that each singer agrees to change, to simplify, a tiny fraction of his own complexity, and grants it to the group to create a coherent collective. We all have different backgrounds, disparate musical cultures, and our technical skills are very diverse. No wonder everyone has different ideas of what Ginga is or should be. Should we spend time recording a CD, or rather focus on public performances? Should written scores be respected at all costs, or can they be modified to make them more interesting or simpler to interpret? Clearly, these differences must be reconciled in order to arrive at a viable group and a common interpretation of each

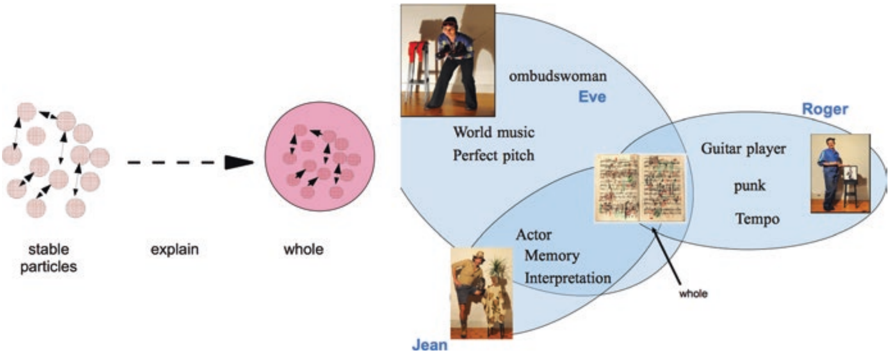


Fig. 2 (Left) In the usual "atomic" modeling, individuals are simple and stable, and their interactions "explain" the whole, which is large as it aggregates the atoms. (Right) In the complex vision, each individual, represented here as an ellipse with several characteristics, is complex. The whole is represented as their intersection, the annotated score, and is smaller than the individuals. There is no guarantee of stability in neither the individuals nor the whole

piece. In this vision then, parts are complex and the whole is simpler, as it represents a provisional agreement. In this example, a concrete trace of the group would be the annotated score, summarizing all the musical choices made collectively, following the more or less animated discussions that allowed us to appropriate the original piece and modify it according to our personalities (Fig. 2). The singers are thus coordinated, partially simplified by this standardized form, which everyone must respect in order to sing together. The whole is much smaller than the parts. But it enriches them, because none of us would have been able, individually, to sing such a score. To go back to Simondon's vision, this collective work crystallized an amorphous part of each of us, in a coherent way.

Going Further

The Philosophy of Simondon, Between Technology and Individuation, Pascal Chabot, Bloomsbury (2013)

The quotation from Paul Valéry: « Regards sur le monde actuel » (1938), in *Œuvres,* Tome II, Paris, Gallimard, 1960, p. 951

The richer conceptualisation of humans is explained in "Fill in the gap. A new alliance for social and natural sciences", *Journal of Artificial Societies and Social Simulation* (2015).

Annex 3: Small but Relevant Data

Finding the "causes" of what happens is at the heart of our control of the world. In physics, controlled laboratory experiments allow to vary one cause at a time and observe its effects. In social systems, many factors vary simultaneously, making it difficult to attribute to each one its share of the effect. In chapter 14, I presented mathematical techniques that allow to isolate the effect of a single cause. Here, I present two experimental ways of linking causes and effects.

Small Data Is Beautiful

Let's start with an elegant example of a controlled experiment, addressing an important question: how can we explain the persistence of corruption in certain countries? The authors' idea is that young people who grow up in a corrupt country tend to cheat, because their socialization takes place in a context that tolerates some rules violations. But how to measure the level of cheating in a country? As cheating remains socially reprehensible in all cultures, observing it seems difficult because people, when they know they are being observed, tend to change their behavior to meet the researchers' expectations. To succeed in detecting average cheating without observing individuals, the authors have come up with an inventive experiment. Participants are isolated in a cubicle and asked to roll a six-sided die twice. Crucially, die rollings were known only to them. They had to report the number obtained in the *first* roll and were paid accordingly: five euros for side "1,″ ten for side "2,″ and so on until the "5.″ If they announced side "6,″ they received nothing. These experiments were carried out in 23 countries, adapting the currency units to the local purchasing power, so that the motivations were identical. And the idea was to check whether there exists a correlation between individual dishonesty in this experiment and an index of corruption in the country, obtained by mixing several indices, such as the rate of tax evasion or the honesty of electoral processes.

The beauty of the experiment is that, although individual cheating is not detectable, the average dishonesty for each country can be perfectly measured. If everyone were honest, an average demand of 2.5 currency units is expected, since all numbers occur with a probability of one-sixth. At the opposite extreme, if all participants pretend to have obtained 5, one can infer that the five-sixths of them have cheated. The actual results are of course between these two extremes. But there is a difference between countries at the top and bottom of the corruption scale. Nottingham's 197 inhabitants (very low corruption index) reported an average of "3." This is a little more than the "2.5" expected for perfect honesty but significantly less than the average "4" reported by the 140 inhabitants of Dar es-Salaam (high corruption index). Half of the latter reported a "5," against a quarter of the English, when the expected figure is one-sixth. Asking for rolling the dice twice allows to interpret the intermediate results observed, as a kind of justified dishonesty, which satisfies the desire to maintain an honest self-image. Lying outright about the result of the roll by reporting "5" when one has obtained two "1" seriously compromises this image. On the other hand, bending the rules a bit, by announcing a relative truth, i.e., not the first throw but the most profitable of the two, can be interpreted as a kind of compromise between total honesty, which would foolishly let an easy win slip through, and the preservation of a good self-image. And the regularity of the dice allows us to conclude that many people follow this intermediate rule, as it reproduces the distribution of the die rolls reported in many countries.

For the authors, these results suggest that an atmosphere of corruption in a country has a strong impact on the “intrinsic” honesty – in the sense of honesty in situations without legal control – of those who are socialized there. However, outright lying remains psychologically costly everywhere. Of course, this elegant experience remains questionable, because it is not obvious that people from such different countries interpret this “game” in the same way, knowing that it was proposed by rich Westerners with whom there existed no social ties and who would not lose money anyway. But it seems to me a good example of the fact that a few thousand results, obtained in well-designed experiments, sometimes make it possible to learn more about our societies than millions of “found data” that have been collected for another purpose.

Randomized Trials

Randomized experiments were first developed in education, medicine, and agriculture, before becoming popular in economics. Suppose one wants to evaluate the effect of reducing class size on educational outcomes. To do so, for example, we would have to reduce the size of a class, evaluate the students, and answer the question: what would their results have been *if* this reduction had not occurred? This impossible task is replaced by a less rigorous (but possible!) approach: take many classes and create two groups, making them as identical as possible, by *randomly* deciding which class goes in each group. The class size reduction is then applied only to one group, and average results of the two groups are compared at the end of the school year. As groups were created randomly, one may hope that the only remaining difference between the groups is class reduction, and we could in principle attribute the divergence in average results to this cause.

However extending this approach to the social sciences poses a large number of methodological and political problems. In short, the method, which is often very expensive, provides interesting lessons for one-off initiatives but cannot guarantee that its results will hold when moving from a controlled trial to widespread application. As we have already explained in chapter 11, there are many reasons for this lack of transferability. The most known is that people tend to behave differently when they know that they participate in an experiment. As an experienced observer notes, this is even more pronounced when a rich “foreign agency comes in with its heavy boots and deep pockets” to carry a test of a single cause. In general, “there tends to be a lot going on other than the [single cause].” Moreover, as we have seen in chapter 14, the interweaving of social mechanisms differs from region to region, in ways that support or block different causal relations and thus render a trial in one setting useless in another. And contrary to the mathematical approach described in the same chapter, randomized trials tell us little about why results happen. Finally, their focus on specific problems tends to obscure important policy issues, such as fiscal, monetary, or industrial strategies.

Going Further

On cheating and corruption, Simon Gächter and Jonathan F. Schulz, “Intrinsic honesty and the prevalence of rule violations across societies”, *Nature,* March 2016

Randomized trials are critically reviewed by Angus Deaton and Nancy Cartwright, in “Understanding and misunderstanding randomized controlled trials”, *Social Science & Medicine* **210**, pp. 2–21 (2018). The quote from the “experienced observer” is taken from that paper.